HANDBOOK OF

Game Theory and Industrial Organization

VOLUME I

Theory

Edited by

Luis C. Corchón • Marco A. Marini

Материал, защищенный авторским правом

HANDBOOK OF GAME THEORY AND INDUSTRIAL ORGANIZATION, VOLUME I

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Handbook of Game Theory and Industrial Organization, Volume I

Theory

Edited by

Luis C. Corchón

Department of Economics, Universidad Carlos III de Madrid, Spain

Marco A. Marini

Department of Social and Economic Sciences, Università di Roma La Sapienza, Italy

Edward Elgar
PUBLISHING

Cheltenham, UK • Northampton, MA, USA

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

© Luis C. Corchón and Marco A. Marini 2018

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical or photocopying, recording, or otherwise without the prior permission of the publisher.

Published by
Edward Elgar Publishing Limited
The Lypiatts
15 Lansdown Road
Cheltenham
Glos GL50 2JA
UK

Edward Elgar Publishing, Inc.
William Pratt House
9 Dewey Court
Northampton
Massachusetts 01060
USA

A catalogue record for this book
is available from the British Library

Library of Congress Control Number: 2017950473

This book is available electronically in the Elgaronline
Economics subject collection
DOI 10.4337/9781785363283

ISBN 978 1 78536 327 6 (cased)
ISBN 978 1 78536 328 3 (eBook)
ISBN 978 1 78811 279 6 (2 volume set)

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Contents

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Contributors

Rabah Amir, Department of Economics, University of Iowa, USA

Andrea Attar, Toulouse School of Economics (CNRS–CRM), France and Università di Roma Tor Vergata, Italy

Gian Italo Bischi, Departments of Economics, Society, Politics, University of Urbino, Italy

Francis Bloch, Université Paris 1 and Paris School of Economics, France

Luis C. Corchón, Universidad Carlos III de Madrid, Spain

Sergio Currarini, University of Leicester, UK, Università di Venezia, Italy

Claude d'Aspremont, CORE, Université Catholique de Louvain, Belgium

Francesco Feri, Royal Holloway, University of London, UK and Università di Trieste, Italy

Jean J. Gabszewicz, CORE, Université Catholique de Louvain, Belgium

Martin Kaae Jensen, School of Economics, University of Surrey, UK

Ludovic A. Julien, EconomiX, Université Paris Nanterre, France

Fabio Lamantia, Department of Economics, Statistics and Finance, University of Calabria, Italy and School of Social Sciences, University of Manchester, UK

Inés Macho-Stadler, Universitat Autònoma de Barcelona and Barcelona GSE, Spain

Marco A. Marini, Department of Social and Economic Sciences, Università di Roma La Sapienza, Italy and CREI, Italy

David Pérez-Castrillo, Universitat Autònoma de Barcelona and Barcelona GSE, Spain

Carlos Pimienta, School of Economics, University of New South Wales, Sydney, Australia

Michele Polo, Bocconi University, IEFE and IGIER, Milan, Italy

Davide Radi, University of Pisa, Department of Economics and Management, Pisa, Italy

Robert A. Ritz, Judge Business School and Faculty of Economics, University of Cambridge, UK

Klaus Ritzberger, Department of Economics, Royal Holloway, University of London and Vienna Graduate School of Finance, Austria

Ornella Tarola, Department of Social and Economic Sciences, Università di Roma La Sapienza, Italy

Jacques-François Thisse, Columbia University, USA, National Research University Higher School of Economics, Russia, CORE-UC Louvain, Belgium, and CEPR

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Amparo Urbano, Department of Economic Analysis and ERI-CES, University of Valencia, Spain

Philip Ushchev, National Research University Higher School of Economics, Russia

Xavier Vives, IESE Business School, Barcelona, Spain

Jingang Zhao, Department of Economics, University of Saskatchewan, Canada

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Foreword

The publication of this *Handbook*, bringing together game theory and industrial organization, is an occasion worth celebrating. After all, industrial organization (IO) – the study of how firms in a given market behave – was game theory's first systematic application to economics, and the success of that application had much to do with giving game-theoretic ideas the prominent place they now have in the economics profession more generally.

Before game theory remade industrial organization in the 1970s, most IO analyses focused on two extreme but simple kinds of markets: perfectly competitive and monopolistic. In a perfectly competitive market, there are many small sellers (all selling the same kind of good) and many small buyers ("small" here means that the quantities sold by a seller and bought by a buyer are tiny compared with the totals for the market). One might guess that large numbers of traders would make analysis complicated, but they actually simplify matters. If each seller is small relative to the market, its own behavior can't affect other sellers appreciably. So when figuring out what it should do, it needn't worry about how the others anticipate it might behave – their anticipations aren't relevant. In other words, a seller doesn't have to be strategic (symmetrically, neither does a buyer). And consequently an economic analyst has a relatively easy job predicting the seller's behavior as the solution to a simple profit-maximization problem. Indeed, because the seller is selling the same sort of good as all its competitors, it will take the good's market price as given (if it chooses a higher price, it will have no customers – since they can get a perfect substitute for less; and it will be overwhelmed by customers if it chooses a lower price). In other words, the seller has no market power.

In a monopolistic market, by contrast, there is just one seller (I shall continue to assume throughout that there are many small buyers). Thus, as with perfect competition, the seller doesn't have to worry about what other sellers are thinking about it – this time because there are no other sellers. And so, again, the seller's optimization exercise as well as the analyst's prediction exercise are quite straightforward (although the seller now does have market power; its own behavior determines the market price).

However, the intermediate case, oligopoly – where there is more than one seller, but not so many that a single seller has no effect on competitors – is more difficult. Think of the American automobile industry as it used to be, consisting primarily of General Motors (GM), Ford, and Chrysler. When GM worked out which models to manufacture, how many units of each model to produce, and what prices to set, it had to take into account what it anticipated Ford and Chrysler would do, and their actions depended on their forecasts about GM. Clearly, grappling with these anticipatory interactions between firms is essential to understanding the automobile industry. Yet such interactions are potentially very complex. Specifically, when an oligopolistic firm A tries to predict what its rival, firm B, will do, it must anticipate what B anticipates A will do, and what B anticipates A anticipates B will do, and so on. That's why Nash equilibrium (Nash, 1950) was such a breakthrough: it cuts through this potentially infinite sequence of mutual anticipations.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

A situation like the automobile industry can be modeled as a game (more precisely, a "non-cooperative" game) in which the firms are players, a rule for how a player behaves constitutes its strategy, and players' strategy choices jointly determine their payoffs. Nash proposed that a good prediction for how players will behave in such a game is that they will choose Nash equilibrium strategies: a configuration of strategies from which no individual player gains by deviating. If each player chooses a strategy to maximize its payoff given its anticipation of others' strategies, then a Nash equilibrium is simply a fixed point of these optimizations. That is, in equilibrium, players' anticipations about other players are correct and thus the infinite sequence of anticipations is circumvented.

Nash equilibrium was, without doubt, the central foundation on which the game-theoretic literature in industrial organization (and, later, other fields in economics) was erected (and it was for this contribution that John Nash shared the 1994 Nobel Memorial Prize in Economics). Nevertheless, Nash's work had at least two important economic precursors.

First, Cournot (1838) and Bertrand (1883) analyzed particular instances of duopoly (an oligopoly in which there are just two firms in the industry) in a game-theoretic way, even though game theory wasn't to be developed formally until the twentieth century. Indeed, both Cournot and Bertrand used what amounted to Nash equilibrium to make their predictions of how firms will behave. Still, remarkable though they are, Cournot's and Bertrand's highly stylized analyses lacked Nash's great generality. Thus, the fact that they had far less influence than Nash (1950) is quite understandable.

The other notable pre-Nash development was monopolistic competition, whose literature was initiated by Chamberlin (1933) and Robinson (1933). Like an oligopoly, a monopolistically competitive market is intermediate between monopoly and perfect competition. And like an oligopolist, a monopolistically competitive firm has market power (normally because the good it sells is not a perfect substitute for other sellers' goods). However, the firm is presumed to be too small to affect its rivals' behavior, and so the strategic interactions of oligopoly are absent.

It may seem surprising that Nash's work, rather than von Neumann and Morgenstern's foundational volume, *Theory of Games and Economic Behavior*, published six years before Nash (1950), had the primary impact on the industrial organization literature. I suspect that von Neumann and Morgenstern (1944) failed to make much of a dent in economics because it is largely devoted to cooperative game theory, which studies games where players can enter into binding coalitions and which normally presumes that the coalition of all players (the grand coalition) forms. This sort of theory is, unfortunately, unsuited to most real-life markets, where typically the grand coalition does not form. Indeed, even if it does arise (the OPEC cartel in the oil market may have been a reasonable approximation of a grand coalition), IO theorists want to understand why this happens and how the coalition sustains itself; they do not usually take the grand coalition for granted, contrary to the presumption of cooperative game theory.

Another surprise is that once Nash's paper appeared, another 20 years passed for game-theoretic work in any economic field – let alone in industrial organization – to take off; there was remarkably little game theory in economics in the 1950s and 1960s (one important exception was Schelling's, 1960, use of game-theoretic ideas to illuminate international relations). Here again, I can only speculate on the reasons, but I conjecture that two important extensions of Nash – Harsanyi's (1968) treatment of games of incomplete information (in particular, his concept of Bayesian equilibrium) and Selten's

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

(1965) treatment of intertemporal games (in particular, his concept of subgame perfect equilibrium) – needed to be understood and digested by economists before they could make good use of game theory in their work.

In any event, the big game-theoretic applications to IO in the 1970s generally involved multiple periods and/or incomplete information; there was a flood of papers on topics such as tacit collusion by oligopolists, market entry by new firms, and limit pricing and predation by incumbent firms, all of which drew heavily on innovations by Harsanyi (1968) and Selten (1965) (who both shared the 1994 Nobel with Nash).

By the early 1980s, game theory had been such a success in industrial organization that it started being used in political economy, international economics, finance, and other areas of economic theory. And at the close of that decade, there was scarcely a self-respecting economics department that didn't offer game theory as an important component of its curriculum.

Industrial organization and game theory together led a revolution in economics. I am truly delighted that there is now a *Handbook* devoted to this transformative partnership.

REFERENCES

Bertrand, J. (1883), "Review of Walras's Théorie Mathématique de la Richesse Sociale and Cournot's Recherches sur les Principes Mathématiques de la Théorie des Richesses", *Journal des Savants*, September, 499–508.

Chamberlin, E. (1933), *The Theory of Monopolistic Competition: A Re-orientation of the Theory of Value*, Cambridge, MA: Harvard University Press.

Cournot, A. (1838), *Recherches sur les Principes Mathématiques de la Théorie des Richesses*, Paris: L. Hachette.

Harsanyi, J. (1968), "Games with Incomplete Information Played by Bayesian Players, I–III", *Management Science*, 14, 159–82, 320–34, 486–502.

Nash, J. (1950), "Equilibrium Points in n-Person Games", *Proceedings of the National Academy of Sciences of the United States of America*, 36, 48–9.

Robinson, J. (1933), *The Economics of Imperfect Competition*, London: Macmillan.

Schelling, T. (1960), *The Strategy of Conflict*, Oxford: Oxford University Press.

Selten, R. (1965), "Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit", *Zeitschrift für die gesamte Staatswissenschaft*, 12, 201–324.

von Neumann, J. and O. Morgenstern (1944), *Theory of Games and Economic Behavior*, Princeton, NJ: Princeton University Press.

Eric Maskin
Adams University Professor
Harvard University

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

1. Introduction

Luis C. Corchón and Marco A. Marini

Game theory lies at the heart of modern industrial organization. Over the second half of the last century, it has provided a sound foundation to the main equilibrium concepts adopted in classical industrial economics, as in the Cournot, Bertrand and Stackelberg models. It has also enabled the development of new and rigorous conceptual frameworks for many industrial organization topics, such as product differentiation, predation, delegation, mergers, collusion and R&D in imperfectly competitive markets. Finally, and perhaps most importantly, over the years, game theory has constantly continued to inspire new research areas in the field of industrial organization, which, in some cases, have gone far beyond the scope of the discipline. This occurred, for instance, in the development of dynamic and incomplete information games or in the recent applications of game theory to law and economics, networks, digital economy, auctions, experiments, health economics, intellectual property rights, contests and corruption, just to cite a few. Furthermore, it should be stressed that the relationship between game theory and industrial organization has never been unidirectional.[1]

Thus, it is no exaggeration to say that game theory has become the common language of industrial organization.[2] In particular, the adoption of a sound mathematical language has allowed industrial organization to steadily progress towards new and unexplored fields. As an example, the recent use of experimental game theory in industrial economics has opened the door to behavioral models for the explanation of the bias of consumers and sellers in the market.[3]

Due to the strong and increasing interlink between game theory and industrial organization, the current volume aims to provide a solid introduction to the main topics lying at the crossroads between these two disciplines. In managing such a – seemingly arduous – task, our major contribution as editors was mainly to attract an impressive array of renowned economists to the challenge of producing up-to-date surveys for the volume. As a final result, and especially thanks to the outstanding quality of the contributors, the current *Handbook* appears suitable for both established researchers as well as for graduate and advanced undergraduate students.

Given the wide heterogeneity of topics being at the boundary between game theory and industrial organization, our primary aim in assembling the book was to give a rational structure to the great amount of material gathered for its preparation. In our final plan for the *Handbook*,

[1] As observed by Bagwell and Wolinsky (2002): "First, the needs of industrial organization fed back and exerted a general influence on the agenda of game theory. Second, specific ideas that grew out of problems in industrial organization gained independent importance as game theoretic topics in their own right. Third, it is mostly through industrial organization that game theory was brought on large scale into economics and achieved its current standing as a fundamental branch of economic theory" (Bagwell and Wolinsky, 2002, p. 1852).

[2] This is confirmed, if ever needed, by Fudenberg and Tirole's long survey contained in the *Handbook of Industrial Economics* (1989), actually encompassing most of the relevant topics in non-cooperative game theory.

[3] See the survey on "Experimental Industrial Organization" by Brandts and Potters in Chapter 17, Volume 2 of this *Handbook*.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

we judged it appropriate to divide the book into two volumes where, while this first volume is mainly devoted to presenting the major game-theoretic modeling tools currently in use in modern industrial organization, the second is specifically aimed at applying those tools to a wide range of industrial organization topics.

More specifically, the first volume is organized into four parts corresponding to four distinct topics: (I) Basic Games in Industrial Organization; (II) Dynamic Games in Industrial Organization; (III) Games of Collusion in Industrial Organization; and (IV) Information Games. Part I of this first volume of the *Handbook* aims at providing an overview of the basic game-theoretic tools currently used in modern industrial organization, such as lattice techniques, aggregative games, monopolistic competition models, oligopoly models with product differentiation, welfare analysis and contest theory. Part II introduces the state of the art in the literature, applying dynamic games to well-known dynamic industrial organization topics such as Stackelberg, entry, and evolutionary games. Part III aims at surveying the main cooperative and non-cooperative games commonly adopted for the analysis of horizontal mergers and collusion. Finally, Part IV provides an overview of some important classes of models dealing with informational issues in imperfectly competitive markets, such as trading under asymmetric information, principal–agent under moral hazard, learning in markets and information sharing in oligopoly. In the next sections we describe in more detail the content of each chapter comprising the *Handbook*.

PART I BASIC GAMES IN INDUSTRIAL ORGANIZATION

In Chapter 2 on strategic complementarities in oligopoly, Xavier Vives provides a detailed overview of many recent results obtained by applying the techniques of supermodular games to the analysis of firm behavior in imperfect competitive markets. Besides offering an excellent introduction to the recent lattice-theoretic methods, the chapter reviews the results obtained in the existence and comparative statics of the Cournot, Bertrand, R&D, advertising, multidimensional and multimarket competition models. In addition, it introduces the use of supermodularity for the analysis of well-known classes of two-stage dynamic games such as entry, dynamic strategic incentives and both Markov and incomplete information games applied to voluntary disclosures and auctions.

In Chapter 3 on Cournot and Bertrand oligopolies and the theory of supermodular games, Rabah Amir surveys two important strands of literature in oligopoly theory, one dealing with the existence of Cournot equilibrium in the general asymmetric and symmetric cases and the other with the effects of exogenous entry on market performance in a Cournot industry. This chapter emphasizes that these two strands of literature share one important unifying common feature: both are achieved via the application of lattice-theoretic methods. This also provides a bridge to the previous chapter.

In Chapter 4 on aggregative games, Martin Kaae Jensen nicely complements the first two chapters by introducing three important classes of widely used aggregative games: (1) linearly aggregative games; (2) generalized aggregative games; and (3) quasi-aggregative games. These games are very useful in industrial organization since they drastically simplify the analysis of the existence, comparative statics and uniqueness of Nash equilibria and unify a vast amount of literature since they apply to a wide array of models like Cournot and Bertrand oligopoly, tournaments, work in teams, contests, patent races and network games.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

In Chapter 5 on monopolistic competition without apology, Jacques-François Thisse and Philip Ushchev review what has been accomplished under the heading of monopolistic competition in industrial organization and in other closely related economic fields. Among other things, the authors argue that monopolistic competition is a market structure in its own right, which encompasses a much broader setup than the celebrated constant elasticity of substitution (CES) model. Also, although oligopolistic and monopolistic competition compete for adherents within the economics profession, the authors explain how such dichotomy is, to a large extent, unwarranted, in that both models are complements rather than substitutes.

In Chapter 6 on oligopoly and product differentiation, Jean J. Gabszewicz and Ornella Tarola overview old and new oligopoly models on product differentiation characterized by local competition. Starting from the microeconomic theory of consumer demand based on characteristics, as introduced by Gorman (1956 and 1980) and then popularized by Lancaster (1966), they present horizontal product differentiation as based on Hotelling (1929) and vertical product differentiation as based on Gabszewicz and Thisse (1979). Finally, they review the model nesting both horizontal and vertical product differentiation and propose two applications of this approach, one based on network externalities and the other on environmental economics.

In Chapter 7 on oligopolistic competition and welfare, Robert A. Ritz nicely reviews the recent developments in the study of social welfare in oligopoly markets. In particular, the chapter covers the usefulness of the rate of cost pass-through for the analysis of market performance and includes a careful analysis of welfare losses due to market power in various widely used models (with symmetric and asymmetric firms, and with or without endogenous entry).

PART II DYNAMIC GAMES IN INDUSTRIAL ORGANIZATION

This second part of the book offers a collection of chapters focusing on the use of dynamic games in a set of well-known industrial organization issues.

Chapter 8 by Klaus Ritzberger on dynamic games sets the scene by introducing three alternative definitions of game trees and extensive forms and also discussing in detail their pros and cons. In the following sections, the author returns to the normal form associated with the extensive form and explains the concept of perfect recall and its significance for economic applications.

In Chapter 9 on strategic refinements, Carlos Pimienta examines the classical literature on equilibrium refinements. Starting with Nash's definition of equilibrium, the chapter presents a comprehensive review of the most successful equilibrium concepts adopted in economic applications as well as the most recent contributions to the subject. The chapter explains in detail how a few decision-theoretic criteria – admissibility, backwards induction, forward induction and invariance – shape the definition of stable sets of equilibria and how they translate into their corresponding mathematical formulation.

In Chapter 10 on Stackelberg games, Ludovic Julien introduces three classes of deterministic non-cooperative Stackelberg games with increasing levels of generality. The first is the basic duopoly game. The second is the oligopoly multiple leader–follower game. The third extends the multiple leader–follower setup within the framework of bilateral oligopoly to describe a multicommodity market. In each case, the author defines and characterizes the

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

equilibrium and the welfare consequences of market power. The chapter also considers the issues of endogenous timing, merging and free entry.

In Chapter on 11 entry games and free entry equilibria, Michele Polo reviews how a wide range of symmetric oligopoly models share some common comparative statics properties. Individual profits and quantities decrease in the number of firms, and tend to competitive or monopolistic competitive equilibria when the number of firms increases indefinitely. The maximum number of firms sustainable in a symmetric long-run equilibrium is shown to depend on technology (economies of scale), preferences (market size) and strategies (toughness of price competition). On the normative side, in homogeneous product markets the business-stealing effect drives the result of excessive entry, whereas adding product differentiation and the utility from variety may reverse this result. In addition, the author considers asymmetric free-entry equilibria using the aggregative nature of many oligopoly models as we noted before. Finally, he discusses the issue of endogenous sunk costs, persistent concentration and frictionless entry in contestable markets.

In Chapter 12 on evolutionary oligopoly games with heterogeneous adaptive players, Gian Italo Bischi, Fabio Lamantia and Davide Radi analyze the properties of evolutionary switching models in oligopoly games, where boundedly rational agents can follow different behavioral rules (or heuristics) to update their production through repeated adaptive decisions. In particular, they focus on well-known heuristics such as best replies with naive expectations, local monopolistic approximation and gradient dynamics on marginal profits. Hence, the chapter examines some specific examples of evolutionary systems where the coexistence of heterogeneous behaviors and of oscillatory time patterns are obtained as possible outcomes.

PART III GAMES OF COLLUSION IN INDUSTRIAL ORGANIZATION

In this part of the volume we gather two theoretical contributions on collusion and mergers in oligopolies. In the first chapter, Chapter 13, on coalitions and networks in oligopolies, Francis Bloch reviews the models on endogenous formation of coalitions and networks in oligopolies. It weaves together a literature in game theory on cooperation and a literature in industrial organization on the formation of groups of oligopolistic firms. The discussion of cooperation in oligopolies starts with a brief presentation of the game-theoretic models used to predict the formation of coalitions and networks. Two different forms of cooperation are considered: (1) cartels and horizontal mergers; and (2) strategic alliances, which encompass both research joint ventures and information exchange platforms.

In Chapter 14 on transferable utility (TU) oligopoly games and industrial cooperation, Jingang Zhao surveys the existing results on TU cooperative games applied to oligopolies and lists nine promising future areas for TU oligopoly games. On the theoretical side, TU oligopoly games are shown to make advances on the refinements and applications of the core, one of the most important solution concepts in cooperative game theory. On the empirical side, the author shows how cooperative games can enable the analysis of industrial cooperation and, hence, the understanding of all forces at work behind industrial changes with and without regulatory policies.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

PART IV INFORMATION GAMES

The final part of the Volume I of the *Handbook* looks at the various effects occurring in markets when relaxing the assumption of symmetric information.

In Chapter 15 on trading under asymmetric information: positive and normative implications, Andrea Attar and Claude d'Aspremont mainly focus their attention on screening models. They divide the chapter into two sections. The first section adopts a simple mechanism design approach with only one mechanism designer. When the mechanism designer is an outsider (say a public authority), all traders may have private information and play simultaneously. When the mechanism designer is an insider (a principal, buyer or seller), he or she is uninformed and has no private information. Three illustrative applications are taken into account: bilateral trade, auctions and insurance. In the second part of the chapter these models are extended to the case of several principals who are uninformed and have no private information but compete by designing mechanisms.

In Chapter 16 on moral hazard: base models and two extensions, Inés Macho-Stadler and David Pérez-Castrillo analyze first the optimal contracts in static moral hazard situations, where the agent's effort is not verifiable. Then, they present the main trade-offs of the principal–agent model. Furthermore, they cover in detail the trade-off of incentives (motivation) vs risk-sharing (efficiency), incentives vs rents (when the agent is protected by limited liability), incentives to a task vs incentives to another (in a multitask situation), and incentives to the agent vs incentives to the principal (when both exert a non-verifiable effort). Finally, they discuss how the predictions of the classical moral hazard model are affected when: (1) there are behavioral biases of individuals; and (2) in presence of a matching market.

In Chapter 17 on learning in markets, Amparo Urbano surveys the problem of market learning as well as that of experimentation (or active market learning) in dynamic models incorporating a Bayesian expectation revision mechanism. Through the lens of this perspective, she reviews the extensive literature on this topic. The experimentation literature has by and large focused on broadly defined bandit models, and thus the starting point is the monopolist experimentation with the classic two-armed bandit problem. The chapter extends to surveying the impact of price competition on experimentation, the role of externalities in social learning and learning in experience good markets.

In Chapter 18 on information sharing in oligopoly, Sergio Currarini and Francesco Feri review the theoretical literature on information sharing in oligopoly and discuss some recent contributions extending the traditional multilateral model to encompass the possibility of bilateral sharing agreements. In the first part of the survey the authors revisit the early insights of the literature, stressing the role of quantity vs price competition and of common vs private values. In the second part, following some more recent contributions, they discuss the bilateral model of information sharing, stressing the role of signals' correlation for the emergence of information sharing in equilibrium and its effect on the architecture of sharing networks. Finally, they conclude the analysis by discussing the emergence of core–periphery networks when firms possess asymmetric information.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

REFERENCES

Bagwell, K. and A. Wolinsky (2002), "Game theory and industrial organization", in R.J. Aumann and S. Hart (eds), *Handbook of Game Theory with Economic Applications, Vol. 3* (1st edition), Amsterdam: Elsevier, pp. 1851–95.

Fudenberg, D.M. and J. Tirole (1989), "Noncooperative game theory for industrial organization: An introduction and overview", in R. Schmalensee and R. Willig (eds), *Handbook of Industrial Organization, Vol. 1*, Amsterdam: Elsevier, pp. 259–327.

Gabszewicz, J.J. and J.F. Thisse (1979), "Price competition, quality and income disparities", *Journal of Economic Theory*, **20**(3), 340–59.

Gorman, W.M. (1956), "The demand for related goods", *Journal Paper J-3129*, Iowa Agricultural Station.

Gorman, W.M. (1980), "A possible procedure for analysing quality differentials in the egg market", *The Review of Economic Studies*, **4**(5), 843–56.

Hotelling, H. (1929), "Stability in competition", *The Economic Journal*, **39**(153), 41–57.

Lancaster, K.J. (1966), "A new approach to consumer theory", *Journal of Political Economy*, **74**(2), 132–57.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

PART I

BASIC GAMES IN INDUSTRIAL ORGANIZATION

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

2. Strategic complementarities in oligopoly

*Xavier Vives**

1 INTRODUCTION

Oligopoly theory is closely connected with game theory. Indeed, oligopoly competition is the leading example of strategic interaction and it should suffice to mention that Cournot's equilibrium concept is just the modern Nash equilibrium. Modeling strategic interaction presents formidable problems as the founders of oligopoly theory (Cournot, Bertrand, Edgeworth, Chamberlin, Robinson and Hotelling) made clear. The oligopoly problem was to establish where prices would settle when market conditions were neither monopoly nor perfect competition. Technical problems in the analysis include lack of quasi-concavity and smoothness of payoffs, indivisibilities, and complex strategy spaces. A Nash equilibrium may not exist, at least in pure strategies. Or, instead, there may be multiple equilibria: How do players coordinate on one of them? How can policy intervention ensure that changing a parameter will have the desired effect? Classical comparative statics analysis provides ambiguous results in the presence of multiple equilibria and imposes strong regularity conditions. These regularity conditions become particularly strong when applied to games with complex functional strategy spaces, such as dynamic or Bayesian games.

Complementarities are intimately linked to multiple equilibria and have a deep connection with strategic situations, and the concept of strategic complementarity is at the center stage of game-theoretic analyses. Examples abound from price games with differentiated products, R&D races, technology adoption, and store and brand location.

Lattice-theoretic methods provide the appropriate toolbox to deal with the problems encountered in oligopoly theory, in particular when complementarities are involved. The theory of monotone comparative statics and supermodular games exploits both order and monotonicity properties (Topkis, 1978, 1979 and further developed and applied to economics by Vives, 1985a, 1990a and Milgrom and Roberts, 1990). By now it has proved useful not only in oligopoly theory but also in all fields of economics from macroeconomics and finance to development and international trade. It continues to be extended at the frontier of research – for example, to dynamic games and games of incomplete information.

Monotone comparative statics analysis provides conditions under which optimal solutions to optimization problems move monotonically with a parameter. This approach exploits order and monotonicity properties, in contrast to classical convex analysis. A central focus of attention are games of strategic complementarities, where the best response of a player to the actions of rivals is increasing in their level, and monotone comparative statics results that allow the extension of the approach to more general games. In this chapter I provide an

* This chapter draws from my previous work on the topic, in particular the book *Oligopoly Pricing* (1999) and the surveys in Vives (2005a, 2005b). I am grateful to project CECO2015-63711-P of the Spanish Ministry of Economy and Competitiveness for support.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

introduction to this methodology and then apply it to the study of strategic interaction in the presence of complementarities in oligopoly games.

The achievements of this approach are as follows. First, it provides a framework for thinking rigorously about complementarities, identifying key parameters in the environment (e.g., what are the critical properties of the payoffs and action spaces?). Second, it simplifies the analysis, clarifying the drivers of the results (e.g., what regularity conditions are really needed to obtain the desired comparative static results?). Third, it encompasses the analysis of multiple equilibria situations by ranking equilibria and helping us to understand how potential equilibria move with the parameters of interest. Finally, it easily incorporates complex strategy spaces, including indivisibilities and functional spaces, such as those arising in dynamic games and games of incomplete information. More specifically, the approach:

- ensures the existence of equilibrium in pure strategies (without requiring quasi-concavity of payoffs or smoothness assumptions) in games of strategic complementarities and beyond;
- allows a global analysis of the equilibrium set, which has an order structure with largest and smallest elements, equilibrium has useful stability properties, and there is an algorithm to compute extremal equilibria;
- permits the use of monotone comparative statics analysis with minimal assumptions by either focusing on extremal equilibria or considering best-response dynamics after the perturbation; and
- extends results beyond the class of games with strategic complementarities.

However, we should be aware also that the lattice-theoretic approach is not a panacea and cannot be applied to everything. Indeed, the approach builds on a set of assumptions.

The chapter provides an introduction to the tools of supermodular games and a range of applications to industrial organization. Section 2 provides an introduction to the theory and basic results. Section 3 provides applications to oligopoly and comparative statics in the context of Cournot, Bertrand, R&D, advertising, multidimensional and multimarket competition. Section 4 deals with dynamic games, studying entry, and characterizing strategic incentives in two-stage and Markov games. Applications to menu and adjustment costs are provided. Section 5 studies games of incomplete information, characterizing equilibria in pure strategies and comparative statics properties, with applications (among others) to games of voluntary disclosure and auctions. The Appendix provides a brief recollection of the most important definitions and results of the lattice-theoretic method.

2 THE FRAMEWORK: SUPERMODULAR GAMES AND MONOTONE COMPARATIVE STATICS

In this section we provide a brief introduction to the tools and main results of the theory of monotone comparative statics and supermodular games. Those tools are based on lattice-theoretic results that exploit order and monotonicity properties of action sets and payoffs. Assumptions are put on strategy sets and payoffs so that best responses are increasing and move monotonically with the parameters under study (following Topkis, 1978). Tarski's (1955) fixed-point theorem delivers the existence of equilibrium and orders properties of

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

the equilibrium set. This section provides the background to the rest of the chapter and the Appendix contains technical definitions and intermediate results.[1]

For simplicity, I provide a definition of a supermodular game in a smooth context. Consider a game $(A_i, \pi_i; i \in N)$, where N is the set of players, $i = 1, \ldots, n$; for player $i \in N$, A_i is the strategy set, a compact cube in Euclidean space, $a_i \in A_i$, and π_i his payoff (defined on the cross-product of the strategy spaces of the players A). Let a_{-i} denote the strategy profile $(a_1, \ldots, a_n)$ excepting the ith element, $a_{-i} \in \prod_{j \neq i} A_j$. Let a_{ih} denote the hth component of the strategy a_i of player i. The game $(A_i, \pi_i; i \in N)$ is *smooth supermodular* if, for all i, $\pi_i\,(a_i, a_{-i})$ is twice continuously differentiable:

1. supermodular in a_i for fixed a_{-i} or $\partial^2\pi_i/\partial a_{ih}\partial a_{ik} \geq 0$ for all $k \neq h$; and
2. with increasing differences in (a_i, a_{-i}) or $\partial^2\pi_i/\partial a_{ih}\partial a_{jk} \geq 0$ for all $j \neq i$ and for all h and k.

The game is smooth *strictly* supermodular if the inequality in (2) is strict. Condition (1) is a complementarity property in own strategies: the marginal payoff to any strategy of player i is increasing in the other strategies of the player. Condition (2) is a strategic complementarity property in rivals' strategies a_{-i}: the marginal payoff to any strategy of player i is increasing in any strategy of any rival player.

In a supermodular game, general strategy spaces can be allowed, including indivisibilities as well as functional strategy spaces, such as those arising in dynamic or Bayesian games (as we will see in Sections 4.3 and 5). Regularity conditions (such as concavity and interior solutions) can be dispensed with.[2]

In a supermodular game each player i has a largest, $\overline{\Psi}_i\,(a_{-i}) = \sup \Psi_i\,(a_{-i})$, and a smallest, $\underline{\Psi}_i\,(a_{-i}) = \inf \Psi_i\,(a_{-i})$, best reply and they are increasing in the strategies of the other players. If the game is strictly supermodular, then any selection from the best-reply correspondence is increasing.[3]

The (weaker) concept of *game of strategic complementarities* (GSC), under our maintained assumptions, is a game where: (a) strategy sets are compact cubes (or "complete lattices"); (b) the best reply of any player has extremal (largest and smallest) elements; and (c) those extremal elements are increasing in the strategies of rivals. Similarly, a game of strict strategic complementarities would have, in addition, that any selection from the best reply of any player is increasing in the strategies of the rivals.[4] The results stated below will hold, replacing (strictly) supermodular game by GSC (game of strict SC).

The following results in Sections 2.1–2.4 hold in a supermodular game. Let $\overline{\Psi} = \left(\overline{\Psi}_1, \ldots, \overline{\Psi}_n\right)$ and $\underline{\Psi} = \left(\underline{\Psi}_1, \ldots, \underline{\Psi}_n\right)$ denote the extremal best-reply maps. Consider as a maintained example a Bertrand oligopoly with differentiated gross substitutable products, with each firm producing a different variety and constant marginal costs.

---

[1] See Chapter 2 of Vives (1999) and Topkis (1998) for a more thorough treatment.

[2] In the general formulation of a supermodular game, strategy spaces need only be "complete lattices," only continuity (not differentiability) of payoffs is needed, and properties (1) and (2) are stated in non-differential terms. A continuity requirement is needed to ensure the existence of best replies. See the Appendix for the general definitions of lattices, supermodularity, increasing differences, and supermodular game.

[3] The basic monotone comparative statics result states that the set of optimizers of a function $u(x, t)$ that is parameterized by t, supermodular in x, and with increasing differences in x and t has a largest and a smallest element and that both are increasing in t. See Lemma 1 in the Appendix for a precise statement of the result.

[4] This definition was used in Vives (1985a). See the Appendix for a more formal definition along those lines.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

2.1 Existence and Characterization of the Equilibrium Set

There always exist extremal equilibria in a supermodular game: a largest equilibrium $\overline{a} = \sup\left\{a \in A : \overline{\Psi}(a) \geq a\right\}$ and a smallest equilibrium $\underline{a} = \inf\left\{a \in A : \underline{\Psi}(a) \leq a\right\}$ of the equilibrium set (Topkis, 1979).[5]

In the Bertrand oligopoly when the payoffs fulfill the complementary conditions (to be discussed in Section 3.1) then it follows that extremal price equilibria do exist.

Symmetric games Consider a symmetric supermodular game (exchangeable against permutations of the players), then:[6]

- The extremal equilibria $\overline{a}$ and $\underline{a}$ are symmetric. Hence, if there is a unique symmetric equilibrium then the equilibrium is unique (since $\overline{a} = \underline{a}$). This result proves useful, for example, to show uniqueness in standard versions of symmetric Bertrand oligopoly models.
- All equilibria are symmetric equilibria if the game is strictly supermodular and the strategy spaces of the players are one-dimensional (or, more generally, completely ordered).

Welfare In a supermodular game, if the payoff to a player has positive spillovers (i.e., it is increasing in the strategies of the other players) then the largest (resp., smallest) equilibrium point is the Pareto best (resp., worst) equilibrium. This is at the basis of finding equilibria that can be Pareto ranked in games with strategic complementarities (Milgrom and Roberts, 1990, Vives, 1990a). For example, in the Bertrand oligopoly case, the profits associated with the largest price equilibrium are also the highest for every firm.

2.2 Stability and Rationalizability

In a supermodular game with continuous payoffs:

1. Simultaneous response best-reply dynamics (Vives, 1990a):

   – approach the "box" $\left[\underline{a}, \overline{a}\right]$ defined by the smallest and the largest equilibrium points of the game;
   – converge monotonically downward (upward) to an equilibrium when starting at any point in the intersection of the upper (lower) contour sets of the largest (smallest) best replies of the players $A^+ \equiv \left\{a \in A : \overline{\Psi}(a) \leq a\right\}$ $(A^- \equiv \left\{a \in A : \underline{\Psi}(a) \geq a\right\})$.

2. The extremal equilibria $\underline{a}$ and $\overline{a}$ correspond to the largest and smallest serially undominated strategies (Milgrom and Roberts, 1990).

[5] The result is shown by applying Tarski's fixed-point theorem (which implies that an increasing function from a compact cube into itself has a largest and a smallest fixed point; see Appendix) to the extremal selections of the best-reply map $\overline{\Psi}$ and $\underline{\Psi}$, which are monotone by the strategic complementarity assumptions. There is no reliance on quasi-concave payoffs and convex strategy sets to deliver convex-valued best replies, as is required when showing existence using Kakutani's fixed-point theorem. Furthermore, the equilibrium set of a supermodular game is a complete lattice (see Vives, 1985a, 1990a, Problem 2.5 in Vives, 1999, and Zhou, 1994).

[6] See Vives (1985a, 1999).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

This result implies that all relevant strategic action is happening in the box $[\underline{a}, \overline{a}]$ defined by the smallest and largest equilibrium points.[7] Results extend to a large class of adaptive dynamics, of which best-reply dynamics are a particular case. A corollary is that if the equilibrium is unique then it is globally stable and dominance solvable. An example is the Bertrand oligopoly market with linear, constant elasticity, or logit demands, where the equilibrium is unique.

2.3 Comparative Statics

Consider an n-player supermodular oligopoly game with payoff for firm i, $\pi_i(a_i, a_{-i}; t)$, parameterized by a vector $t = (t_1, \ldots, t_n)$. If π_i has increasing differences in (a_i, t) (i.e., $\partial^2\pi_i/\partial a_{ih}\partial t_j \geq 0$ for all h and j) then with an increase in t:

(i) the largest and smallest equilibrium points increase; and
(ii) starting from any equilibrium, best-reply dynamics lead to a (weakly) larger equilibrium following the parameter change.[8]

Increasing actions by one player reinforce the desire of all other players to increase their actions, and the increases are mutually reinforcing (i.e., they exhibit positive feedback). We can think in terms of multiplier effects as pointed out in Vives (2005a). Indeed, a multiplier effect in the parameter t_j obtains if the equilibrium reaction of each player to a change in the parameter is strictly larger than the reaction of the player keeping the strategies of the other players constant. This will happen, for example, in a smooth strictly supermodular game with one-dimensional strategy spaces for which $\partial^2\pi_i/\partial a_i\partial t_j \geq 0$ with strict inequality for at least one firm if we consider extremal equilibria (or following best-reply adjustment dynamics after a parameter change).[9]

As an example consider the Bertrand oligopoly fulfilling the complementarity conditions. There, extremal equilibrium price vectors will be increasing in an excise tax t since $\partial^2\pi_i/\partial p_i\partial t > 0$ whenever demand is strictly downward sloping.

Multiplier effects can be related to the *LeChatelier-Samuelson principle* in a strategic environment. This principle states that the response of an agent to a shock will be smaller in the short run than in the long run when other related actions can also be adjusted. Alexandrov and Bedre-Defolie (2017) show, indeed, that the principle holds for extremal equilibria of supermodular games as in the result both for idiosyncratic and common shocks; and in other games under more restrictive conditions for idiosyncratic shocks.[10]

In games with strategic complementarities, we have a multidimensional global version of Samuelson's (1979) *correspondence principle*. This principle links unambiguous comparative statics with stable equilibria and is obtained with standard calculus methods applied to interior and stable one-dimensional models. In GSC, unambiguous monotone comparative statics obtain if we concentrate on stable equilibria.

---

7 For example, rationalizable outcomes (Bernheim, 1984, Pearce, 1984) and supports of mixed-strategy and correlated equilibria must lie in the box $[\underline{a}, \overline{a}]$.

8 The result holds for a class of adaptive dynamics, including fictitious play and gradient dynamics. Furthermore, continuous equilibrium selections that do not increase monotonically with t predict unstable equilibria (Echenique, 2002). The comparative statics result is generalized in Milgrom and Shannon (1994).

9 See Peitz (2000) for sufficient conditions for a price game to display multiplier effects.

10 See also Milgrom and Roberts (1996).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

2.4 Duopoly with Strategic Substitutability

Consider a duopoly ($n = 2$) where there is (a) strategic complementarity in own strategies, with π_i supermodular in a_i or $\partial^2\pi_i/\partial a_{ih}\partial a_{ik} \geq 0$ for all $k \neq h$, and (b) strategic substitutability in rivals' strategies, with π_i with decreasing differences in (a_i, a_j) or $\partial^2\pi_i/\partial a_{ih}\partial a_{jk} \leq 0$ for all $j \neq i$ and for all h and k. Then the transformed game with new strategies $s_1 = a_1$ and $s_2 = -a_2$ is (smooth) supermodular (Vives, 1990a). It follows that all the results stated previously apply to this duopoly game as well. However, the extension to the strategic substitutability case for n players does not apply since the transformation does not work for $n > 2$.

A typical example is a Cournot duopoly with gross substitutes, where typically – but not always – best replies are decreasing. In this case, if for some players payoffs are increasing in the strategies of rivals and for other players they are decreasing, then the largest equilibrium is best for the former and worst for the latter. We have that the preferred equilibrium for a firm is the one in which its output is largest and the output of the rival lowest.

2.5 Extensions to Non-supermodular Games

Totally ordered strategy spaces For totally ordered strategy spaces (e.g., one-dimensional, say a subset of the real line) existence of symmetric equilibrium in n-player games can be obtained, relaxing the monotonicity requirement of best responses (which characterizes supermodular games). The result follows from Tarski's intersection point theorem[11] between a quasi-increasing function and a quasi-decreasing function when they both have the same domains and ranges (which are totally ordered) and the first starts above and ends below the second. A quasi-increasing function cannot have jumps down and, under the assumptions, will necessarily have an intersection with a quasi-decreasing function (which cannot have jumps up). The result can be used to show existence of symmetric Cournot equilibrium since the identity function (the 45^o line) is quasi-decreasing (the theorem is then a fixed-point theorem for quasi-increasing functions). The first result has been successively rediscovered in economics to show existence of equilibrium in a class of symmetric Cournot games starting with the work of McManus (1962, 1964), Roberts and Sonnenschein (1976), and Milgrom and Roberts (1994).

The result can also be used to show existence of equilibrium in two-player asymmetric games by applying the fixed-point theorem to the composition of the best replies assuming one is quasi-increasing and the other continuous and increasing (noting that the composition of the two functions will be itself quasi-increasing). This situation, where one player displays continuous strategic complementarity and the other limited strategic substitutability, is considered by Amir and De Castro (2017).[12]

Aggregative games Results can be extended to aggregative games where the payoff to a player depends on his strategy and an aggregator (typically an additive separable function) of the strategies of all the players. A key tool is the cumulative best reply (or backwards response correspondence) of a player introduced by Selten (1970) and used by Bamon and

---

[11] Tarski (1955). See the Appendix and Section 2.3.1 in Vives (1999).

[12] The results also apply to the dual case where the best reply of one player is continuous and decreasing and of the other quasi-decreasing.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Frayssé (1985) and Novshek (1985) to show existence of a Cournot equilibrium when outputs are strategic substitutes and best-reply correspondences are decreasing (see Theorem 2.7 in Section 2.3.2 of Vives, 1999, Kukushkin, 1994, and Chapter 4 by Jensen in this *Handbook*). For example, consider a symmetric n-player game where for each player the strategy space is a compact interval of the realms and the payoff depends only on his own strategy and the aggregate strategy of the rivals. Then if the best reply of a player has no jumps down, symmetric equilibria exist.[13] In this approach uniqueness of equilibrium is obtained with the requirement that best-reply correspondences (depending on a linear aggregate of the strategies of rivals) have slopes strictly larger than -1 (see Theorem 2.8 in Section 2.3.2 of Vives, 1999). Under smoothness and regularity conditions ("nice aggregative games", see Jensen, Chapter 4 in this *Handbook*, which require concavity assumptions) existence of equilibrium and monotone comparative statics results are obtained without substitutability or complementarity requirements.

Large games The results obtained so far apply also to large games (e.g., non-atomic games with a continuum of players) with some technical caveats. In this case existence of equilibrium can be shown under standard continuity and compactness requirements without requiring quasi-concavity or supermodularity of payoffs because of the convexifying effect of the continuum of players, formulation (see Schmediler, 1973). For example, consider our definition of a GSC and note that it applies to games with an infinite number of players (be it countable or uncountable).

2.6 The Scope of the Theory

If not everything is a game of strategic complementarities, where are the bounds of the theory? If we take the view that the order of the strategy spaces is part of the description of the game or that there is a "natural" order in the strategy spaces, then there are many games that are not of strategic complementarities (as we will see in the next section). In many games, best responses are non-monotone, e.g., they are increasing in a part of the strategy space and decreasing in another. However, if we take the view that the order of the strategy sets of the players is a modeling choice at the convenience of the researcher (and this is what we have done to extend the reach of the theory to duopolies with strategic substitutes) the answer may be different. In fact, if we allow the construction of this order ex post, with knowledge of the equilibria of the game, then most games *are* of strategic complementarities. To put it another way, complementarities alone, in the weak stated sense, do not have much predictive power unless coupled with additional structure (Echenique, 2004a). However, this procedure, with a priori knowledge of the equilibria and the defined order, may not be "natural."

13 The argument is simple. Let $\pi_i\left(a_i, a_{-i}\right) = \pi\left(a_i, \sum_{j\neq i} a_j\right)$ for any i, as in a Cournot game with homogeneous product and identical cost functions. Existence of symmetric equilibria follows then from the stated result if the best-reply Ψ_i of a player (identical for all i due to symmetry) has no jumps down. This is in fact true if costs are convex in the Cournot game. Symmetric equilibria are given by the intersection of the graph of $a_i = \Psi_i\left(\sum_{j\neq i}\right) a_j$ with the line $a_i = \left(\sum_{j\neq i} a_j\right) / (n-1)$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

3 STATIC OLIGOPOLY GAMES AND COMPARATIVE STATICS

This section provides a brief review of some of the basic applications to static models of oligopoly competition. It surveys Cournot and Bertrand markets, including comparative statics results, patent races, and multidimensional competition (including extensions of the methods to games that do not display global complementarities). The analysis highlights the power and applicability of the approach.

3.1 Cournot and Bertrand Markets

Leading oligopoly models fit, in natural but not universal specifications, the assumptions made in supermodular games. This is the case for a Cournot oligopoly with complementary products. In this case, the strategy sets are compact intervals of quantities and the complementarity assumptions are natural. A second case is a Bertrand oligopoly with differentiated substitutable products, with each firm producing a different variety. The demand for variety i is given by $D_i(p_i, p_{-i})$, where p_i is the price of firm i and p_{-i} denotes the vector of the prices charged by the other firms. A linear demand system with gross substitutes will satisfy the complementarity assumptions.

Considering increasing transformations of the payoffs the application of the theory is extended (since this operation does not change the equilibrium set of the game). The game is *log-supermodular* if $\pi_i \geq 0$ and if $\log \pi_i$ fulfills the complementarity conditions (1) and (2) of Section 2. In the Bertrand oligopoly example, $\pi_i = (p_i - c_i) D_i(p_i, p_{-i})$, where c_i is the constant marginal cost, is log-supermodular in (p_i, p_{-i}) whenever $\partial^2 \log D_i / \partial p_i \partial p_j \geq 0$. This holds when the own-price elasticity of demand η_i is decreasing in p_{-i}, as with constant elasticity, logit, or constant expenditure demand systems.[14] However, not all Bertrand games with product differentiation are supermodular games. Examples include games with payoffs that are not single-peaked as well as with avoidable fixed costs, and the Hotelling model where firms are located close to each other. In those cases, at some point best replies jump down and a price equilibrium (in pure strategies) fails to exist.[15] Even with goods that are gross substitutes, prices may not be strategic complements since the own-price elasticity of demand need not decrease in the prices charged by rivals.[16] An instance where strategic price substitutability among prices may arise is in the presence of strong network externalities.[17] Furthermore, even a linear Bertrand oligopoly game with continuous best replies and more than two firms need not be supermodular or satisfy single-crossing conditions when demands have kinks and some firms may not produce.[18]

---

14 See Chapter 6 of Vives (1999).

15 See Roberts and Sonnenschein (1977), Friedman (1983), and Vives (1999, Sec. 6.2). However, in the modified Hotelling game in Thisse and Vives (1992) best responses may be discontinuous but are increasing.

16 A price increase by rival j may lead to an *increase* in the own-price elasticity of demand for firm i because it makes consumers of brand i who do not have a strong preference for any product – that is, who are more price sensitive – more willing to switch brands (see Berry, Levinsohn, and Pakes, 1999 for some empirical support for this effect).

17 For example, in the logit model with network externalities (Anderson, De Palma, and Thisse, 1992, Ch. 7), increasing the price set by a rival raises the value for consumers of the network of firm i, so it may pay this firm to cut prices in order to increase this lead if network externalities are large enough.

18 See Cumbul and Virág (2014).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

The lattice-theoretic methods can be further extended to non-supermodular price games. Consider, for example, a Bertrand duopoly with differentiated gross substitute products where firm 1 has concave costs (increasing returns) and supermodularity fails with competition being of the "strategic quasi-substitutability" type (that is, demands are such that the best reply of firm 2 is continuous and decreasing, and that of firm 1 is quasi-decreasing). In this case an equilibrium exists using Tarski's intersection point theorem (see Appendix). Another example is provided by the mixed price–quantity duopoly of Singh and Vives (1984) where firm 1 is a price setter and firm 2 a quantity setter. Assume constant marginal costs and suppose that demands are such that the payoff of firm 1 in quantities is submodular and the payoff of firm 2 in prices is supermodular and quasi-concave in own price. Then the mixed duopoly displays "strategic quasi-complementarity" (with the best reply of firm 1 quasi-increasing and the best reply of firm 2 continuous and increasing), and an equilibrium exists (generalizing the results of Singh and Vives, 1984 with continuous best replies).[19]

3.1.1 Comparison of Cournot and Bertrand equilibria

Consider the n-firm Bertrand oligopoly case with firm i producing q_i of variety i at cost $C_i(q_i)$. In the Bertrand game firms compete in prices and $\pi_i = p_i D_i(p_i, p_{-i}) - C_i(D_i(p_i, p_{-i}))$. If firms compete in quantities in the same market then profits for firm i are given by $P_i(q_i, q_{-i})\, q_i - C_i(q_i)$, where $P_i(q_i, q_{-i})$ is the inverse demand for firm i. The lattice-theoretical approach makes precise in what sense Bertrand equilibria are more competitive than Cournot equilibria and what drives the result. With gross substitute, or complementary products, if the price game is supermodular and quasi-concave (that is, π_i is quasi-concave in p_i for all i) then at any interior Cournot equilibrium prices are higher than the smallest Bertrand equilibrium price vector. A dual result holds also. With gross substitute, or complementary, products, if the quantity game is supermodular and quasi-concave, then at any interior Bertrand equilibrium outputs are higher than the smallest Cournot equilibrium quantity vector (Vives, 1985b, 1990a).

3.1.2 Comparative statics in Cournot markets

The standard Cournot game displays strategic substitutability and, therefore, the game is supermodular only in the duopoly case (by changing the sign of the strategy space of one player), as discussed in Section 2.4. We can also obtain results with n firms with the lattice-theoretic approach even if the game is not supermodular. The standard approach (Dixit, 1986) assumes quasi-concavity of payoffs, downward-sloping best replies, and that the equilibrium analyzed is unique and stable to derive comparative statics results. The classical approach has several problems. First of all, it requires unnecessary regularity conditions to deliver results. Second, it is silent when payoffs are not quasi-concave. Third, it is problematic for comparative static analysis when there are multiple equilibria. For example, if the uniqueness condition for symmetric equilibria does not hold and there are multiple symmetric equilibria, changing n may either cause the equilibrium considered to disappear or introduce more equilibria.[20]

---

[19] See Section 2.5 and Amir and De Castro (2017).

[20] See Amir (1996a), Vives (1999), Amir and Lambson (2000) and Chapter 3 by Amir in this Handbook for the results.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

3.1.3 Monopolistic competition

Monopolistic competition is characterized by firms that are negligible with respect to the overall market but still retain market power on the differentiated product supplied. The appropriate modeling of such a situation is with a continuum of firms each supplying a product. Whenever the complementarity assumptions are fulfilled the price game will be supermodular and the results of Section 2 will apply (and, indeed, even with heterogeneous firms, e.g., Yang and Qi, 2014). It is worth noting that in monopolistic competition, and with no uncertainty, Cournot and Bertrand equilibria deliver the same outcome (that is, quantity or price competition are equivalent; see Section 6.6 in Vives, 1999).[21]

3.2 Patent Races

Consider an n-firm oligopoly engaged in a memoryless patent race. All firms have access to the same R&D technology. An innovating firm obtains prize V and losers obtain nothing. If a firm spends x continuously then the (instantaneous) probability of innovating is given by $h(x)$ (where h is a smooth concave function with $h(0) = 0$, and $h' > 0$, $\lim_{x\to\infty} h'(x) = 0$, $h'(0) = \infty$, a region of increasing returns for small x may be allowed). With no innovation, the normalized profit of firms is null. We have then that the expected discounted profits (at rate r) of firm i investing x_i if rival j invests x_j is

$$\pi_i = \frac{h(x_i)\,V - x_i}{h(x_i) + \sum_{j\neq i} h\left(x_j\right) + r}.$$

The best response of a firm by $x_i = R\left(\sum_{j\neq i} h(x_j) + r\right)$ is well defined under the assumptions. Restricting attention to symmetric Nash equilibria of the game, under a stability condition at a symmetric equilibrium x^*, $R'((n-1)\,h(x^*))\,(n-1)\,h'(x^*) < 1$, x^* increases with n (Lee and Wilde, 1980). However, this approach requires assumptions to ensure a unique and stable symmetric equilibrium and cannot rule out the existence of asymmetric equilibria. Alternatively, the following mild assumptions, $h(0) = 0$ and h is strictly increasing in $[0, \overline{x}]$, with $h(x)\,V - x < 0$ for $x \geq \overline{x} > 0$, ensure that the game is strictly log-supermodular. It follows then from Section 2.1 that equilibria exist and are symmetric. It follows that at extremal equilibria the expenditure intensity x^* is increasing in n (strictly if h is smooth with $h' > 0$ and $h'(0) = \infty$). Furthermore, starting at any equilibrium, an increase in n will raise the research intensity, with out-of-equilibrium adjustment according to best-reply dynamics. This will be so even if some equilibria disappear or new ones appear as a result of increasing n.

3.3 Multidimensional Competition

The lattice-theoretic approach can readily handle multidimensional strategy spaces. I consider Cournot competition with cost reduction, advertising and pricing, and multimarket oligopoly.

[21] See Thisse and Uschev, Chapter 5 in this *Handbook* for a survey of monopolistic competition models. Vives (1985b) and Gabaix, Laibson, and Li (2016) provide approximations of margins in large monopolistically competitive markets.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

3.3.1 Cournot competition with cost reduction

Consider an n-firm Cournot market for a homogeneous product with smooth inverse demand $P(\cdot)$, $P' < 0$. Firm i can invest z_i to reduce its constant marginal cost of production c_i according to a smooth function $c_i = c(z_i)$ with $c(z) > 0$, $c'(z) < 0$, and $c''(z) > 0$ for all $z > 0$. The profit to firm i is given by

$$\pi_i = P(Q)\, q_i - c(z_i)\, q_i - z_i,$$

where q_i is the output of the firm and Q is total output. Firms simultaneously choose output and cost-reduction effort. Using lattice-theoretic methods we do not need to invoke regularity conditions to obtain the existence of equilibrium and comparative statics results on the number of firms, as long as we restrict attention to extremal equilibria. Under the assumptions plus some mild boundary conditions interior extremal equilibria (q^*, z^*) exist and q^* and z^* are strictly decreasing (increasing) in n if Cournot best replies are strictly decreasing (increasing) (Vives, 2008a).

3.3.2 Advertising, prices, and quantities

I examine complementarities between advertising and other strategic variables, considering first a price game and then a quantity and cost-reduction game, both with differentiated products.

In the price game the demand of firm i $D_i(p; t_i)$ increases on advertising effort t_i, $\partial D_i / \partial t_i > 0$ with cost $F_i(t_i)$ and $F_i' > 0$, so that $\pi_i = (p_i - c_i)\, D_i(p; t_i) - F_i(t_i)$. Suppose that goods are gross substitutes, $\partial D_i / \partial p_j \geq 0$ for $j \neq i$, and that demand is downward sloping, $\partial D_i / \partial p_i < 0$. The action of the firm is $a_i = (p_i, t_i)$, lying in a compact rectangle. A sufficient condition for π_i to be strictly supermodular in a_i is that $\partial^2 D_i / \partial p_i \partial t_i \geq 0$ since

$$\frac{\partial^2 \pi_i}{\partial p_i \partial t_i} = (p_i - c_i)\, \frac{\partial^2 D_i}{\partial p_i \partial t_i} + \frac{\partial D_i}{\partial t_i} > 0.$$

The condition requires advertising to increase customers' willingness to pay. If $\partial^2 D_i / \partial p_i \partial p_j \geq 0$ for $j \neq i$ (noting that $\partial D_i / \partial p_i \partial t_j = 0$, $j \neq i$), π_i has increasing differences in $((p_i, t_i), (p_{-i}, t_{-i}))$. Under these assumptions, the game is supermodular and the largest (smallest) equilibrium displays high (low) prices and high (low) advertising levels. In a symmetric model and with a linear demand system, multiple equilibria obtain, with t_i increasing the demand intercept if F' is concave enough. Under these conditions high advertising levels are associated with high prices.

Immordino (2009) considers a Cournot oligopoly with differentiated product launching. Consumers become aware of the products via advertising and firms decide simultaneously on production, advertising expenditure and cost-reducing investment. It is assumed that consumers with higher willingness to pay are more likely to be receptive to advertising; that marginal consumer awareness is increasing in advertising effort; that the consumer awareness of product i decreases in the intensity of advertising of firm j, and the marginal effectiveness of the advertising of firm i is decreasing in the advertising effort of firm j. Furthermore, the marginal cost of advertising is decreasing by using more specialized media. It is shown that in a strategic substitutes duopoly where firm i benefits from the improvements in advertising technology, but not firm j, all strategic variables at extremal equilibria increase in targeted

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

advertising for firm i and decrease for firm j. In an oligopoly with complementary products (and where consumer awareness of product i increases in the intensity of advertising of firm j, and the marginal effectiveness of the advertising of firm i is increasing in the advertising effort of firm j) and strategic complements, all variables at extremal equilibria increase with a move towards targeted advertising.

3.3.3 Multimarket oligopoly

The approach allows the study of multiproduct firms and even of price games that are neither supermodular nor log-supermodular. I provide three applications: multimarket oligopoly pricing, two-sided markets, and the pricing of components.

Multimarket oligopoly In a standard multiproduct logit oligopoly pricing model best responses are increasing and there is a unique Bertrand equilibrium despite the fact that payoffs are single-peaked (not quasi-concave) and neither supermodular nor log-supermodular in own actions or prices (Spady, 1984). However, strategic complementarity across prices of different firms holds. A similar, and more general, result is obtained by Nocke and Schutz (2015) who, using a discrete/continuous choice framework with independent and identically distributed (IID) type 1 extreme-value taste shocks, introduce a class of demand systems for multiproduct firms that nests the cases of multinomial logit and constant elasticity of substitution (CES).[22] The demand for product $k \in N$, where N is the set of differentiated products, is given by $D_k(p) = \frac{-h_i'(p_i)}{\sum_{j \in N} h_j(p_j)}$ where $p \in R_{++}^N$, $h_j' < 0$ and h_j is log convex. The set of firms is a partition N. Suppose that firm i produces goods in the set N_i. With constant (and positive) marginal costs this defines an aggregative pricing game since the profit of firm i depends only on $(p_k)_{k \in N_i}$ and on $H \equiv \sum_{j \in N} h_j\left(p_j\right)$. Under the assumption that the relative degree of convexity of h_j is non-decreasing in price[23] a Bertrand equilibrium exists (and it is unique under stronger conditions). The result is obtained even though profits are not quasi-concave in own prices, but they are single-peaked, and the price game is not supermodular. Monotone comparative static results can be derived at extremal equilibria (with largest and smallest H). For example, with an increasing outside option H^0 at extremal equilibria, profits and prices of all firms decrease and consumer surplus increases (with expansion of the set of products sold). With CES demands, an algorithm to compute the price equilibrium with multiproduct firms is provided. The results allow the characterization of the dynamic optimality of myopic merger policy.

A multimarket mixed oligopoly featuring products demand complements within the firm and substitutes across firms provides another example. This situation obtains in two-sided markets, where two groups of market participants benefit from interaction via a platform or intermediary, or when final products are combinations of components.

---

22 The authors also show that these demand systems are integrable with quasi-linear preferences.

23 In the monopolistic competition case where H is taken as given this corresponds to a non-decreasing perceived price elasticity of demand.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Two-sided exclusive intermediation Consider two groups of participants in platforms[24] where each participant joins one of the two existing intermediaries. The utility derived by a member of a group from joining a particular intermediary is increasing in the number of members of the other group joining the same intermediary. With linear demands arising from Hotelling-type preferences for the intermediaries, we have that prices charged by intermediaries are strategic complements across firms but strategic substitutes within the firm. The multimarket oligopoly game is therefore not a supermodular game. However, best replies will be increasing as long as the demand complementarity among the products of the same platform/intermediary is not very strong. With linear demands and small and symmetric network effects, best replies are increasing and there is a unique symmetric equilibrium.

Pricing of components Consider now a situation where each of a finite set of end products uses one or more components and where no two products share a component. Each component is produced by a separate monopolist who sets its price, and the price of a product is the sum of the prices of its components. The price game is not supermodular since the prices of the different components are strategic substitutes. However, Quint (2014) provides conditions on the distribution of consumer valuations for a discrete-choice demand system to yield demand for each product that is log-concave in price, and has log-increasing differences in own and another product's price. This leads to the consideration of an auxiliary game in product prices, with an equilibrium directly linked to the equilibrium of the pricing components game, which is supermodular and from which we can derive comparative static properties in terms of costs, qualities and entry of new products, as well as derive the effects of mergers between firms. The results apply to retail competition, licensing of intellectual property, and patent pools.

4 DYNAMIC GAMES

This section examines dynamic games, building on the stated comparative statics results. I examine entry, a generalization of the taxonomy of strategic behavior of Fudenberg and Tirole (1985), conditions under which increasing or decreasing dominance occurs in oligopoly, and Markov games and Markov perfect equilibria (MPE). I characterize conditions for dynamic strategic complementarity and the link between static and dynamic complementarities, and the existence of MPE.

4.1 Entry

Consider a two-stage game where first firms decide whether to enter or not in the market, paying an entry cost, and then compete in quantities, and study subgame-perfect equilibria. Amir, De Castro, and Koutsougeras (2014) extend the Mankiw and Whinston (1986) excess

[24] Examples are numerous and include readers/viewers and advertisers in media markets, cardholders/consumers and merchants/retailers in payment systems such as credit cards, consumers and shops in shopping malls, authors and readers in academic journals, borrowers and depositors in banking, "subscription to a network" and "number of calls made to a network" in telecom markets, and in general buyers and sellers put together with the help of intermediaries (in real estate, financial products, or auction markets). The interaction between the two sides gives rise to complementarities or externalities between groups that are not internalized by end users. See Armstrong (2006) for a survey of two-sided competition.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

entry results in symmetric Cournot oligopoly with free entry to allow for limited increasing returns to scale using lattice-theoretic methods. The authors assume that inverse demand is downward sloping ($P' < 0$) and costs are strictly increasing ($C' > 0$), and both smooth, with $-P' + C'' > 0$, and price is below average cost for high enough outputs. Under the assumptions, Cournot extremal equilibria exist for any n and the authors show that there is excessive entry in the sense that at most there is one firm too few in the market solution (with respect to the structural second best where the number of firms is decided by the planner) whenever there is business stealing. This is always the case when outputs are globally strategic substitutes. When $-P' + C'' < 0$ then only one firm should enter but the market solution will allow entry (weakly) more and with no entry cost it would allow entry of an infinite number of firms.

Anderson, Erkal, and Piccinin (2016) study free entry in aggregative oligopoly games with potentially asymmetric firms (including potential asymmetric entry costs).[25] They make the observation that a Bertrand pricing game is aggregative if demands satisfy the independence of irrelevant alternatives (IIA) property, e.g., CES or logit (the converse assertion is not true). The authors derive neutrality results (where the aggregate stays the same) across market structures and the corresponding policy implications for merger analysis.

Mrazova and Neary (2016) study selection effects with *heterogeneous* firms in the decision of whether and how to enter a market using lattice-theoretic methods. They find that "first-order" selection effects (in terms of firms entering or not) are very robust while "second-order" effects (in terms of the entry mode, exporting or foreign direct investment [FDI], conditional on entry) are much less so. More efficient firms select the entry mode with lower market-access costs if firms' profits are supermodular in production and market-access costs but need not do so otherwise. The authors derive microfoundations for supermodularity to hold in a range of standard models and show how supermodularity may fail with FDI when demands are less convex than CES, with fixed costs increasing with productivity, and with threshold effects in R&D (that is, when the average cost function is first convex and then concave in investment).

4.2 Taxonomy of Strategic Behavior

The taxonomy of strategic behavior provided by Fudenberg and Tirole (1984) in the context of a two-stage game between an incumbent (firm 1) and an entrant (firm 2) illustrates the use of the approach. At the first stage the incumbent can make an observable investment t. The incumbent can influence the market outcome at the second stage by affecting the equilibrium behavior of the rival at the second stage. At the (market) stage payoffs are, respectively, $\pi_1(a_1, a_2; t)$ and $\pi_2(a_1, a_2)$ where a_i is the market action of firm i. We want to sign the strategic effect, taking as benchmark behavior where the incumbent when deciding about t only takes into account the direct effect of the investment on his payoff. This corresponds to the open-loop equilibrium of the two-stage game, which is equivalent to the game with simultaneous choice by the incumbent of t and a_1.

The standard approach assumes that at the second stage there are well-defined best-response functions for both firms, and that there is a unique and (locally) stable Nash equilibrium that depends smoothly on t, $a^*(t)$. A taxonomy of strategic behavior (see Table 2.1) can be

[25] See Corchón (1994) for an early analysis and Chapter 11 by Polo in this *Handbook*.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Table 2.1 Taxonomy of strategic behavior

	Investment makes player 1:	
	Tough	Soft
Strategic Substitutes	Overinvest (top dog)	Underinvest (lean and hungry)
Strategic Complements	Underinvest (puppy dog)	Overinvest (fat cat)

provided depending on whether competition is of the strategic substitutes $\left(\frac{\partial^2\pi_2}{\partial a_1\partial a_2} < 0\right)$ or complements $\left(\frac{\partial^2\pi_2}{\partial a_1\partial a_2} > 0\right)$ variety and on whether investment makes firm 1 soft $\left(\frac{\partial\pi_1}{\partial a_2}\frac{\partial^2\pi_1}{\partial t\partial a_1} > 0\right)$ or tough $\left(\frac{\partial\pi_1}{\partial a_2}\frac{\partial^2\pi_1}{\partial t\partial a_1} < 0\right)$.[26] The top dog strategy obtains if competition is of the strategic substitutes type and investment makes firm 1 tough, then the incumbent wants to overinvest to push the entrant down his best-response curve. Cournot competition and investment in cost reduction are an example. The puppy dog strategy obtains if competition is of the strategic complements type and investment makes firm 1 tough, then the incumbent wants to underinvest to move the entrant up his best-response curve. Price competition with differentiated products and investment in cost reduction provide an example. We can define similarly the strategies "lean and hungry" and "fat cat".

The taxonomy follows from minimal assumptions, the character of competition and investment, as applied to extremal equilibria in the lattice-theoretic version of the result (Section 7.4.3, Vives, 1999) with no need to impose strong restrictions to obtain a unique and stable equilibrium at the market stage. Indeed, if the market game is supermodular and $\partial^2\pi_1/\partial a_1\partial t \geq 0$ then extremal equilibria are increasing in t. If the game is of strategic substitutes (submodular) then extremal duopoly equilibrium strategies for firm 1(2) are increasing(decreasing) in t if $\partial^2\pi_1/\partial a_1\partial t \geq 0$. The results are reversed if $\partial^2\pi_1/\partial a_1\partial t \leq 0$. The taxonomy follows for extremal equilibria: $sign\frac{\partial a_2^*}{\partial t} = sign\left(\frac{\partial^2\pi_2}{\partial a_1\partial a_2}\frac{\partial^2\pi_1}{\partial t\partial a_1}\right)$ when a_2^* is an extremal equilibrium. What if at the market stage firms are sitting on a non-extremal equilibrium? Then if out of equilibrium adjustment is governed by best-reply dynamics the sign of the impact of a change in t is the same as with an extremal equilibrium.

4.3 Markov Games

We explore in this section dynamic complementarities and their relation to static ones in discrete-time Markov games. A Markov strategy depends only on state variables that condense the direct effect of the past on the current payoff. Denote by $\pi_i(x, y)$ the current payoff of player i, where x is the current action profile vector and y is the state that evolves according to the law of motion $y = f(x^-, y^-)$, with x^- and y^- (respectively) the lagged action profile vector and the lagged state. A Markov perfect equilibrium (MPE) is a subgame-perfect equilibrium in Markov strategies. That is, an MPE is a set of strategies optimal for any firm, and for any state of system, given the strategies of rivals.

[26] Indeed, if $\frac{\partial\pi_i}{\partial a_j} < 0, j \neq i$, an increase in the market action of firm j hurts firm i, and if $\frac{\partial^2\pi_1}{\partial t\partial a_1} > 0$ an increase in t will shift the best-response function of firm 1 out and this will represent an aggressive move.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Let us speak of "contemporaneous" strategic complementarity (SC) when the value function at an MPE $V_i(y)$ displays SC (V_i has increasing differences in (y_i, y_{-i})). "Intertemporal" SC obtains when a player raising her state variable today increases the state variable of her rival tomorrow. "Intertemporal" strategic substitutability (SS) obtains when a player raising her state variable today decreases the state variable of her rival tomorrow. I restrict attention to a class of simple dynamic Markov games that admits two-stage games, simultaneous move games with adjustment costs, and alternating moves games. Consider an n-player game in which the actions of player i in any period lie in A_i, a compact cube of Euclidean space; $\pi_i(x, y)$ is the current payoff for player i, continuous in both $y \in A$, the action profile in the previous period (state variables) and in $x \in A$, the current action profile.

I take in turn the issues of contemporaneous SC in two-stage games and intertemporal SC or SS in infinite-horizon games. I end the section with results on the existence of MPE in stochastic games.

4.3.1 Two-stage games

Let $y \in A$ be the action profile in the first stage and $x \in A$ the action profile in the second stage. The contemporaneous SC property obtains under two conditions: (a) if at the second stage, for any actions y in the first stage, payoffs $\pi_i(x, y)$ display SC and (b) if the SC property is preserved when payoffs are folded back at the first stage in a subgame-perfect equilibrium. Suppose that $\pi_i(x, y)$ displays increasing differences (or is supermodular) in any pair of variables. Let $x^*(y)$ be an extremal equilibrium in the second stage (they exist at the second stage for any y because the second-stage game is supermodular). $V_i(y) \equiv \pi_i(x^*(y), y)$ is the first-period reduced-form payoff for player i. Vives (2009) shows that $V_i(y)$ is supermodular in y provided that for any player i: π_i is increasing and convex in each component of $x_j, j \neq i$, and each component of $x_j^*(y)$ is supermodular in y. The result can be generalized to Markov finite-horizon multistage games with observable actions (e.g., Fudenberg and Tirole, 1991), where the payoff to each player displays increasing differences in any two variables.[27]

An example of the result is provided by the linear demand Bertrand oligopoly with advertising when advertising levels are chosen in a first stage and are observable. Under the assumptions made (Section 3.3.2), profits are supermodular in any pair of arguments, and the first-stage value function at extremal equilibria is supermodular (that is, advertising expenditures are strategic complements). Indeed, the assumptions are fulfilled in the classical linear gross substitutes products Bertrand competition model with constant marginal costs when either advertising or investment in product quality raises the demand intercept of the firm exerting the effort (Vives, 1985a) or increases the willingness to pay for the product of the firm by lowering the absolute value of the slope of demand $|\partial D_i / \partial p_i|$ (Vives, 1990b). In this case, for a given advertising effort there is a unique price equilibrium at the second stage.[28]

The result can be extended to a duopoly case in which, for all i, $\pi_i(x, y)$ has increasing differences in $(x_i, -x_j)$, $(y_i, -y_j)$, and $(x_i, (y_i, -y_j))$, $j \neq i$. An example is provided by a linear demand and cost Cournot duopoly in which outputs are strategic substitutes and y_i is the cost-reduction effort by firm i. Let $\pi_i = P_i(x_1, x_2)x_i - C_i(x_i, y_i)$ with $\partial^2 C_i / \partial x_i \partial y_i \leq 0$.

[27] Nonetheless, the result cannot be extended to the case where each payoff function $\pi_i(x, y)$ fulfills the ordinal complementarity conditions or the single-crossing property in any pair of variables (Echenique, 2004b).

[28] If firms invest in cost reduction, the second-stage SC is transformed into a first-stage SS. The same happens with investments in models of vertical quality differentiation when the market is covered (Shaked and Sutton, 1982).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Then the assumptions are fulfilled because $\partial^2\pi_i/\partial x_i\partial y_i \geq 0$, and $\partial^2\pi_i/\partial x_i\partial y_j = \partial^2\pi_i/\partial y_i\partial y_j = 0$ for $j \neq i$. We then have that cost-reduction investments are strategic substitutes at the first stage.[29]

With some further restrictions we can find conditions for increasing or decreasing dominance, that is, whether an initial dominance is reinforced by subsequent market actions (Athey and Schmutzler, 2001). Examples are provided by the Bertrand differentiated oligopoly model with learning by doing or, alternatively, with production adjustment costs, or even with switching costs. Similar results can be obtained in the Cournot model with network demand externalities (Katz and Shapiro, 1986).

4.3.2 Infinite-horizon games

Consider an infinite-horizon simultaneous move game with discount factor δ, and let $V_i(y)$ be the continuous value function associated to player i at a stationary MPE. Player i solves

$$\max_{x_i} \{\pi_i(x, y) + \delta V_i(x)\}.$$

Assume $x^*(y)$ is the unique contemporaneous Nash equilibrium given y. We have that $x^*(y)$ is increasing in y (i.e., we have intertemporal SC: x_i^* increases with y_j for $j \neq i$) if for all i:

1. $\pi_i(x, y) + \delta V_i(x)$ has increasing differences in (x_i, x_{-i}); and
2. π_i has increasing differences in (x_i, y).

In order for (1) to hold it is sufficient that both π_i and V_i have increasing differences in (x_i, x_{-i}).

Similarly, we have the corresponding result for a duopoly with strategic substitutability. We have that x_i^* increases in $(y_i, -y_j)$ (i.e., we have intertemporal SS: x_i^* decreases with y_j for $j \neq i$) if for all i:

1. $\pi_i(x, y) + \delta V_i(x)$ has increasing differences in $(x_i, -x_j)$, $j \neq i$; and
2. π_i has increasing differences in $(x_i, (y_i, -y_j))$.

We check the fulfillment of the conditions in an adjustment cost model (see Vives, 2005a for the alternating move duopoly). With simultaneous moves and adjustment costs, the payoff to player i is given by

$$\pi_i(x, y) = u_i(x) - F_i(x, y),$$

where $u_i(x)$ is the current profit in the period and $F_i(x, y)$ is the convex adjustment cost in going from past actions (y) to current actions (x) with $F_i(x, x) = 0$, $i = 1, 2$; that is, when actions are not changed, there is no adjustment cost. We can interpret actions as either prices or quantities and correspondingly let price or production bear the adjustment cost. Models with price adjustment costs, or "menu costs," are commonly used in macroeconomics.

[29] With linear demand there is a unique equilibrium at the second stage (see Vives, 1990b for a computed example where investment reduces the slope of marginal costs and for a reinterpretation in terms of firms that invest in expanding their own market).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Our conditions are fulfilled in a linear-quadratic specification. With price competition (and static SC) and menu costs, the marginal profit for firm i is increasing in the price y_i charged by the firm in the previous period and is independent of the price y_j charged by the rival in the previous period. Furthermore, the value function V_i displays SC. With quantity competition (static SS) and production adjustment costs, the marginal profit for firm i is increasing in the production y_i of the firm in the previous period and independent of the production y_j of the rival in the previous period. The value function displays SS in the duopoly case.

In these two cases, static SC or SS is transformed into intertemporal SC or SS. However, this need not be always the case. Jun and Vives (2004) fully characterize the linear and stable MPE in a symmetric differentiated duopoly model with quadratic payoffs and adjustment costs in a continuous time infinite-horizon differential game. They also find that contemporaneous (dynamic) SC or SS are inherited from static SC or SS but intertemporal SC or SS obtains depending on what variable bears the adjustment cost. If production is costly to adjust then intertemporal SS obtains, whereas if price is costly to adjust then intertemporal SC obtains. In particular, for the mixed case of price competition with production adjustment costs, the static SC is transformed into intertemporal SS.[30] Having intertemporal SC or SS matters because it governs strategic incentives at the MPE with respect to non-strategic behavior at the open-loop equilibrium. Indeed, with intertemporal SC (SS), steady-state prices at the MPE are above (below) the stationary open-loop equilibrium prices. This provides a generalization of the taxonomy of strategic behavior in two-stage games of Section 4.2 to the full-blown infinite-horizon game.

4.3.3 Existence of MPE in stochastic games

Existence of MPE in deterministic dynamic games has been shown only in particular models such as the linear-quadratic, and general results in stochastic games have been difficult to come by and rely on strong assumptions (particularly on transition probabilities) precluding deterministic transitions when actions spaces are uncountably infinite.[31] Lattice-theoretic methods are of help when there is enough monotonicity in the problem under study.

The existence of (stationary) MPE of stochastic games with complementarities in discrete time and infinite horizon is studied by Curtat (1996) under strong assumptions. He considers multidimensional action spaces and a multidimensional state evolving according to a transition probability as a function of the current state and action profile. Payoffs are smooth and display per-period complementarities and positive spillovers (the payoff to a player is increasing in the actions of rivals and the state); the transition distribution function is smooth, displays complementarities, and is stochastically increasing in actions and states. Furthermore, the payoff to a player as well as the transition distribution function fulfill a strict dominant diagonal condition.[32] These assumptions allow the collapse of the multiperiod

30 The reason – as in the learning curve model with price competition – is that a firm wants to make the rival small today in order to induce it to price softly tomorrow. A cut in price today will therefore bring a price increase by the rival tomorrow.

31 See the discussion of the literature in Duggan (2012).

32 The continuity assumptions on the transition probability are akin to Amir (1996b) and Nowak (2007) who also proved the existence of stationary MPE in games possessing strategic complementarities with uncountable state and action spaces.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

problem to a reduced-form static game (with continuation value functions increasing in the state variable), which can be shown to be supermodular. An equilibrium can then be found with value functions increasing in the state. An example fulfilling the assumptions is a dynamic version of a Cournot oligopoly with complementary products and learning by doing, where a high level of accumulated output by one firm yields stochastically higher levels of cumulated experience and lower production costs to the firm (learning by doing) and to the rivals (spillovers).

Balbus, Reffett, and Wozny (2014) consider an n-player discounted infinite-horizon stochastic game in discrete time. They allow for multidimensional action spaces and a multidimensional state (compact cube in Euclidean space with smallest point at 0) evolving according to a transition probability as a function of the current state and action profile. Payoffs are continuous and display per-period complementarities and positive spillovers (the payoff to a player is increasing in the actions of rivals); the transition distribution function is continuous, displays complementarities and is stochastically increasing in actions, and fulfills a strong mixing assumption, with a positive probability of setting the state to zero. The authors show existence of a largest and a smallest stationary Markov-Nash equilibrium. The assumptions relax smoothness (Lipschitz continuity) conditions of Curtat (1996) as well as the increasing differences assumptions between actions and states. That is, they do not require *monotone* Markov equilibrium to obtain existence. The results can be applied to studying supermodular price competition with durable goods (and convex costs). Here the state is a demand shock and the assumptions on transition probabilities mean that there is a positive probability that the market disappears and that high prices today result in a high probability of future positive demand, and there is no need for the monotonicity assumption at a high demand state today to translate stochastically into a high demand state tomorrow. Under the assumptions extremal strategies display intertemporal strategic complementarity.

Sleet (2001) considers a version of the adjustment cost model of the previous section in an infinite-horizon discrete game with a continuum of heterogeneous players and symmetric payoffs. This a dynamic monopolistic competition model with menu costs where firms interact repeatedly over an infinite horizon and each firm receives an idiosyncratic demand or cost shock every period. The demand for the product of a firm may depend on the average price charged in the market or on a price index. The assumptions are fulfilled with linear or constant elasticity demands, quadratic or constant elasticity production costs (subject to a multiplicative shock), and quadratic costs of price adjustment.

5 UNCERTAINTY AND PRIVATE INFORMATION: BAYESIAN OLIGOPOLY GAMES

The lattice-theoretical approach allows for general strategy spaces and payoff functions and therefore is apt for games of incomplete information. I present in this section a framework for Bayesian games and three approaches to characterize equilibria in pure strategies together with applications: supermodular games (Vives, 1990a), single-crossing properties (Athey, 2001), and "monotone supermodular" games (Van Zandt and Vives, 2007). The last two approaches deliver conditions for equilibria to be monotone in type, a desirable property in applications.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

5.1 A Framework for Bayesian Games

Let T_i be the set of possible types t_i of player i, a subset of Euclidean space. The types of the players are drawn from a common prior distribution μ on $T = \prod_{i=0}^{n} T_i$, where T_0 is interpreted as unobserved residual uncertainty. In a game of incomplete information, the type of a player embodies all the decision-relevant private information. The action space of player i is a compact cube of Euclidean space A_i, and his payoff is given by the (measurable and bounded) function $\pi_i : A \times T \to \mathbb{R}$. The (ex post) payoff to firm i when the profile of actions is $a = (a_1, \ldots, a_n)$ and the realized types $t = (t_1, \ldots, t_n)$ is thus $\pi_i(a; t)$. Action spaces, payoff functions, type sets, and the prior distribution are common knowledge. The Bayesian game is fully described by $(A_i, T_i, \pi_i; i \in N)$.

A (pure) strategy for player i is a (measurable) function $\sigma_i : T_i \to A_i$ that assigns an action to every possible type of the player. Let Σ_i denote the strategy space of player i (and identify strategies σ_i and τ_i if they are equal with probability 1). Denote the expected payoff to player i, when agent j uses strategy σ_j, by $U_i(\sigma) = E\pi_i(\sigma_1(t_1), \ldots, \sigma_n(t_n); t)$ where $\sigma = (\sigma_1, \ldots, \sigma_n)$. A Bayesian Nash equilibrium is a Nash equilibrium of the game $(\Sigma_i, U_i, i \in N)$ where the strategy space and payoff function of player i are denoted Σ_i and U_i, respectively.[33]

The formulation of the Bayesian game encompasses common and private values as well as perfect or imperfect signals. With pure private values, allowing for correlated types, say costs of firms, we have $\pi_i(a; t_i)$. With a common value case, say a demand shock where firm i observes component t_i only, we may have that $\pi_i(a; t) = v_i(a; \Sigma_i t_i)$. For an example of imperfect signals, suppose firms observe with noise their cost parameters. In this case t_0 could represent the n-vector of firms' cost parameters and t_i the private cost estimate of firm i, allowing for correlation among the cost parameters as well as the error terms in the private signals.

5.2 Equilibrium Existence in Pure Strategies

To show existence of pure-strategy equilibria in games of incomplete information with a continuum of types and/or actions has proved difficult. Known sufficient conditions for existence include typically conditionally independent types, finite action spaces, and atomless distributions for types.[34] Under these assumptions existence of mixed strategy equilibria is shown first and then equilibria are purified. The lattice-theoretic approach has provided results:

1. for supermodular games with general action and type spaces (Vives, 1990a);
2. for games satisfying single-crossing properties in which each player uses a strategy that is monotone (increasing) in type in response to monotone (increasing) strategies of rivals (Athey, 2001, McAdams, 2003, 2006, Reny, 2011); and
3. for "monotone" supermodular games with general action and type spaces (Van Zandt and Vives, 2007).

---

33 Denote by $\beta_i : \Sigma_{-i} \to \Sigma_i$ player i's best-reply correspondence in terms of strategies. Then a Bayesian Nash equilibrium is a strategy profile σ such that $\sigma_i \in \beta_i(\sigma_{-i})$ for $i \in N$. We can define a natural order in the strategy space Σ_i: $\sigma_i \leq \sigma_i'$ if $\sigma_i(t_i) \leq \sigma_i'(t_i)$, in the usual component-wise order, with probability 1 on T_i.

34 See Radner and Rosenthal (1982) and Milgrom and Weber (1985). Khan and Sun (1995) show existence of pure-strategy equilibria when types are independent, payoffs continuous, and action sets countable.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

In the first approach, existence of pure-strategy Bayesian equilibria follows directly from supermodularity of the underlying family of games defined with the ex post payoffs for given realizations of the types of the players. A key observation is that supermodularity of this underlying family of games is inherited by the Bayesian game.[35] Existence of extremal pure strategy Bayesian equilibria then follows from the general versions of the results in Section 2 (see also Vives, 1990a, 1999, Sec. 2.7.3). This existence result holds for multidimensional action spaces and requires no distributional restrictions. Applications of this approach can be found beyond oligopoly games in Diamond's (1982) search model, and natural resource exploration games with private information (Hendricks and Kovenock, 1989 and Milgrom and Roberts, 1990).

5.2.1 Single-crossing properties

In this approach, conditions are imposed so that an equilibrium in monotone increasing strategies in types can be found. Suppose that both action A_i and type sets T_i for any player i are compact subsets of the real line and that types have a joint density μ that is bounded, atomless, and log-supermodular (i.e., types are affiliated). Suppose also that $\pi_i\,(a, t)$ is continuous and supermodular in a_i and has increasing differences in (a_i, a_{-i}) and (a_i, t) or, alternatively, that $\pi_i\,(a, t)$ is non-negative and log-supermodular in (a, t). Then the Bayesian game has a pure-strategy equilibrium in increasing strategies (Athey, 2001). Note that under the assumptions the first approach outlined already delivers existence of a pure-strategy equilibrium.[36] An example of the result in the differentiated Bertrand oligopoly has firm i with random marginal cost t_i with both $D_i(p_i, p_{-i})$ and the joint density of $(t_1, \ldots, t_n)$ log-supermodular. Then if the strategies of rivals, $p_j\,(\cdot)$, $j \neq i$, are increasing in types, $E(\pi_i \mid t_i) = (p_i - t_i)E(D_i(p_i, p_{-i}(t_{-i})) \mid t_i)$ is log-supermodular in (p_i, t_i) and the best-reply map of player i is increasing in t_i.

The approach can also be used in games that are not of strategic complementarities and with discontinuous payoffs. The existence of monotone equilibria in pure strategies can be shown for first-price auctions with heterogeneous (weakly) risk-averse bidders characterized by private affiliated values or common value and conditionally independent signals (Athey, 2001), as well as for uniform-price auctions featuring multiunit demand, interdependent values and independent types (McAdams, 2003, 2006).[37] Reny (2011) extends those results using a fixed-point theorem by Eilenberg and Montgomery (1946) that replaces the requirement in Kakutani or Glicksberg of convex-valued correspondences used by Athey

[35] Let π_i be supermodular in a_i and have increasing differences in (a_i, a_{-i}). Then $U_i(\sigma)$ is supermodular in σ_i and has increasing differences in (σ_i, σ_{-i}), because supermodularity and increasing differences are preserved by integration. Furthermore, strategy spaces in the Bayesian game Σ_i can be shown to have the appropriate order structure (i.e., they are complete lattices). Then the game $(\Sigma_i, U_i, i \in N)$ is a GSC and for all $\sigma_{-i} \in \Sigma_{-i}$, $\beta_i(\sigma_{-i})$ contains extremal elements $\bar{\beta}_i(\sigma_{-i})$ and $\underline{\beta}_i(\sigma_{-i})$.

[36] The proof of these results relies on the standard Kakutani fixed-point theorem, based on convex-valued correspondences since with discrete action spaces and under the prevailing assumptions, best-response correspondences are convex valued. A key step in the proof is to show that if the rivals of player i use increasing strategies then the payoff to player i fulfills an appropriate single-crossing property (e.g., is log-supermodular or has increasing differences) in action and type. This ensures that a player uses a strategy that is increasing in his type as a best response to increasing strategies of rivals. The existence result for discrete action spaces can then be used to show existence with a continuum of actions via a purification argument.

[37] McAdams (2006) uses a discrete bid space and atomless types to show existence of monotone equilibria with risk-neutral bidders checking that the single-crossing condition in Athey (2001) used in the single-object case extends to multiunit auctions.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

(2001) by contractible-valued ones.[38] Reny (2011) weakens previous conditions on interim payoff functions for the monotone best-reply condition to hold. The conditions allow for infinite-dimensional type and action spaces, general joint distributions over types, general partial orders on both action and type spaces (this is useful since single crossing may fail for one partial order but hold for another), and dispense with single crossing (although it remains very useful).

The results are applied to prove existence of monotone equilibrium in uniform-price multiunit auctions with weakly risk-averse bidders and interdependent values (and where bids are restricted to a finite grid) and to oligopoly pricing using judicious partial orders over types. The oligopoly application considers n firms competing with differentiated products with random constant marginal costs and random demand. The firms are partitioned into two groups with goods being substitutes within each group. Firms have private information about both cost and demand conditions and marginal costs are affiliated and information about demand may be correlated across firms. It is shown that a pure-strategy price equilibrium exists with prices monotone in costs (which is the coordinate in which strict single crossing holds, and also in the demand signal according to the defined partial order).

5.2.2 Monotone supermodular games

For "monotone" supermodular games with multidimensional action spaces, and type spaces, a strong result is provided by Van Zandt and Vives (2007). Let $\Delta(T_{-i})$ be the set of probability distributions on T_{-i} and let player i's posteriors be given by the (measurable) function $p_i : T_i \to \Delta(T_{-i})$, consistent with the prior μ. The following properties define a monotone supermodular game:

1. *supermodularity and complementarity between action and type* with π_i supermodular in a_i, and with increasing differences in (a_i, a_{-i}) and in (a_i, t);
2. *monotone posteriors* with $p_i : T_i \to \Delta(T_{-i})$ increasing with respect to the partial order on $\Delta(T_{-i})$ of first-order stochastic dominance (a sufficient but not necessary condition is that μ be affiliated).

The result is that in a monotone supermodular game there is a largest and a smallest Bayesian equilibrium and each one is in monotone strategies. There might be other equilibria that are in non-monotone strategies but, if so, they will be "sandwiched" between the largest and the smallest ones. Furthermore, extremal equilibria are increasing in posteriors (according to first-order stochastic dominance) and can be obtained through the iterative application of the best-reply map β.

The assumptions on action and type spaces can be considerably weakened (beyond Euclidean spaces), and there is no need to assume a common prior, but the result cannot

[38] A set is contractible if, within itself, it can be continuously deformed to a single point. Convex sets are, indeed, contractible but the converse is not true. The author finds conditions for the existence of a monotone (pure) equilibrium whenever monotone (pure) best replies are non-empty and join-closed in response to rivals using monotone best replies. Those conditions require, among other technical conditions, payoffs to be bounded, jointly measurable in actions and types, continuous in actions for every type, and the marginal distribution of types to be atomless.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

be extended to log-supermodular payoffs. Yang and Qi (2014) provide an extension of the results to non-atomic games.[39]

Monotone supermodular games fit a variety of problems.[40] We provide here applications to strategic information revelation and endogenous information acquisition.

Comparative statics and strategic information revelation If payoffs display positive spillovers (π_i is increasing in a_{-i}), then increasing the posteriors increases the equilibrium expected payoffs. This is a consequence of extremal equilibria being increasing in posteriors and it implies that the expected payoff of each player in an extremal equilibrium is increasing in the posteriors of the other players. The result can easily be strengthened to "strictly increasing" under certain regularity assumptions (including some smooth strict complementary conditions and requiring π_i to be strictly increasing in a_j).

Okuno-Fujiwara, Postlewaite, and Suzumura (1990) have provided conditions under which fully revealing equilibria obtain in duopoly games of voluntary disclosure of information when information is verifiable. The conditions involve regularity assumptions such as one-dimensional actions, concavity of payoffs, uniqueness and interiority of equilibrium, and independent types for the players.[41] Once we realize that their framework is within the realm of monotone supermodular games, it is only necessary that the marginal payoff of an action of a player is strictly increasing in the actions of rivals and in the types of players. The results extend to n-player GSC games and to a duopoly with strategic substitutability, multidimensional actions, affiliated types, and possibly multiple non-interior extremal equilibria. Mensch (2016) provides an extension of the existence of monotone equilibrium results in Reny (2011) to dynamic games (such as stopping games) and uses them to weaken the assumptions to obtain a full separating equilibrium in voluntary disclosure games.[42]

Endogenous information acquisition Amir and Lazzati (2016) study covert endogenous information acquisition in the framework of common-value monotone supermodular games. The authors use the supermodular stochastic order to arrange the information structures (joint distribution of state of the world and signals) and show that better information increases expected payoffs. If a convexity assumption is added, implying that increasing the quality of information raises informativeness with increasing returns, then the value of information for a player is convex in its quality. This leads to extreme behavior of agents with choices of a highest- or a lowest-quality signal. The results contrast with models with linear-quadratic payoffs and Gaussian information structures with concave values of information (Vives, 2008b).[43]

39 Under their assumptions monotone equilibria form a complete lattice.

40 Van Zandt and Vives (2007) present an application to the discrete setup of an adoption game on a graph with local network effects, and so-called "global games" are typically monotone supermodular (see Section 7.2 in Vives, 2005a).

41 The basic intuition for the result is that in equilibrium inferences are skeptical: if a player reports a set of types, others believe the worst (i.e., others believe that the player is of the most unfavorable type in the reported set). This unravels the information.

42 The author constructs an auxiliary static game to deal with the endogeneity of beliefs in the dynamic game preserving at the same time the continuity of payoffs. Then he uses the methods in Athey (2001) and Reny (2011) to show that there exist monotone best replies to monotone strategies by rivals.

43 Myatt and Wallace (2016) provide an analysis of information use and acquisition in a strategic complements price-setting oligopoly with differentiated products with a continuum of goods, where a finite number of suppliers

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

5.3 Complementarities in Uniform-price Divisible Good Auctions and Behavioral Traders

Progress in the characterization of equilibria in uniform-price divisible good auctions when there is incomplete information and market power has been made in a linear-Gaussian specification. Kyle (1989) considers a Gaussian model of a divisible good double auction where some bidders are privately informed and others are uninformed. Vives (2011) shows how increased correlation in the values of the traders increases their market power and how private information generates market power over and above the full information level. Bergemann, Heumann, and Morris (2015) generalize the information structure in Vives (2011), keeping the symmetry assumption. Rostek and Weretka (2012) partially relax the symmetry assumption in Vives (2011) and replace it with a weaker "equicommonality" assumption on the matrix correlation among the agents' values. Manzano and Vives (2016) consider the case of two types of bidders.

Consider the case of supply function competition to fix ideas. This corresponds to competition in the wholesale market for electricity in many countries. It is worth noting that restricting attention to linear supply functions and with complete information there is strategic complementarity in the supply slopes (e.g., Klemperer and Meyer, 1989 and Akgün, 2004) despite the game not being supermodular. The reason is that if rivals of a firm increase the slope of supply then the residual demand left for the firm becomes steeper and induces this firm to also set a steeper supply. Interestingly, with uncertainty and incomplete information where prices convey information (say about an uncertain common cost component), there is an inference effect that moderates, and may even reverse, the strategic complementarity in slopes. This is so since when costs are positively correlated, the inference effect moderates the reaction to the price (a high price means high costs), the more so, the more rivals react to the price. The reason is that a higher reaction to the price by rivals induces a trader to also give a higher weight to the price in the estimation of his cost and hence it increases the magnitude of the inference effect. However, the equilibrium happens at a point where there are strategic complementarities in slopes (see Bayona, Brandts, and Vives, 2016).

When traders or firms have trouble retrieving the information from the price, be it because they neglect the correlation between random variables or because of other types of bounded rationality, then the outcome is more competitive since high prices are not interpreted as signals of high costs. If there is a proportion of such naive traders in the market then sophisticated traders, who take into account information in prices as well as the presence of naive traders who bid relatively flat supply schedules, will respond also with flatter supply schedules and the outcome will be more competitive than predicted by the Bayesian equilibrium in supply functions with fully rational traders. This is in fact what happens in the experiment conducted in Bayona et al. (2016) and that leads to the similar behavior of naive and sophisticated sellers being not so distinct. The general phenomenon was first noted by Camerer and Fehr (2006) in the context of games characterized by strategic complementarities and in the presence of sophisticated and boundedly rational subjects. When actions are strategic complements then sophisticated players align their actions with those of naive

have access to multiple sources of information about the uncertain common demand level. Myatt and Wallace (2015) examine similar issues in the context of a Cournot model. Bonatti, Cisternas, and Toikka (2016) examine learning and signaling in a dynamic Cournot game with firms having private information on costs and observing the (noisy) market price.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

players, and the former do not provide a check on the effects of naive strategies on outcomes. Instead, when actions are strategic substitutes then sophisticated players counteract the actions of naive ones on aggregate behavior. The result is that under strategic substitutes a small proportion of sophisticated agents may be sufficient to lead to aggregate outcomes not far from the equilibrium predictions with fully rational agents, while under strategic complements, a not so large proportion of naive agents may generate outcomes far from the rational agent equilibrium predictions. This means that rational agent equilibrium analysis in games of strategic complementarities may be less robust than in games of strategic substitutability.

6 CONCLUDING REMARKS

In this chapter I have provided a selective survey of the theory and applications of the lattice-theoretic approach in the study of oligopoly games. The approach has proved fruitful well beyond the domain of games of strategic complementarities. Indeed, in many situations, patterns of complementarity and substitutability are present but still the approach is useful and delivers existence and characterization of equilibrium results as well as comparative statics analysis. In fact, the approach has proved useful in all domains of economic theory and is being progressively incorporated in the standard toolbox of economics, including empirical studies. The research agenda ahead is challenging in terms of continuing to push the frontiers of the theory with a view toward applications including heterogeneous agents and fully dynamic games with incomplete information, as well as developing the empirical analysis more fully.[44]

REFERENCES

Akgün, U. 2004. "Mergers with Supply Functions." *Journal of Industrial Economics* 52(4): 535–546.

Alexandrov, A., and Ö. Bedre-Defolie. 2017. "LeChatelier-Samuelson Principle in Games and Pass-through of Shocks. " *Journal of Economic Theory* 168: 44–54.

Amir, R. 1996a. "Cournot Oligopoly and the Theory of Supermodular Games." *Games and Economic Behavior* 15(2): 132–148.

Amir, R. 1996b. "Continuous Stochastic Games of Capital Accumulation with Convex Transitions." *Games and Economic Behavior* 15: 111–131.

Amir, R., and L. de Castro. 2017. "Nash Equilibrium in Games with Quasi-monotonic Best-responses." *Journal of Economic Theory* 172: 220–246.

Amir, R., and V. Lambson. 2000. "On the Effects of Entry in Cournot Markets." *Review of Economic Studies* 67(2): 235–254.

Amir, R., and N. Lazzati. 2016. "Endogenous Information Acquisition in Bayesian Games with Strategic Complementarities." *Journal of Economic Theory* 163: 684–698.

Amir, R., L. de Castro, and L. Koutsougeras. 2014. "Free Entry versus Socially Optimal Entry." *Journal of Economic Theory* 154: 121–135.

Anderson, S., A. de Palma, and J.F. Thisse. 1992. *Discrete Choice Theory of Product Differentiation*. Cambridge, MA and London: MIT Press.

Anderson, S., N. Erkal, and D. Piccinin. 2016. "Aggregative Oligopoly Games with Entry." Mimeo.

[44] Progress on the empirical front in dynamic oligopoly has been made, among others, by Ericson and Pakes (1995), Bajari, Lanier Nenkard, and Levin (2007), Fershtman and Pakes (2012), and Ifrach and Weintraub (2017). An example of empirical analysis of complementarities is provided by Miravete and Pernías (2006), and of markets with multiple equilibria by Sweeting (2006), Ciliberto and Tamer (2009), and Galichon and Henry (2011). A promising approach based on monotone comparative statics is provided by Echenique and Komunjer (2009).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Armstrong, M. 2006. "Competition in Two-sided Markets." *RAND Journal of Economics* 37(3): 668–691.
Athey, S. 2001. "Single Crossing Properties and the Existence of Pure Strategy Equilibria in Games of Incomplete Information." *Econometrica* 69(4): 861–890.
Athey, S., and A. Schmutzler. 2001. "Investment and Market Dominance." *RAND Journal of Economics* 32(1): 1–26.
Bajari, P., C. Lanier Nenkard, and J. Levin. 2007. "Estimating Dynamic Models of Imperfect Competition." *Econometrica* 75(5): 1331–1370.
Balbus, L., K. Reffett, and L. Wozny. 2014. "A Constructive Study of Markov Equilibria in Stochastic Games with Strategic Complementarities." *Journal of Economic Theory* 150: 815–840.
Bamon, R., and J. Fraysśe. (1985). "Existence of Cournot Equilibrium in Large Markets." *Econometrica* 53(3): 587–597.
Bayona, A., J. Brandts, and X. Vives. 2016. "Supply Function Competition, Private Information and Market Power: A Laboratory Study." *CEPR Discussion Paper* 11378.
Bergemann, D., T. Heumann, and S. Morris. 2015. "Information and Volatility." *Journal of Economic Theory* 158: 427–465.
Bernheim, D. 1984. "Rationalizable Strategic Behavior." *Econometrica* 52(4): 1007–1028.
Berry, S., J. Levinsohn, and A. Pakes. 1999. "Voluntary Export Restraints on Automobiles: Evaluating a Trade Policy." *American Economic Review* 89(3): 400–430.
Bonatti, A., G. Cisternas, and J. Toikka. 2016. "Dynamic Oligopoly with Incomplete Information." *Review of Economic Studies* 84(2): 503–546.
Camerer, C., and E. Fehr. 2006. "When Does 'Economic Man' Dominate Social Behavior?" *Science* 311: 47–52.
Ciliberto, F., and E. Tamer. 2009. "Market Structure and Multiple Equilibria in Airline Markets." *Econometrica* 77(6): 1791–1828.
Corchón, L. 1994. "Comparative Statics for Aggregative Games: The Strong Concavity Case." *Mathematical Social Sciences* 28: 151–165.
Cumbul, E., and G. Virág. 2014. "Non-supermodular Price Setting Games." Mimeo.
Curtat, L. 1996. "Markov Equilibria of Stochastic Games with Complementarities." *Games and Economic Behavior* 17(2): 177–199.
Diamond, P. 1982. "Aggregate Demand Management in Search Equilibrium." *Journal of Political Economy* 90(5): 881–894.
Dixit, A. 1986. "Comparative Statics for Oligopoly." *International Economic Review* 27: 107–122.
Duggan, J. 2012. "Noisy Stochastic Games." *Econometrica* 80(5): 2017–2045.
Echenique, F. 2002. "Comparative Statics by Adaptive Dynamics and the Correspondence Principle." *Econometrica* 70(2): 833–844.
Echenique, F. 2004a. "A Characterization of Strategic Complementarities." *Games and Economic Behavior* 46(2): 325–347.
Echenique, F. 2004b. "Extensive-form Games and Strategic Complementarities." *Games and Economic Behavior* 46(2): 348–364.
Echenique, F., and I. Komunjer. 2009. "Testing Models with Multiple Equilibria by Quantile Methods." *Econometrica* 77: 1281–1297.
Edlin, A., and C. Shannon. 1998. "Strict Monotonicity in Comparative Statics." *Journal of Economic Theory* 81(1): 201–219.
Eilenberg, S., and D. Montgomery. 1946. "Fixed Point Theorems for Multi-valued Transformations." *American Journal of Mathematics* 68(2): 214–222.
Ericson, R., and A. Pakes. 1995. "Markov-perfect Industry Dynamics: A Framework for Empirical Work." *Review of Economic Studies* 62(1): 53–82.
Fershtman, C., and A. Pakes. 2012. "Dynamic Games with Asymmetric Information: A Framework for Empirical Work." *Quarterly Journal of Economics* 127(4): 1611–1661.
Friedman, J. 1983. *Oligopoly Theory*. Cambridge, UK: Cambridge University Press.
Fudenberg, D., and J. Tirole. 1984. "The Fat Cat Effect, the Puppy Dog Ploy and the Lean and Hungry Look." *American Economic Review* 74: 361–368.
Fudenberg, D., and J. Tirole. 1985. "Preemption and Rent Equalization in the Adoption of New Technology." *Review of Economic Studies* 52(3): 383–401.
Fudenberg, D., and J. Tirole. 1991. *Game Theory*. Cambridge, MA and London: MIT Press.
Gabaix, X., D. Laibson, and D. Li. 2016. "The Impact of Competition on Prices with Numerous Firms." *Journal of Economic Theory* 165: 1–24.
Galichon, A., and M. Henry. 2011. "Set Identification in Models with Multiple Equilibria." *Review of Economic Studies* 78: 1264–1298.
Hendricks, K., and D. Kovenock. 1989. "Asymmetric Information, Information Externalities, and Efficiency: The Case of Oil Exploration." *RAND Journal of Economics* 20(2): 164–182.
Ifrach, B., and G. Weintraub. 2017. "A Framework for Dynamic Oligopoly in Concentrated Industries." *Review of Economic Studies* 84(3): 1106–1150.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Immordino, G. 2009. "Advertising and Cost Reduction." *The B.E. Journal of Theoretical Economics* 9(1): 1–15.
Jun, B., and X. Vives. 2004. "Strategic Incentives in Dynamic Duopoly." *Journal of Economic Theory* 116(2): 249–281
Katz, M., and C. Shapiro. 1986. "Technology Adoption in the Presence of Network Externalities." *Journal of Political Economy* 94: 822–841.
Khan, A., and Y. Sun. 1995. "Pure Strategies in Games with Private Information." *Journal of Mathematical Economics* 24: 633–653.
Klemperer, P., and M. Meyer. 1989. "Supply Function Equilibria in Oligopoly Under Uncertainty." *Econometrica* 57(6): 1243–1277.
Kukushkin, N. 1994. "A Fixed-point Theorem for Decreasing Mappings." *Economics Letters* 46(1): 23–26.
Kyle, A. 1989. "Informed Speculation with Imperfect Competition." *Review of Economic Studies* 56(3): 317–355.
Lee, T., and L. Wilde. 1980. "Market Structure and Innovation: A Reformulation." *Quarterly Journal of Economics* 94(2): 429–436.
Mankiw, G., and M. Whinston. 1986. "Free Entry and Social Inefficiency." *RAND Journal of Economics* 17(1): 48–58.
Manzano, C., and X. Vives. 2016. "Market Power and Welfare in Asymmetric Divisible Good Auctions." Mimeo.
McAdams, D. 2003. "Isotone Equilibrium in Games of Incomplete Information." *Econometrica,* 71(4): 1191–1214.
McAdams, D. 2006. "Monotone Equilibrium in Multi-unit Auctions." *Review of Economic Studies* 73: 1039–1056.
McManus, M. 1962. "Numbers and Size in Cournot Oligopoly." *Yorkshire Bulletin of Social and Economic Research* 14.
McManus, M. 1964. "Equilibrium, Numbers and Size in Cournot Oligopoly." *Yorkshire Bulletin of Social Science and Economic Research* 16.
Mensch, J. 2016. "On the Existence of Monotone Pure-strategy Perfect Bayesian Equilibrium in Games with Complementarities." Mimeo.
Milgrom, P., and J. Roberts. 1990. "Rationalizability, Learning and Equilibrium in Games with Strategic Complementarities." *Econometrica* 58: 1255–1277.
Milgrom, P., and J. Roberts. 1994. "Comparing Equilibria." *American Economic Review* 84(3): 441–459.
Milgrom, P., and J. Roberts. 1996. "The LeChatelier Principle." *American Economic Review* 86(1): 173–179.
Milgrom, P., and C. Shannon. 1994. "Monotone Comparative Statics." *Econometrica* 62: 157–180.
Milgrom, P., and R. Weber. 1985. "Distributional Strategies for Games with Incomplete Information." *Mathematics of Operations Research* 10(4): 619–632.
Miravete, E., and J.C. Pernías. 2006. "Innovation Complementarity and Scale of Production." *Journal of Industrial Economics* 54(1): 1–29.
Mrazova, M., and P. Neary. 2016. "Not So Demanding: Demand Structure and Firm Behavior." *CEPR Discussion Paper* 11119.
Myatt, D., and C. Wallace. 2015. "Cournot Competition and the Social Value of Information." *Journal of Economic Theory* 158: 466–506.
Myatt, D., and C. Wallace. 2016. "Information Use and Acquisition in Price-setting Oligopolies." Mimeo.
Nocke, V., and N. Schutz. 2015. "An Aggregative Games Approach." *CEPR Discussion Paper* 11539.
Novshek, W. 1985. "On the Existence of Cournot Equilibrium." *Review of Economic Studies* 52(1): 85–98.
Nowak, A. 2007. "On Stochastic Games in Economics." *Mathematical Methods of Operations Research* 66: 513–530.
Okuno-Fujiwara, M., A. Postlewaite, and K. Suzumura. 1990. "Strategic Information Revelation." *Review of Economic Studies* 57, 25–47.
Pearce, D. 1984. "Rationalizable Strategic Behavior and the Problem of Perfection." *Econometrica* 52(4): 1029–1050.
Peitz, M. 2000. "Aggregation in a Model of Price Competition." *Journal of Economic Theory* 90(1): 1–38.
Quint, D. 2014. "Imperfect Competition with Complements and Substitutes." *Journal of Economic Theory* 152: 266–290.
Radner, R., and R. Rosenthal. 1982. "Private Information and Pure-strategy Equilibria." *Mathematical Operations Research* 7: 401–409.
Reny, P. 2011. "On the Existence of Monotone Pure-strategy Equilibria in Bayesian Games." *Econometrica* 79(2): 499–553.
Roberts, J., and H. Sonnenschein. 1976. "On Existence of Cournot Equilibrium without Concave Profit Functions." *Journal of Economic Theory* 13(1): 101–113.
Roberts, J., and H. Sonnenschein. 1977. "On the Foundations of the Theory of Monopolistic Competition." *Econometrica* 45(1): 101–113.
Rostek, M., and M. Weretka. 2012. "Price Inference in Small Markets." *Econometrica* 80(2): 687–711.
Samuelson, P. 1979. *Foundations of Economic Analysis*. New York: Atheneum (first published in 1947 by Harvard University Press, Cambridge, MA).
Schmediler, D. 1973. "Equilibrium Points of Nonatomic Games." *Journal of Statistical Physics* 7(4): 295–300.
Selten, R. 1970. *Preispolitik der Mehrproduktenunternehmung in der Statischen Theorie*. Berlin: Springer Verlag.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Shaked, A., and J. Sutton. 1982. "Relaxing Price Competition Through Product Differentiation." *Review of Economic Studies* 49: 3–13.

Singh, N., and X. Vives. 1984. "Price and Quantity Competition in a Differentiated Duopoly." *RAND Journal of Economics* 15(4): 546–554.

Sleet, C. 2001. "Markov Perfect Equilibria in Industries with Complementarities." *Economic Theory* 17(2): 371–397.

Spady, R. 1984. "Non-cooperative Price-setting by Asymmetric Multiproduct Firms." Mimeo. Bell Laboratoires.

Sweeting, A. 2006. "Coordination, Differentiation, and the Timing of Radio Commercials." *Journal of Economics and Management Strategy* 15(4): 909–942.

Tarski, A. 1955. "A Lattice-theoretical Fixpoint Theorem and its Applications." *Pacific Journal of Mathematics* 5: 285–308.

Thisse, J., and X. Vives. 1992. "Basing Point Pricing: Competition versus Collusion." *Journal of Industrial Economics* 40: 249–260.

Topkis, D. 1978. "Minimizing a Submodular Function on a Lattice." *Operations Research,* 26: 2.

Topkis, D. 1979. "Equilibrium Point in Nonzero-sum n-Person Submodular Games." *SIAM Journal of Control and Optimization* 17: 773–787.

Topkis, D. 1998. *Supermodularity and Complementarity: Frontiers of Economic Research Series*. Princeton, NJ: Princeton University Press.

Van Zandt, T., and X. Vives. 2007. "Monotone Equilibria in Bayesian Games of Strategic Complementarities." *Journal of Economic Theory* 134: 339–360.

Vives, X. 1985a. "Nash Equilibrium in Oligopoly Games with Monotone Best Responses." *CARESS W.P.* #85–10, University of Pennsylvania.

Vives, X. 1985b. "On the Efficiency of Bertrand and Cournot Equilibria with Product Differentiation." *Journal of Economic Theory* 36(1): 166–175.

Vives, X. 1990a. "Nash Equilibrium with Strategic Complementarities." *Journal of Mathematical Economics* 19(3): 305–321.

Vives, X. 1990b. "Information and Competitive Advantage." *International Journal of Industrial Organization* 8: 17–35.

Vives, X. 1999. *Oligopoly Pricing: Old Ideas and New Tools*. Boston, MA: MIT Press.

Vives, X. 2005a. "Complementarities and Games: New Developments." *Journal of Economic Literature* 43(2): 437–479.

Vives, X. 2005b. "Games with Strategic Complementarities: New Applications to Industrial Organization." *International Journal of Industrial Organization* 23: 625–637.

Vives, X. 2008a. "Innovation and Competitive Pressure." *Journal of Industrial Economics* 56(3): 419–469.

Vives, X. 2008b. *Information and Learning in Markets: The Impact of Market Microstructure*. Princeton, NJ: Princeton University Press.

Vives, X. 2009. "Strategic Complementarity in Multi-stage Games," *Economic Theory* 40: 151–171.

Vives, X. 2011. "Strategic Supply Function Competition with Private Information." *Econometrica* 79(6): 1919–1966.

Yang, J., and X. Qi. 2014. "The Nonatomic Supermodular Game." *Games and Economic Behavior* 82: 609–620.

Zhou, L. 1994. "The Set of Nash Equilibria of a Supermodular Game is a Complete Lattice." *Games and Economic Behavior* 7: 295–300.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

APPENDIX: BRIEF SUMMARY OF LATTICE-THEORETIC METHODS[45]

Definitions

A binary relation $\geq$ on a non-empty set X is a partial order if $\geq$ is reflexive, transitive, and anti-symmetric. An upper bound on a subset $A \subset X$ is $z \in X$ such that $z \geq x$ for all $x \in A$. A greatest element of A is an element of A that is also an upper bound on A. Lower bounds and least elements are defined analogously. The greatest and least elements of A, when they exist, are denoted $\max A$ and $\min A$, respectively. A supremum (resp., infimum) of A is a least upper bound (resp., greatest lower bound); it is denoted $\sup A$ (resp., $\inf A$).

A *lattice* is a partially ordered set $(X, \geq)$ in which any two elements have a supremum and an infimum. A lattice $(X, \geq)$ is *complete* if every non-empty subset has a supremum and an infimum. A subset L of the lattice X is a *sublattice* of X if the supremum and infimum of any two elements of L belong also to L.

Let $(X, \geq)$ and $(T, \geq)$ be partially ordered sets. A function $f\colon X \to T$ is *increasing* if, for x, y in X, $x \geq y$ implies that $f(x) \geq f(y)$.

A function $g\colon X \to \mathbb{R}$ on a lattice X is *supermodular* if all x, y in X, $g(\inf(x, y)) + g(\sup(x, y)) \geq g(x) + g(y)$. It is *strictly supermodular* if the inequality is strict for all pairs x, y in X that cannot be compared with respect to $\geq$ (i.e., neither $x \geq y$ nor $y \geq x$ holds). A function f is (*strictly*) *submodular* if $-f$ is (strictly) supermodular; a function f is (*strictly*) *log-supermodular* if $\log f$ is (strictly) supermodular.

Let X be a lattice and T a partially ordered set. The function $g\colon X \times T \to R$ has (*strictly*) *increasing differences* in (x, t) if $g(x', t) - g(x, t)$ is (strictly) increasing in t for $x' > x$ or, equivalently, if $g(x, t') - g(x, t)$ is (strictly) increasing in x for $t' > t$. Decreasing differences are defined analogously. If X is a convex subset of $\mathbb{R}^n$ and if $g\colon X \to R$ is twice-continuously differentiable, then g has increasing differences in (x_i, x_j) if and only if $\partial^2 g(x)/\partial x_i \partial x_j \geq 0$ for all x and $i \neq j$.

Results

Supermodularity is a stronger property than increasing differences: if T is also a lattice and if g is (strictly) supermodular on $X \times T$, then g has (strictly) increasing differences in (x, t). The two concepts coincide on the product of linearly ordered sets: if X is such a lattice, then a function $g\colon X \to \mathbb{R}$ is supermodular if and only if it has increasing differences in any pair of variables.

The complementarity properties are robust in the sense that they are preserved under addition or integration, pointwise limits, and maximization (with respect to a subset of variables, preserving supermodularity for the remaining variables).[46]

[45] More complete treatments can be found in Vives (1999, Ch. 2) and Topkis (1998).

[46] Supermodularity and increasing differences can be weakened to define an "ordinal supermodular" game, relaxing supermodularity to the weaker concept of quasi-supermodularity and increasing differences to a single-crossing property (see Milgrom and Shannon, 1994). However, such properties (unlike supermodularity and increasing differences) have no differential characterization and need not be preserved under addition or partial maximization operations.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Lemma 1 Monotonicity of optimal solutions *Let X be a compact lattice and let T be a partially ordered set. Let $u\colon X \times T \to \mathbb{R}$ be a function that (a) is supermodular and continuous on the lattice X for each $t \in T$ and (b) has increasing differences in (x, t). Let $\varphi(t) = \arg\max_{x \in X} u(x, t)$. Then:*

1. *$\varphi(t)$ is a non-empty compact sublattice for all t;*
2. *φ is increasing in the sense that, for $t' > t$ and for $x' \in \varphi(t')$ and $x \in \varphi(t)$, we have $\sup(x', x) \in \varphi(t')$ and $\inf(x', x) \in \varphi(t)$; and*
3. *$t \mapsto \max\phi(t)$ and $t \mapsto \min\phi(t)$ are well-defined increasing functions.*

Remark If u has strictly increasing differences in (x, t), then all selections of φ are increasing.

Remark If $X \subset \mathbb{R}^m$, solutions are interior, and $\partial u/\partial x_i$ is strictly increasing in t for some i, then all selections of φ are strictly increasing (Edlin and Shannon, 1998).

Theorem 2 (Tarski, 1955) *Let A be a complete lattice (e.g., a compact cube in $\mathbb{R}^m$). Then an increasing function $f : A \to A$ has a largest $\sup\{a \in A : f(a) \geq a\}$ and a smallest $\inf\{a \in A : a \geq f(a)\}$ fixed point.*

Supermodular game The game $(A_i, \pi_i; i \in N)$ is *supermodular* if, for all i, the following statements hold:

- A_i is a compact lattice.
- $\pi_i\,(a_i, a_{-i})$ is continuous:
  1. is supermodular in a_i; and
  2. has increasing differences in (a_i, a_{-i}).

Game of strategic complementarities Given a set of players N, strategy spaces A_i, and (non-empty) best-reply maps Ψ_i, $i = 1, \ldots, n$, we define a *game of strategic complementarities* (GSC) as one in which, for each i, A_i is a complete lattice and Ψ_i is increasing and has well-defined extremal elements.

Let $X \subset \mathbb{R}$. A function $f : X \to \mathbb{R}$ is *quasi-increasing* if for every $x \in X$, $\limsup_{y \uparrow x} f(y) \leq f(x) \leq \liminf_{y \downarrow x} f(y)$; f is *quasi-decreasing* if $-f$ is quasi-increasing. The following is a real-valued version of Theorem 3 in Tarski (1955).

Theorem 3 (Tarski's intersection point theorem) *If $f : [a, b] \longrightarrow \mathbb{R}$ is quasi-increasing, $g : [a, b] \longrightarrow \mathbb{R}$ is quasi-decreasing, $f(a) \geq g(a)$ and $f(b) \leq g(b)$, then the set $\{x \in [a, b] : f(x) = g(x)\}$ is non-empty, and has as largest element sup $\{x \in [a, b] : f(x) \geq g(x)\}$ and as smallest element inf $\{x \in [a, b] : f(x) \leq g(x)\}$.*

Corollary *Let $X = [a, b]$. Then a quasi-increasing function $f : X \to X$ has a largest $\overline{x} \equiv$* $\sup\{x \in X : f(x) \geq x\}$ and a smallest $\underline{x} \equiv \inf\{x \in X : x \geq f(x)\}$ fixed point.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

The result is easy to grasp considering a function f: $[0, 1] \to [0, 1]$ which when discontinuous jumps up but not down. The function must then cross the 45° line at some point. Indeed, suppose that it starts above the 45° line (otherwise, 0 is a fixed point); then it either stays above it (and then 1 is a fixed point) or it crosses the 45° line.

Comparative statics analysis is trivial when the function f is (strictly) increasing in a parameter t (for t in a partially ordered set T). Then $\overline{x}(t)$ and $\underline{x}(t)$ are (strictly) increasing in t. This follows since $\overline{x}(t) = sup\ \{x \in X\colon f(x; t) \geq x\}$, $\underline{x}(t) = inf\ \{x \in X\colon f(x; t) \leq x\}$, and f is (strictly) increasing in t. It is worth remarking that as t varies, the number of equilibria may change, but still the largest and the smallest equilibrium will be increasing in t.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

3. On the Cournot and Bertrand oligopolies and the theory of supermodular games*

Rabah Amir

1 INTRODUCTION

After Cournot's (1838) path-breaking work, and Bertrand's (1883) well-known critical book review, the field of oligopoly theory remained largely dormant for a long time. This survey will trace its rebirth to studies starting in the 1960s, but will focus mostly on the more recent developments associated with the birth of modern industrial organization theory in the 1980s, in particular on the literature linking Cournot oligopoly to the theory of supermodular games.

These games, also known as the class of games with strategic complementarities, turn out to encompass many of the commonly studied games in industrial organization and applied micro-economics, often under broadly satisfied and economically meaningful assumptions on the primitives of these models.[1] The class of supermodular games has played an important unifying role at the interface between applied game theory and industrial organization. Among the important general properties of such games are first the existence of pure-strategy Nash equilibrium, and the lattice structure of the equilibrium set, in particular the existence of minimal and maximal equilibrium points.[2] The latter extremal bounds also determine the range of equilibrium behavior for such games according to several solution concepts, such as mixed-strategy Nash equilibrium, correlated equilibrium, and rationalizable strategies. Another property is that the same bounds determine the relevant interval for the convergence of a wide class of adaptive learning algorithms, including best-reply Cournot dynamics and fictitious play. Finally, for parametrized classes of supermodular games, a simple additional complementarity assumption leads to clear-cut comparative statics properties of Nash equilibrium strategies with respect to parameters.[3] Early on, industrial economists understood the importance of the nature of strategic interaction, expressed in terms of relationships of strategic complementarity or substitutability, in many representative problems of industrial organization (see e.g., Bulow, Geanakoplos, and Klemperer, 1985, and Fudenberg and Tirole, 1984).

Under a broad set of economically meaningful assumptions, the Bertrand oligopoly is a game of strategic complements (or a supermodular game). On the other hand, Cournot

* It is a pleasure to acknowledge helpful conversations over the years on the theory of oligopoly, broadly construed, with Claude d'Aspremont, Francis Bloch, Rodolphe Dos Santos Ferreira, Jacques Dreze, David Encaoua, Christian Ewerhart, Jean Gabszewicz, Val Lambson, Yassine Lefouili, Laurent Linnemer, Diego Moreno, Jean-François Mertens, Heracles Polemarchakis, Jacques Thisse, and Xavier Vives.

[1] Introduced by Topkis (1978, 1979), this class of games was subsequently thoroughly investigated by Vives (1990), Milgrom and Roberts (1990, 1994) and Milgrom and Shannon (1994). There are also book treatments by Topkis (1998) and Vives (1999), as well as general surveys by Amir (2005) and Vives (2005a, 2005b).

[2] In addition, mixed-strategy Nash equilbria of supermodular games are unstable under a broad class of learning dynamics (Edlin and Echenique, 2004), just as is the case for general coordination games (e.g., Battle of the Sexes).

[3] This comparative statics result forms the basis for the modern treatment of the correspondence principle that goes back to Samuelson (1947): see Echenique (2002) for details.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

oligopoly is instead a game of strategic substitutes (or a submodular game) under broad conditions, such as the log-concavity of inverse demand. Unfortunately, the class of games of strategic substitutes need not possess any of the aforementioned strong properties of supermodular games. In particular, even the existence of pure-strategy Nash equilibrium is not guaranteed. Fortunately, the Cournot game is also an aggregative game, in the sense that a firm's payoff depends only on its actions and on the sum of all rivals' actions. The conjunction of the aggregation and strategic substitutes properties is sufficient to guarantee existence of Cournot equilibrium, via a construction that goes back to Selten (1970) and Szidarovszki (1970). For the special case of duopoly, the Cournot game may be viewed as a supermodular game, with the added benefit that all the nice properties then apply directly. However, this special property of the duopoly case does not extend to n-firm oligopoly in general.

As the existence results dealing with the Cournot game as a game of strategic substitutes were developed sequentially, the presentation of the results follows the same chronological manner. At the same time, we shed light on the connections between them to the greatest extent.

The other strand of literature on Cournot oligopoly reviewed here deals with the effects of entry on equilibrium prices and profits. The main question is to investigate, in as much generality as possible, when more competition in the sense of more identical firms entering the market leads to lower or higher Cournot equilibrium prices. It turns out that the attending existence question for symmetric Cournot equilibria is of interest in its own right here, in light of the fact that the key property that governs the said comparative statics question is whether the slopes of the reaction curve of a firm are globally lower than -1. The same issue arises naturally when one tries to settle the associated existence question in as much generality as possible. Thus for this part, the key properties of strategic substitutes and strategic complements do not play a role in the analysis.

Recently, new attention has been devoted to the broad class of aggregative games, often satisfying strategic substitutes as well. These are games with the property that a player's payoff depends on own action and on just an aggregate of other players' actions.[4] Since this class of games includes Cournot oligopoly, some of the results extend and unify those derived for Cournot oligopoly. Nevertheless, this strand of literature is beyond the scope of this survey.

For Bertrand competition with differentiated products, supermodularity of the game follows under the conditions of convex costs and log-supermodular demand. The latter condition enjoys the exact economic interpretation that the price elasticity of demand for one good is increasing in the price of any other good. Under these natural assumptions, the Bertrand model satisfies all the general properties of supermodular games.

We investigate the connection between the relationship of the differentiated products (gross substitutes or complements) on the one hand, and the strategic complementarity or substitutability of the associated Bertrand game on the other. Recall that in the common special case of linear demand, gross substitute goods translate exactly to a game of strategic complements while gross complements are equivalent to the game of strategic substitutes (e.g., Singh and Vives, 1984). However, for non-linear demand, this one-to-one connection

[4] See Corchón (1994), Anderson and Renault (2003), Dubey, Haimanko, and Zapelchelnyuk (2006), Acemoglu and Jensen (2009), Jensen (2010), Amir, Garcia, and Knauff (2010), Cornes and Hartley (2012), Roy and Sabarwal (2012) and Anderson, Erkal, and Piccinin (2014), among others.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

need not hold. It is then natural to explore the extent to which, say, gross complements could be compatible with strategic complements.

This survey is organized as follows. In Section 2, we begin with a basic review of the main tools of supermodular optimization in a simplified setting adapted to the needs of this review. In Section 3, we summarize the main existence results for general Cournot oligopoly. In Section 4, the comparative statics of entry and the existence of symmetric equilibrium are tackled together. Section 5 is devoted to the Bertrand model. A brief conclusion forms Section 6.

2 A SIMPLIFIED VERSION OF TOPKIS'S MONOTONICITY THEOREM

This section provides a simplified exposition of Topkis's framework in the special case where both the parameter and the decision sets are real intervals. While this is a very special case of the general theory, it will result in substantial gains in accessibility of this material to a wide readership, while at the same time covering our oligopoly applications.

Topkis considered the following parametrized family of constrained optimization problems,[5] where $A_s \subset A$, with the intent of deriving sufficient conditions on the objective and constraint set that yield monotone optimal solutions:

$$a^*(s) = \arg\max\{F(s,a) : a \in A(s)\}. \tag{3.1}$$

For our purposes, we shall consider S (the parameter set) and A (the action set) as real intervals.

A function $f : S \times A \to R$ has (strictly) increasing differences in (s,a) if

$$f(s,a) - f(s,a')(>) \geq f(s',a) - f(s',a'),\ \forall a > a', s > s' \tag{3.2}$$

or in other words if the difference $f(\cdot,a) - f(\cdot,a')$ is an increasing function.[6] If f is smooth, (3.2) is equivalent to $\partial^2 f(a)/\partial a \partial s \geq 0$. There are no restrictions on partials of the form $\partial^2 f(a)/\partial a_i^2$.

Increasing differences are interpreted as formalizing the notion of (Edgeworth) complementarity, i.e., higher values in any subset of the decisions or of the parameters, respectively, increase the marginal returns to higher values in the remaining decisions.

A version of Topkis's Monotonicity theorem is now given. Though a special case of the original result, it is adequate for the applications covered here.

Topkis's theorem *Assume that*:

(i) *F has increasing differences in (s,a); and*
(ii) *$A(s) = [h_i(s), g_i(s)]$ where $h_i, g_i : S \to R$ are increasing functions with $h_i \leq g_i$.*

[5] It is assumed throughout that F is upper semi-continuous in a, so the maximum in (3.1) is attained.
[6] Throughout this chapter, a function $f : S \to R$ is increasing (strictly increasing) if $x \geq y \Rightarrow f(x) \geq (>) f(y)$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Then the maximal and minimal selections[7] *of $a^*(s)$ are increasing functions. Furthermore, if (i) is strict, then every selection of $a^*(s)$ is increasing.*

Proof Let $\overline{a}$ be the maximal selection of a^* and assume it is not increasing. Then for some $s > s'$, $\overline{a}(s) < \overline{a}(s')$. Consider, as *(ii)* implies that $\overline{a}(s) \in A(s)$ and $\overline{a}(s') \in A(s')$,

$$\begin{aligned} 0 &\geq F[s, \overline{a}(s')] - F[s, \overline{a}(s)], \text{ since } \overline{a}(s) \in a^*(s) \\ &\geq F[s', \overline{a}(s')] - F[s', \overline{a}(s)], \text{ by increasing differences} \\ &\geq 0 \text{ , since } \overline{a}(s') \in a^*(s'). \end{aligned}$$

This implies that equality holds throughout, and thus that $\overline{a}(s') \in a^*(s)$. This contradicts the fact that $\overline{a}(s)$ is the largest argmax at s. Hence, $\overline{a}(s) \geq \overline{a}(s')$. If *(i)* is strict, the middle inequality above is strict, yielding a contradiction (as this shows $0 > 0$). A similar argument works for the minimal selection $\underline{a}(s)$. ■

Thus, in the one-dimensional case with smoothness, it is sufficient for monotone comparative statics that the objective satisfy $\partial^2 F(s,a)/\partial a \partial s \geq 0$ and the constraint set be a compact interval whose end points increase in the parameter.

The main advantages over the classical method, based on applying the implicit function theorem to the first-order conditions, are that (a) no form of concavity is required, and (b) the argmax may switch between a boundary and an interior solution without prejudice to its global monotonicity.

There is an order-dual to Topkis's theorem: if one changes only "increasing differences" to "decreasing differences" and h_i and g_i to decreasing functions in Topkis's theorem, then the conclusion is that the extremal selections of $a^*(s)$ are decreasing functions.[8] Likewise, the same duality applies to the strict statement in Topkis's theorem.

A further strengthening that is useful below is proved in Amir (1996b) and Edlin and Shannon (1998). If one adds that $F(s,a)$ is C^1 in a, and changes "increasing differences" to the condition that $\partial F(s,a)/\partial a$ is strictly increasing in s, then the conclusion is that, if interior, every selection of $a^*(s)$ is strictly increasing.

An alternative way to think of the dual statements is that it can be derived from the original statement simply by considering the parameter to be $-s$ instead of s. In other words, increasing in $-s$ is the same as decreasing in s. This connection will prove useful in some applications below.

The fixed-point result used to prove existence of pure-strategy Nash equilibrium, once monotonicity of the best-response mapping is established via Topkis's theorem, is due to Tarski (1955):

Tarski's theorem *Let F be an increasing mapping from $X_{i=1}^n [a_i, b_i]$ to itself. Then the set of fixed points of F is a non-empty complete lattice.*

---

[7] These are $\overline{a}^*(s) \triangleq \max\{a^*(s)\}$ and $\underline{a}^*(s) \triangleq \min\{a^*(s)\}$ and are always well defined when F is u.s.c. (upper semi-continuous) in a.

[8] Writing in the tradition of the operations research literature, Topkis (1978) actually considered the problem of *minimizing* a submodular objective (one with decreasing differences).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

3 THE COURNOT OLIGOPOLY

This section provides an overview of the literature on Cournot oligopoly, with an emphasis on results on existence, the submodularity of Cournot oligopoly, and properties of Cournot equilibrium as they relate to the theory of supermodular games.

3.1 The Basic Model

Consider a market for a homogeneous product characterized by an inverse demand function $P(Q)$, where Q is the total output of the n firms that serve this market. Firm i produces output level q_i with a cost function $C_i(q_i)$, and $Q = \sum_{i=1}^{n} q_i$. Let Q_{-i} denote the total output of the firms other than firm i, i.e., $Q_{-i} = Q - q_i$.

The profit function of firm i is

$$\Pi_i(q_i, Q_{-i}) = q_i P(q_i + Q_{-i}) - C_i(q_i). \tag{3.3}$$

Standard assumptions The following are in effect throughout this section:

(A1) $P(\cdot)$ *is twice continuously differentiable and* $P'(\cdot) < 0$.
(A2) $C(\cdot)$ *is lower semi-continuous and increasing on* $(0, \infty)$.

We shall assume throughout that firm i chooses an output $q_i \in [0, K_i]$, either due to a production capacity or a boundedness property of P (relative to C_i). While some results rely on this boundedness assumption for convenience, it is redundant for the main existence results.[9] The reaction correspondence of firm i is defined by $r_i(Q_{-i}) \triangleq \arg\max \{\Pi_i(q_i, Q_{-i}) : q_i \in [0, K_i]\}$.

Although we assume smoothness of the inverse demand function throughout, we stress that this is only for convenience and ease of interpretation of the needed sufficient conditions.

3.2 A Brief Early Literature Review

Starting with Hahn (1962) and Frank and Quandt (1963), a rich literature has developed over the years dedicated to the existence of Cournot equilibrium in the general case (with asymmetric firms), and to related issues of uniqueness or characterization of equilibrium as well. All the studies that comprise the early literature derived existence from the classical topological approach based on fitting the assumptions of the Brouwer/Kakutani fixed-point theorem. In other words, the assumptions on the primitives of the Cournot oligopoly model included some version of concavity of either inverse demand or the revenue function, and convexity of the cost functions, together with some condition that guarantees bounded outputs. Under such assumptions, the best-response mapping is continuous and existence follows directly from the Brouwer/Kakutani fixed-point theorem. This approach to existence and uniqueness of pure-strategy Nash equilibrium for general strategic games with real action sets was pioneered by Rosen (1965).

A list of the studies that are part of this literature on Cournot oligopoly include Frank and Quandt (1963), Szidarovszky (1970), Szidarovszky and Yakowitz (1977, 1982), Novshek

[9] See Ewerhart (2014) for more details on this point.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

(1980, 1984), and Roberts and Sonnenschein (1977), among others.[10] The literature dealing more specifically with existence and the comparative statics of entry in the symmetric case is discussed in the next section. Finally, several surveys and books offer some treatment of related issues in oligopoly theory: see Friedman (1977, 1982), Shapiro (1989), and Vives (2005a, 2005b).

3.3 Existence of Cournot Equilibrium in the General Case

The first existence result, due to Novshek (1985), can be viewed as being directly based on the supermodular approach, even though the author was then unaware of these methods, and thus had to derive all the needed results from first principles.[11]

Proposition 1 *Assume that inverse demand satisfies*

$$P'(Q) + QP''(Q) \leq 0 \text{ for all } Q \geq 0. \tag{3.4}$$

Then the Cournot game is a game of strategic substitutes and a Cournot equilibrium exists.

Proof Using the smooth characterization of decreasing differences, $\Pi_i(q_i, Q_{-i})$ has this property if and only if

$$\partial^2 \Pi_i(q_i, Q_{-i})/\partial q_i \partial Q_{-i} = P'(q_i + Q_{-i}) + q_i P''(q_i + Q_{-i}) \leq 0 \text{ for all } q_i, Q_{-i} \geq 0 \tag{3.5}$$

Since with $P' \leq 0$, it is easy to see that (3.5) holds if and only if (3.4) holds. It follows that the game is submodular if (3.4) holds. This also holds in the n-firm case, for all n.

In addition to being a submodular game, the Cournot model is an aggregative game, i.e., each firm's profit depends only on its output and on the aggregate output of the rivals. These two properties together imply existence of a Cournot equilibrium, via a well-known result (see Novshek, 1985, or Kukushkin, 1994). ■

In terms of economic interpretation, Condition (3.5) says that a firm's marginal revenue decreases with rivals' aggregate output. This is quite a natural condition for Cournot competition.

Condition (3.4) has been widely used in the literature, and is quite general. Nevertheless, an important class that does not satisfy this condition is the class of hyperbolic inverse demands

$$P(Q) = \frac{a}{Q^\alpha}, \text{ for } a > 0, \alpha > 0. \tag{3.6}$$

[10] Szidarovszky and Okuguchi (2008) deal with existence in a generalized version of a standard Tullock rent-seeking game, but this is closely related to the Cournot model.

[11] Novshek's work represents one of the early instances where an author of an important advance in economic theory discovers the powerful results from the theory of supermodular games in a specific context, independently of the work of Topkis on parametric monotonicity. Another example is the theory of convex (coalitional) games (Shapley, 1971).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Importantly, there is no requirement whatsoever on the cost functions, besides lower semi-continuity. (The latter is needed only to ensure that the profit functions are upper semi-continuous so that the reaction correspondences are non-empty valued.) This irrelevance of the nature of returns to scale in production to the existence of Cournot equilibria underscores the most important of the novel benefits of using the lattice-theoretic approach over the standard topological approach. Indeed, the latter typically requires convex costs.

It is worth recalling here that, in contrast to supermodular games, in general, a submodular game, or a game of strategic substitutes, need not possess a pure-stategy Nash equilibrium. In other words, the aggregation property mentioned in the proof above is critically needed, in conjunction with the submodularity of the Cournot game to guarantee existence.

This special argument, exploiting the special structure of the Cournot model, has a long history. Novshek (1985) rediscovered an argument that exploits the aggregation and submodular properties to establish existence of a Cournot oligopoly, originally given in Selten (1970) and Szidarovszky (1970); also see Bamon and Fraysse (1985). For a more formal proof in the form of a fixed-point theorem, see Kukushkin (1994).[12]

A noteworthy special case is duopoly or $n = 2$. For any two-player submodular game, the existence of a pure-strategy Nash equilibrium is guaranteed via Tarski's fixed-point theorem applied to the composition (of the largest selections, say) of the two reaction curves (Vives, 1990). The argument is as follows. Since both reaction curves are decreasing, their composition is increasing. Its fixed points are easily seen to be Nash equilibria.[13] In addition, if one reverses the order on (say) firm 2's decision, i.e., if one thinks of its action as being $-q_2$ instead of q_2, then under (3.4), the Cournot duopoly becomes a supermodular game. This "order-reversing" trick does not work for three or more firms, i.e., in general, a submodular n-player game is not a supermodular game if $n \geq 3$. It follows that it is only for the special case of two players that the Cournot duopoly under Condition (3.4) enjoys all the strong properties of supermodular games.[14]

The next result is due to Amir (1996a):

Proposition 2 *Assume that inverse demand $P(Q)$ is log-concave or satisfies*

$$P(Q)P''(Q) - P'^2(Q) \leq 0 \textit{ for all } Q \geq 0. \tag{3.7}$$

Then the Cournot game is a game of strategic substitutes and a Cournot equilibrium exists.

The proof, given in the Appendix, shows that Condition (3.7) is sufficient for the Cournot game to be of strategic substitutes, by showing that each firm's profit function satisfies the dual single-crossing condition in own and rivals' outputs.[15] Existence of PSNE then follows

---

[12] Dubey et al. (2008) use another argument and arrive at the same existence result.

[13] This argument applies to Cournot duopoly in particular.

[14] These include the fact that uniqueness of pure-strategy Nash equilibrium (PSNE) implies its coincidence with the mixed-strategy Nash equilibrium (MSNE) set and the correlated equilibrium set, the convergence of a wide class of learning models to the PSNE set, etc.

[15] Partly as a reflection of the absence of any convenient differential test, the single-crossing property can often be quite tedious to prove even in simple settings. The proof that log-concavity of inverse demand implies downward-sloping reaction curves is an elementary proof, but nevertheless one that has inspired similar proofs in other contexts (see e.g., Quah and Strulovici, 2012, and Amir and Lazzati, 2011). It is given here for the sake of completeness and as an illustration of how a proof of the single-crossing property might look.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

from the conjunction of the strategic substitutes and the aggregation properties of the Cournot oligopoly.

The economic interpretation of Condition (3.7) is that a firm's price elasticity of demand decreases with rivals' aggregate output. This is also a natural and general condition for Cournot competition. Like Condition (3.4), it includes as special cases the common linear inverse demand, concave inverse demand, and even allows for limited forms of convexity of inverse demand. Nevertheless, the class of hyperbolic inverse demands also fails Condition (3.7). In fact, these demands satisfy the opposite property of log-convexity, which (under smoothness) is equivalent to

$$P(Q)P''(Q) - P'^2(Q) \geq 0 \text{ for all } Q \geq 0. \tag{3.8}$$

It is natural to wonder whether Conditions (3.7) and (3.4) are nested. The answer is negative. Indeed, $P(Q) = \min\{-\log Q, 0\}, Q \geq 0$, satisfies (3.4) but not (3.7), and $P(Q) = e^{-Q}, Q \geq 0$, satisfies (3.7) but not (3.4). The latter inverse demand also constitutes the limit case for the property of log-concavity, since it is log-linear. We shall return to this point below.

Finally, while the two previous results provide alternative sufficient conditions on inverse demand alone that ensure strategic substitutability for the Cournot game, another condition on costs (strong concavity) can lead to the same outcome, as will be seen in the next section.

3.4 Some Extensions

In addition to the central question of existence via the property of strategic substitutes, Amir (1996a) addressed a number of other issues of some interest, which are summarized next.

3.4.1 Cournot oligopoly as a game of strategic complements

The first of these issues is whether n-player Cournot oligopoly could ever be a game of strategic complements (in the original order on the output sets). The answer is clearly negative if one uses the cardinal condition of increasing differences. Indeed, the latter condition would entail that Condition (3.5) holds with the reverse sign, which is impossible (to see this, simply let $q_i = 0$ in (3.5)). Nevertheless, the following result still holds:

Proposition 3 *Assume that inverse demand is log-convex or satisfies (3.8), and that $C_i(q_i) \equiv 0$ for all i. Then the Cournot game is a game of strategic complements and a Cournot equilibrium exists.*

Proof Since $\log \Pi_i(q_i, Q_{-i}) = \log q_i + \log P(q_i + Q_{-i})$, it is easy to see that $\log \Pi_i(q_i, Q_{-i})$ has increasing differences in (q_i, Q_{-i}) if and only if $\log P(\cdot)$ is convex.[16] ■

Log-convexity is a very restrictive condition for an inverse demand function, though one that is nevertheless satisfied by the subclass of hyperbolic demand functions (3.6).

A very strong requirement here is the absence of (non-trivial) production costs. To see why this restriction is needed, consider the following two observations, made in the context of

[16] A formal proof of this simple fact appears in Amir (1996c).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

linear costs (for simplicity), $C_i(q_i) = c_i q_i$. First, $P(\cdot) - c_i$ can never be a log-convex function when $c_i > 0$.[17] Second, when facing aggregate output Q_{-i} by its rivals, for firm i, the outputs outside $[0, P^{-1}(c_i) - Q_{-i}]$ are easily seen to be dominated strategies (by the 0 output choice). Since one may view the firm as facing the constraint set $[0, P^{-1}(c_i) - Q_{-i}]$ when choosing its output, and since this constraint set is descending in Q_{-i}, firm i's reaction curve cannot be globally increasing (or, it would hit the constraint and be decreasing thereafter).

In other words, the presence of non-trivial production costs always pushes in favor of the Cournot game having strategic substitutes. If $c = 0$, there is a complete duality in the way log-concavity and log-convexity assumptions on inverse demand translate into a decreasing and an increasing reaction curve, respectively. However, whenever $c > 0$, this duality breaks down, with costs always introducing a bias in favor of downward-sloping reactions. Indeed, $P(\cdot) - c$ is log-concave whenever $P(\cdot)$ is log-concave while the converse need not hold. In fact, $P(\cdot) - c$ may be globally log-concave even when $P(\cdot)$ is globally log-convex.

Here is a simple instructive example:

Remark 1 The inverse demand $P(Q) = ae^{-Q}, a > 0, Q \geq 0$, is log-linear, thus both log-concave and log-convex. It follows from the two propositions above that, if a firm has zero costs, its reaction curve will be both increasing and decreasing, and hence constant. Indeed, it can be verified via a simple computation that the reaction curve is constant at 1; in other words, each firm has a *dominant strategy* to produce one unit.

We close with a simple reformulation of the previous result. The idea here is to bring production costs back into the picture, but limit the effective joint action space in such a way that the failure of global strategic complementarity ceases to be binding:

Proposition 4 *Assume that each firm's cost function is linear, or $C_i(q_i) = c_i q_i$, for $i = 1, 2, \ldots, n$, and that inverse demand is such that $P(\cdot) - \max_i c_i$ is a log-convex function on $\left[0, \sum_{i=1}^{n} K_i\right]$. Then the Cournot game is a game of strategic complements and a Cournot equilibrium exists.*

Although it is certainly of interest to know that the n-firm Cournot oligopoly can be a game of strategic complements in the original order on the output sets, the sufficient conditions leading to this case are quite restrictive. This is so for the restrictiveness of the log-convexity of inverse demand, and in particular by the absence of non-trivial costs. When the assumptions do hold, the Cournot game satisfies all the strong properties that characterize the class of supermodular games.

An ancillary benefit of the foregoing analysis is that it provides a precise sense in which the oft-expressed view, that Cournot oligopoly is naturally a game of strategic substitutes, is justified.

3.4.2 Taking monotone transformations

Since taking strictly monotone transformations of a payoff function will not change the reaction correspondence, it is natural to explore the scope of complementarity that can be

[17] Indeed, $\log P(\cdot) - c_i$ convex amounts to $\left[P(Q) - c_i)P''(Q) - P'^2(Q)\right] \geq 0$ for all $Q \geq 0$, which cannot hold (to see this, let $Q = P^{-1}(c_i)$).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

spanned by taking suitably chosen transforms of a Cournot profit function. Amir (2005) considers the effects of the following choices of transformations:

$$x^{\alpha}/\alpha, \alpha \in R; \;\log(ax+b), a>0, b>0; \; -e^{-ax}, a>0; \text{ and } ax + logx, a > 0.$$

The main idea is easily conveyed via an example:

Example 1 Consider a Cournot firm with $C(q) = 0$ and

$$P(Q) = 10e^{-Q} + 3 - Q, \text{ for } Q \leq 3.35 \text{ and } P(Q) = 0 \text{ for } Q > 3.35.$$

Then $\Pi_i(q_i, Q_{-i}) = 10q_i\left[e^{-(q_i+Q_{-i})} + 3 - q_i - Q_{-i}\right]$. It is easy to verify that

$$P'(Q) + QP''(Q) = 10(x-1)e^{-Q} - 1 > 0 \text{ if and only if } 1.41 < Q < 2.99$$

and

$$P(Q)P''(Q) - P'^2(Q) = -10(x-1)e^{-Q} - 1 > 0 \text{ if and only if } Q < 0.783.$$

Hence, we conclude that Conditions (3.7) and (3.4) both fail to hold globally, so we cannot conclude that the reaction curve is downward-sloping.

Next apply the transformation $h(x) = x + logx$ to $\Pi_i(q_i, Q_{-i})$, and take the cross-partial to get

$$\frac{\partial^2 h \circ \Pi_i(q_i, Q_{-i})}{\partial q_i \partial Q_{-i}} = 10(x-1)e^{-Q} - 1 - 10(x-1)e^{-Q} - 1 = -2 < 0.$$

Since $h \circ \Pi_i(q_i, Q_{-i})$ has decreasing differences, we conclude that $\Pi_i(q_i, Q_{-i})$ has the dual single-crossing property, and therefore, from Topkis's theorem, that the reaction curve is decreasing.

3.4.3 The special case of linear costs

The reason Conditions (3.7) and (3.4) are not nested is that they are both imposed globally, i.e., for all possible total outputs. Yet in general, along a firm's optimal response, not all total outputs will be reachable. Exploring this observation for the the case of linear costs (for simplicity), one arrives at a necessary and sufficient condition for strategic substitutes for Cournot oligopoly (Amir, 2005):

Proposition 5 *Assume that each firm's cost function is linear, or $C_i(q_i) = c_i q_i$, for $i = 1, 2, \ldots, n$. Then the Cournot game is a game of strategic substitutes if and only if inverse demand is such that $P(\cdot) - \min_i c_i$ is a log-concave function on $\left[q_i^m, +\infty\right)$, for all i, where q_i^m denotes firm i's optimal monopoly output.*

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

A formal proof is omitted here (see Amir, 2005), but the main idea behind the proof is provided next.[18] The reason that total outputs in the interval $[0, q_i^m)$ are not reachable along a firm's best response is that the reaction curves have all their slopes larger than -1, which implies that $r_i(Q_{-i}) + Q_{-i}$ is increasing in Q_{-i} and the smallest reachable total output is then $r_i(Q_{-i}) = q_i^m$. We shall return to this property as it plays a key role in the analysis of symmetric Cournot oligopoly in the next section.

We close with a simple closed-form illustrative example of how the shape of a reaction curve depends on the size of the firm's unit cost:

Example 2 Let $P(Q) = \frac{1}{Q+1}$ and $C_i(q_i) = c_i q_i$.

If $c_i = 0$, since P is log-convex, the reaction curve r_i is increasing. As c_i is increased, r_i is increasing up to some value of $Q_{-i} > 0$, but is decreasing thereafter. For $c_i = 1/4$, the reaction curve is easily computed to be

$$r_i(Q_{-i}) = 2(Q_{-i} + 1)^{1/2} - Q_{-i} - 1 \text{ if } Q_{-i} \leq 3 \text{ and } r_i(Q_{-i}) = 0 \text{ if } Q_{-i} > 3.$$

It is easy to see that $r_i(Q_{-i})$ is globally decreasing. (In fact, $1/4$ is the smallest value of c_i for which the reaction curve is decreasing. All higher values of c_i will keep $r_i(Q_{-i})$ decreasing.)

The key point for understanding this outcome is that $P(Q) - 1/4$ is a log-concave function of Q on (q_i^m, ∞), even though $P(Q)$ is a log-convex function. Indeed, to verify, note that $q_i^m = \arg\max\{q_i/(q_i + 1) - q_i/4\} = 1$ here, and

$$\left[P(Q) - \frac{1}{4}\right] P''(Q) - P'^2(Q) = \frac{(1 - Q)}{2(Q + 1)^4} < 0 \text{ for } Q > 1 = q_i^m.$$

3.5 Generalized Concavity

Introduced into oligopoly theory with product differentiation by Caplin and Nalebuff (1991) and Dierker and Podczek (1992) as a way to model consumer heterogeneity, generalized concavity turned out to be relevant to the existence of Cournot equilibrium. Ewerhart (2014) shows that this notion can be used fruitfully to derive a significant extension of the previous existence results based on the lattice-theoretic approach, for general cost functions.

We begin with the definition of the general notion of generalized concavity. With $\varphi_\alpha(x) = x^\alpha/\alpha$ being a class of monotone transformations as α varies, an inverse demand function $P(Q)$ is (α, β) bi-concave if the mapping from $\varphi_\beta(Q)$ to $\varphi_\alpha(P(Q))$ is concave,[19] or $d\varphi_\alpha(P(Q))/d\varphi_\beta(Q)$ is decreasing, or, for all $Q \geq 0$,

$$\Delta_{\alpha,\beta}(Q) \triangleq (\alpha - 1)P'^2(Q) + QP(Q)P''(Q) + (1 - \beta)P(Q)P'(Q) \leq 0. \quad (3.9)$$

Several of the known conditions in the literature on existence of Cournot equilibrium can be nested in (3.9).

[18] Actually, one can write a simple proof of this characterization using the implicit function theorem and the first-order conditions along standard lines.

[19] In words, bi-concavity means that inverse demand becomes concave once the quantity and the price axis are scaled by the transformations φ_β and φ_α.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

The main existence result in Ewerhart (2014) is as follows:

Proposition 6 *Let $P(Q)$ be $(\alpha, 1-\alpha)$-bi-concave for some $\alpha \in [0,1]$. Then the Cournot oligopoly has strategic substitutes and a Cournot equilibrium exists.*

The proof makes use of the scaling maps φ_α on the price axis and φ_β on the quantity axis, and works with the revenue function directly to show that $(\alpha, 1-\alpha)$-bi-concavity of $P(\cdot)$ leads to a profit function that satisfies the dual single-crossing property, provided the cost function is increasing. It follows that the reaction correspondence then has a decreasing minimal selection (see Ewerhart, 2014, p. 45 for details).

It is straightforward to verify that this result generalizes previous existence results on existence (based on strategic substitutes). In particular, Conditions (3.4) and (3.7) can be seen to amount to $(1,0)$ and $(0,1)$ bi-concavity respectively.

We close this section with a simple class of examples illustrating the value added of this existence result:

Example 3 Let $P(Q) = (1-Q^\delta)^{1/\gamma}$ for $Q \leq 1$ and 0 for $Q > 1$.

It can be shown (numerically) via the sign of $\Delta_{\alpha,\beta}(Q)$ for this inverse demand function that $P(Q)$ is (α,β) bi-concave if and only if $\alpha \leq \gamma$ and $\beta \leq \delta$. However, Conditions (3.4) and (3.7) can be seen not to hold if δ and γ are less than, but very close to, 1. Nevertheless, according to the proposition, the reaction curve is downward-sloping here.

For this example, a closed-form solution for the reaction curve is not possible in general.

3.6 On Uniqueness of Cournot Equilibrium

Many of the papers on Cournot oligopoly offered some insights on uniqueness of Cournot equilibrium. Kolstad and Mathiesen (1987) restricted attention to interior and regular Cournot equilibria and use degree theory to derive conditions for uniqueness (see Gaudet and Salant, 1991, for a reformulation). As a summary of the main result by Kolstad and Mathiesen (1987), the uniqueness condition amounts to signing the determinant of the Jacobian of marginal profit, which reduces to the key condition that, at all candidate Cournot equilibria, and including firms with a positive output, one has (with C_i assumed smooth)

$$\left[1 - \sum_i \frac{P' + q_i P''}{C_i'' - P'}\right] \Pi_{i=1}^n (C_i'' - P') > 0.$$

Integrating generalized concavity with the supermodularity approach, Ewerhart (2014) provides an alternative approach and result on uniqueness, making use of the Selten-Novshek aggregation argument again.

4 SYMMETRIC COURNOT OLIGOPOLY AND ENTRY

In this section, we consider the special case of symmetric Cournot oligopoly and consider the twin questions of existence of a Cournot equilibrium and the comparative statics of exogenous

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

entry of firms. For the latter issue, we investigate how industry price and per-firm profit vary with the number of firms in the industry. This amounts to conducting equilibrium comparative statics, the exogenous parameter being the integer number of firms.

4.1 A Brief Literature Review

The problem of existence of symmetric Cournot equilibrium has a long history. McManus (1962, 1964) proved existence of such equilibria assuming only that the cost function is convex. Without any type of concavity assumption, he showed that the reaction curve has all its slopes larger than -1, a property that rules out any downward jumps, but allows upward jumps as well as upward or downward-sloping continuous parts of the graph. He then shows that this property is sufficient for the existence of symmetric Cournot equilibrium in n-firm oligopoly. Unaware of McManus's work, Roberts and Sonnenschein (1976) provided a (more rigorous) version of the same result under the assumption of a linear cost function.

As noted in Vives (1990), the essence for this existence result is another theorem by Tarski (1955), one that was not known to the above authors at the time they were writing, and that has remained less known than Tarski's fixed-point theorem for a long time. Amir and Lambson (2000) used the full lattice-theoretic machinery explicitly to come up with a somewhat more general version of the result, exposited below in some detail. Hoernig (2003) extended their analysis to accommodate product differentiation in quantity competition.

As to the problem of the effects of exogenous entry on market performance, there is an extensive literature in oligopoly theory, in line with the question being such a central and historical part of the overall theory of markets, in particular in partial equilibrium analysis. A limited list of studies dealing specifically with the Cournot model would include McManus (1962, 1964), Frank (1965), Okuguchi (1973), Ruffin (1971), Seade (1980), Szidarovsky and Yakowitz (1982), and Quirmbach (1988), among many others. In addition, Dixit (1986) and Corchón (1994) address similar questions of comparative statics of equilibria in broader settings that include the Cournot model.[20]

The reason the issues of existence and comparative statics of symmetric Cournot equilibria are addressed together in this survey is that, with the lattice-theoretic approach, the minimal assumptions needed to derive definite conclusions about these two issues naturally lead to two distinct cases of analysis: an intuitive case where more competition leads to a lower industry price, and a counter-intuitive case where the opposite happens.[21]

4.2 The Basic Set-up

Since firms are identical, we write $C(q_i)$ for the cost function (assumed smooth here), and $r(Q_{-i})$, for the reaction correspondence of firm i. At equilibrium, these quantities will be indexed by the underlying number of firms n. We explicitly deal with the (possible) non-uniqueness of Cournot equilibria by considering extremal equilibria. Denote the maximal and minimal points of any equilibrium set by an upper and a lower bar, respectively. Thus, for instance, $\overline{Q}_n$ and $\underline{Q}_n$ are the highest and lowest total equilibrium outputs, with corresponding

[20] The problem of endogenous entry deals with somewhat different issues; see Mankiw and Whinston (1986), Suzumura and Kiyono (1987) and Amir, Koutsougeras, and De Castro (2014).

[21] De Meza (1985) provides an example indicating that simply signing the derivative of price with respect to the number of firms (treated as a real number) can lead to ambiguous outcomes.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

equilibrium prices $\underline{p}_n$ and $\overline{p}_n$, and per-firm profits $\underline{\pi}_n$ and $\overline{\pi}_n$, respectively. A priori, our comparative statics conclusions refer to the extremal equilibria, but they may be extended to what are called regular equilibria in Amir et al. (2014).[22]

Instead of the profit function (3.3), one may think of firm i as choosing total output $Q = q_i + Q_{-i}$, given the other firms' cumulative output Q_{-i}, in which case its profit can be rewritten as

$$\widetilde{\Pi}_i\left(Q, Q_{-i}\right) = \Pi\left(Q - Q_{-i}, Q_{-i}\right) = \left(Q - Q_{-i}\right) P\left(Q\right) - C\left(Q - Q_{-i}\right). \tag{3.10}$$

Its augmented reaction curve is then

$$R_i(Q_{-i}) \triangleq \arg\max\left\{\widetilde{\Pi}_i\left(Q, Q_{-i}\right) : Q \geq Q_{-i}\right\}.$$

The argmax's in (3.3) and (3.10) are always related by $R_i(Q_{-i}) = r_i(Q_{-i}) + Q_{-i}$. It follows that $R_i(Q_{-i})$ is increasing (decreasing) in Q_{-i} if and only if $r_i(Q_{-i})$ has slopes $\geq$ $(\leq)1$. This is a key observation in what follows.

A key determinant for both the existence and the comparative statics issues is the sign of $\triangle\left(Q, Q_{-i}\right)$, which denotes the cross-partial derivative of $\widetilde{\Pi}$ with respect to Q and Q_{-i}, i.e.,

$$\Delta\left(Q, Q_{-i}\right) \triangleq \frac{\partial^2 \widetilde{\Pi}_i\left(Q, Q_{-i}\right)}{\partial Q \partial Q_{-i}} = -P^{'}\left(Q\right) + C^{''}\left(Q - Q_{-i}\right).$$

Both $\widetilde{\Pi}$ and $\triangle$ are defined on (the lattice) $\varphi \triangleq \{(Q, Q_{-i}) : Q_{-i} \geq 0, Q \geq Q_{-i}\}$.

The following cumulative best-response mapping is the key object of study for both the existence and the comparative statics questions. Define

$$\begin{array}{rccc} B_n: & [0,(n-1)K] & \longrightarrow & 2^{[0,(n-1)K]} \\ & y & \longrightarrow & \frac{(n-1)}{n}\left[r_i(Q_{-i}) + Q_{-i}\right] \end{array}$$

The qualitative nature of both the existence and the comparative statics results hinges entirely on the global sign of Δ, so that we will distinguish two main cases: $\Delta > 0$ and $\Delta < 0$.[23]

4.3 The $\Delta > 0$ Case

As convexity of the cost function implies that $\Delta > 0$, this corresponds to the standard case commonly studied in the oligopoly literature. Indeed, the effects of entry on price and profit will be seen to correspond to standard intuition. Exploiting the symmetry assumption, both existence and regular comparative statics properties rest only on the one key assumption that $\Delta > 0$. Importantly, the upcoming analysis does not rely on the Cournot game being of strategic substitutes or complements. Yet the lattice-theoretic approach is crucial to the results below.

[22] This issue comes up naturally in any formal treatment or discussion of the correspondence principle (Samuelson, 1947 and Echenique, 2002).
[23] The condition $\Delta > 0$ has a long history in the oligopoly literature, going back to at least Hahn (1962).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

We begin with a simple existence result, which generalizes older work by McManus (1962, 1964) and Roberts and Sonnenschein (1976). The latter results assumed either linear or convex costs, while the condition $\Delta = -P^{'}(Q) + C^{''}(Q - Q_{-i}) > 0$ makes it apparent that it can accommodate some limited increasing returns to scale, even of a global sort, as illustrated in an example below. In terms of economic interpretation, the key condition $\Delta > (<)0$ is quite transparent; it holds that price decreases faster (slower) than marginal cost at all possible output levels.

Proposition 7 *When $\Delta > 0$, there exists at least one symmetric Cournot equilibrium and no asymmetric ones.*

Proof Since $\triangle(Q, Q_{-i}) > 0$, by a strengthening of Topkis's theorem given in Amir (1996b), every selection of $R(Q_{-i})$ is strictly increasing in Q_{-i}. By Tarski's fixed-point theorem applied (say) to the maximal selection $\overline{R}_i$, we conclude that $\overline{R}_i$ has a fixed point $\overline{y}$, which is easily seen to correspond to a symmetric Cournot equilibrium, since $\overline{B}(\overline{y}) = \frac{(n-1)}{n}[r(\overline{y}) + \overline{y}]$ implies that $r(\overline{y}) = \frac{\overline{y}}{n-1}$, which characterizes symmetric Cournot equilibria.

To see that no asymmetric Cournot equilibrium is possible here, recall that the symmetry of the game ensures that permutations of such an equilibrium are also equilibria. Hence, for some i, r_i must have a slope equal to 1, which implies a slope of 0 for R_i, a contradiction to the fact that $R_i(Q_{-i})$ is strictly increasing in Q_{-i}. ■

It is remarkable that both the existence and the intuitive comparative statics below entirely hinge on the fact that $\Delta > 0$:

Proposition 8 *When $\Delta > 0$, as n increases exogenously, for the extremal Cournot equilibria,*

(i) P_n *decreases;*
(ii) π_n *decreases.*

Proof *(i)* As n increases, the fraction $\frac{n-1}{n}$ increases, and hence the (maximal selection) mapping $\overline{B}(y) = \min\{B_n(y)\}$ shifts up (as a function of y). Hence the maximal fixed point, $\overline{y}_n$, increases in n. Since the largest equilibrium output $\overline{Q}_n = R_i(\overline{y}_n) = r_i(\overline{y}_n) + \overline{y}_n$ and $R_i(\cdot)$ is increasing, $\overline{Q}_n$ also increases in n. A similar proof applies to the minimal equilibrium output $\underline{Q}_n$.

(ii) Let $\overline{q}_n$ be the maximal Cournot equilibrium per-firm output, and consider

$$\begin{aligned}
\pi_n &= \overline{q}_n P(n\overline{q}_n) - C(\overline{q}_n) \\
&\geq \overline{q}_{n+1} P\left[\overline{q}_{n+1} + (n-1)\overline{q}_n\right] - C(\overline{q}_{n+1}) \\
&\geq \overline{q}_{n+1}\left\{P\left[\overline{q}_{n+1} + n\overline{q}_{n+1}\right] - C(\overline{q}_{n+1})\right\} \\
&= \overline{q}_{n+1}\left\{P\left[(n+1)\overline{q}_{n+1}\right] - C(\overline{q}_{n+1})\right\} \\
&= \pi_{n+1}
\end{aligned}$$

where the first inequality follows from the Cournot equilibrium property, the second from the facts that P is decreasing and $nx_n \leq (n+1)x_{n+1}$ since $\Delta > 0$.■

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Example 4 Let $P(Q) = 2 - Q, Q \leq 2$, and $C(q) = \log(q + 1)$. Even though C is globally concave, we have, on the relevant range $0 \leq Q_{-i} \leq Q \leq 2$,

$$\Delta\left(Q, Q_{-i}\right) = -P'(Q) + C''(Q - Q_{-i}) = 2 - (1/(Q - Q_{-i} + 1)^2) > 0.$$

A simple calculation shows that

$$r(Q_{-i}) = \frac{1}{4}\left[-Q_{-i} + \sqrt{Q_{-i}^2 - 8Q_{-i} + 8}\right] \text{ for } Q_{-i} \leq 1 \text{ and } 0 \text{ for } Q_{-i} > 1$$

It is easily checked that $r(Q_{-i}) > -1$ for $Q_{-i} \leq 1$, and that the unique equilibrium has $Q_n = \frac{n}{2(n+1)}\left[1 - n + \sqrt{n^2 + 2n + 5}\right]$, which is increasing in n (so price is decreasing in n). Hence, this case fits the usual situation indeed.

We next consider the counter-intuitive case.

4.4 The $\Delta < 0$ Case

The condition $\Delta < 0$ clearly requires the cost function to be strongly concave, and is thus quite restrictive. This case has generally been neglected in the oligopoly literature.

A key property of the reaction curve follows from Topkis's theorem: the augmented reaction curves $R_i(Q_{-i})$ is decreasing, which is the same as saying that, while interior, r satisfies

$$\frac{r(Q'_{-i}) - r(Q_{-i})}{Q'_{-i} - Q_{-i}} < -1, \text{ for all } Q'_{-i}, Q_{-i} \tag{3.11}$$

has all its slopes < -1. It follows that r_i is strongly decreasing, so the Cournot game here is of strategic substitutes.

The existence issue is quite distinct from the $\Delta > 0$ case. The same argument used in the previous section (Novshek, 1985, Kukushkin, 1994) shows abstractly that Cournot equilibria exist here. However, one can be far more explicit here and exhibit an important specific Cournot equilibrium: the monopoly outcome $(q^m, 0, 0, \ldots, 0)$. Indeed, this follows directly from the facts that $r_i(0) = q^m$, and $r_i(q^m) = 0$, with the latter being a direct consequence of (3.11).

As to the possible existence of other equilibria, since (3.11) is consistent with r_i having downward jumps, there may be none if all that is assumed is $\Delta < 0$. To see this, it helps to think of symmetric equilibria in the n-firm case as intersections of $r_i(\cdot)$ with the line $Q_{-i}/(n-1)$. A downward jump of r_i may well lead to the absence of such intersection.

To restore existence of a symmetric equilibrium for every n, one needs to assume that $\Pi_i(q_i, Q_{-i})$ is strictly quasi-concave in q_i, so that r_i is a continuous function. Then it is easily shown that the symmetric equilibrium is unique for every n, since they are intersections of $r_i(\cdot)$ with the line $Q_{-i}/(n-1)$. Furthermore, the only other equilibria are of the form $(q_m, q_m, \ldots, q_m, 0, 0, .., 0)$ for any $m < n$. In other words, each equilibrium of the latter sort involves some number of inactive firms and symmetric behavior among the active firms.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

In light of the structure of the two types of equilibria here, as far as comparative statics properties are concerned, we can without loss of generality confine attention to the symmetric equilibria:

Proposition 9 *When $\Delta < 0$, as n increases, for the unique symmetric Cournot equilibrium,*

(i) P_n *increases;*
(ii) π_n *decreases.*

Proof *(i)* We first argue that, since $\Delta = \partial^2 \widetilde{\Pi}_i (Q, Q_{-i})/\partial Q \partial Q_{-i} < 0$, a combination of Topkis's theorem and an additional argument using the structure of the Cournot model implies that $R_i(\cdot)$ is decreasing and thus that $r_i(\cdot)$ has all slopes < -1, while interior, and that $r_i(\cdot)$ remains at zero once it reaches there (for key details needed because the feasible set $Q \in [Q_{-i}, \infty)$ is ascending, instead of descending, and thus Topkis's theorem is not applicable (see Amir and Lambson 2000, Lemma 3.1, p. 250). The same argument as in the previous result leads to the fixed point, y_n, which is unique here, being increasing in n. Hence, the equilibrium total output $Q_n = R_i(y_n)$, Q_n is decreasing in n, and the price P_n is increasing in n.
(ii) The proof that π_n decreases in n is the same as in the previous proposition. ■

In light of the fact that $\Delta < 0$ requires the cost function to be strongly concave, the quasi-concavity of each profit function in own output becomes a very restrictive condition. Indeed, in the literature on Cournot oligopoly, the latter assumption typically entails convexity of costs. Nevertheless, it is easy to check that the two assumptions are actually compatible, as can also be confirmed by many examples (see Amir and Lambson, 2000).

As previously noted, without the quasi-concavity assumption of each profit function in own output, existence of symmetric Cournot equilibrium may fail for some values of n. Nevertheless, the comparative statics results will hold for those values of n for which said existence holds.

Here is an illustrative example of the $\Delta < 0$ case:

Example 5 Let $P(Q) = 4 - 6Q, Q \leq 2/3$, and $C(q) = 3q - 3q^2 - 3q^3, q \leq K = 1/3$. We have, on the relevant range $0 \leq Q_{-i} \leq Q \leq 2/3$,

$$\Delta (Q, Q_{-i}) = -18(Q - Q_{-i}) \leq 0.$$

A simple calculation shows that $r(Q_{-i}) = \frac{1}{3}\left[1 - \sqrt{6Q_{-i}}\right]$ for $Q_{-i} \leq 1/6$, and 0 for $Q_{-i} > 1/6$. It is easily checked that $r(Q_{-i}) < -1$ for $Q_{-i} \leq 1/6$, and that the unique symmetric equilibrium has total output $Q_n = \frac{n}{3}\left(n - \sqrt{n^2 - 1}\right)$, which is decreasing in n (so price is increasing in n).

The conclusion that more competition increases price in a global sense is certainly one of the most provocative propositions in economics. Yet, it is also quite in line with basic economic intuition. Indeed, as n increases, each firm must produce much less at the new equilibrium (since $r(\cdot)$ falls steeply), and since the average cost curve falls very sharply

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

(when $\Delta < 0$), more competition leads to a sharp decline in production efficiency. Since firms must pass this on to consumers at a Cournot equilibrium, the end result is an increase in price.

Perhaps not surprisingly, the case of a concave cost function has been generally avoided in much of the literature. Yet, this case is certainly of some economic interest. For instance, the $\Delta < 0$ case may be viewed as an alternative definition of natural monopoly, based on both supply side and market considerations. The well-known alternative defines natural monopoly as an industry with a sub-additive cost function, which is based purely on the production side of an industry; see Baumol, Panzar, and Willig (1982). This suggests that the latter definition might be appropriate for regulated monopoly, while the present one would suit unregulated monopoly.

5 BERTRAND COMPETITION WITH DIFFERENTIATED PRODUCTS

This section considers a general model of Bertrand competition with differentiated products, and derives general conditions under which the game satisfies strategic complementarity (see Vives, 1990, and Milgrom and Roberts, 1990). This property is broadly thought as a natural one for price competition to satisfy. In addition, this section explores the general connection between the strategic complementarity (resp, substitutability) of the Bertrand game, and the nature of the inter-product relationship (gross substitutes and gross complements).[24]

An important assumption for the game to be well defined is that each firm is committed to satisfy whatever demand might be forthcoming at the prices quoted by the firms. Thus, even in cases where this results in a loss, firms simply cannot turn consumers away. In other words, some plays of the game might well lead to negative payoffs for some of the firms, but this is obviously an off-equilibrium possibility only.[25] One advantage of this formulation is that it dispenses with the need to specify a rationing scheme, thereby avoiding the possibility that key results might depend on this specification.

Consider an industry consisting of n single-product firms, with the demand for the good produced by firm i denoted $D_i(p_1, p_2, \ldots, p_n)$, $i = 1, 2, \ldots, n$. Firm i is assumed to have a cost function $C_i(\cdot)$.

Since firm i's sales correspond to its demand, the profit function of firm i is defined as usual by

$$\Pi_i(p_i, p_{-i}) = p_i D_i(p_i, p_{-i}) - C_i\left[D_i(p_i, p_{-i})\right]$$

and its reaction correspondence is

$$r_i(p_{-i}) = \arg\max_{p_i} \Pi_i(p_i, p_{-i}).$$

---

[24] Recall that when demands are linear, strategic complementarity (resp, substitutability) of the Bertrand game is equivalent to the products being gross substitutes (resp., gross complements): see e.g., Singh and Vives (1984). This is not true for general demand functions.

[25] This is one of the two commonly used models of price competition. The other model, wherein firms have bounded capacities and equilibrium is in mixed strategies, is the so-called Bertrand-Edgeworth model. This is beyond the scope of the present survey (for a fairly recent treatment, see e.g., Deneckere and Kovenock, 1996).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

We will say that the Bertrand oligopoly is symmetric if the demand functions are symmetric and $c_1 = c_2 = \ldots = c_n \triangleq c$.

We define the effective overall price space as $S_i \triangleq \{(p_1, p_2, \ldots, p_n) \in R^n_+ \mid D_i(p_1, p_2, \ldots, p_n) > 0\}$.

Standard assumptions The following are in effect throughout this section:

B1 D_i and C_i are twice continuously differentiable on S_i, $i = 1, 2, .., n$.

B2 $\dfrac{\partial D_i}{\partial p_i} < 0$, and $C_i' > 0$.

B3 $\sum_{k=1}^n \dfrac{\partial D_i(p_1, p_2, \ldots, p_n)}{\partial p_k} < 0$ over the set S_i.

B4 $\sum_{k=1}^n \dfrac{\partial^2 D_i(p_1, p_2, \ldots, p_n)}{\partial p_i \partial p_k} < 0$ over the set S_i.

These conditions are quite general, and are commonly invoked for differentiated-good demand systems. They have the following meanings and economic interpretations. For **B2**, part (i) is just the ordinary law of demand, and part (ii) says that goods *i* and *j* are substitutes.

As to **B3**, it is a strict dominant diagonal condition for the Jacobian matrix of the demand system (see e.g., Vives, 1999). It says that own price effect on demand exceeds the total cross-price effects. In a similar vein, **B4** says that each demand satisfies a strict dominant diagonal condition for the Hessian matrix of the demand system (see e.g., Vives, 1985, 1999). The usual interpretation is again that own effects of price changes dominate cross-effects, but for the slope of demand.

5.1 The Case of Substitute Goods

The main case considered here will be that of goods that are gross substitutes, defined by the condition $\frac{\partial D_i}{\partial p_j} > 0$ for all $i = 1, 2, .., n$. This says that an increase in the price of one good leads to an increase in the demand for the other good.

The following proposition provides sufficient conditions for the strategic complementarity of the Bertrand game (Vives, 1990):

Proposition 10 *Under assumptions **B1–B2**, assume that C_i is a convex function and that, over the set S_i*

$$\frac{\partial \log D_i}{\partial p_j} + p_i \frac{\partial^2 \log D_i}{\partial p_j \partial p_i} \geq 0, \ for\ j \neq i. \tag{3.12}$$

Then the Bertrand game is of strategic complements, and hence has a Bertrand equilibrium.

Proof The cross-partial of $\Pi_i(p_i, p_{-i})$ with respect to p_j and p_i is

$$\frac{\partial^2 \Pi_i(p_i, p_{-i})}{\partial p_j \partial p_i} = \frac{\partial D_i}{\partial p_j} + (p_i - C_i') \frac{\partial^2 D_i}{\partial p_j \partial p_i} - C_i'' \frac{\partial D_i}{\partial p_i} \frac{\partial D_i}{\partial p_j},$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

which is easily seen to be ≥ 0, provided one restricts attention to the region in price space where prices lie above marginal costs.

This then implies that the game is supermodular and that a Bertrand equilibrium exists (by Tarski's theorem). ■

Alternative sufficient conditions for the strategic complementarity of the Bertrand game are as follows (Milgrom and Roberts, 1990, Milgrom and Shannon, 1994).

Proposition 11 *Under Assumptions* ***B1–B2****, assume that C_i is a convex function and that, over the set S_i,*

$$D_i \frac{\partial^2 D_i}{\partial p_j \partial p_i} - \frac{\partial D_i}{\partial p_j}\frac{\partial D_i}{\partial p_i} \geq 0, \textit{ for } j \neq i. \tag{3.13}$$

Then the Bertrand game is of strategic complements, and hence has a Bertrand equilibrium.

Proof Milgrom and Shannon (1994) show that $\Pi_i(p_i, p_{-i})$ satisfies the single-crossing property in (p_i, p_j) for any $j \neq i$. This then implies that the game has strategic complements, and that a Bertrand equilibrium exists (by Tarski's theorem). ■

A natural question about these two sufficient conditions for the Bertrand game to satisfy strategic complementarity is whether they are nested or otherwise closely related. The two conditions can be seen by inspection to be quite similar, and quite general when the goods are substitutes (so $\frac{\partial D_i}{\partial p_j} > 0$). Both conditions rule out cases where the demand function is very submodular (i.e., has a very large and negative cross-partial derivative with respect to own price and any rival's price).[26]

Nevertheless, by way of examples, Amir and Grilo (2003) demonstrate that Conditions (3.12) and (3.13) are not nested. Indeed, the demand function $D_1(p_1, p_2) = e^{-p_1} + p_2/p_1$ satisfies Condition (3.12) but not Condition (3.13), while the demand function $D_1(p_1, p_2) = e^{(p_2 - p_1)}$ satisfies the latter but not the former (the verification details are left out). The reason that Conditions (3.12) and (3.13) are not nested is that both conditions are assumed to hold globally, including at prices that may not be reached along a best response.

Condition (3.13) has an exact economic interpretation: that the price elasticity of demand with respect to one good is increasing in the price of the other good. This is a very intuitive condition for price competition. In addition, it is seen by inspection that this condition treats the goods i and j symmetrically (in the sense that it holds for the two goods (i, j) if and only if it holds for the two goods (j, i)). In contrast, in order to guarantee that Condition (3.12) also treats the goods i and j symmetrically, one needs to assume that, as is commonly done, the demand functions are derived from maximizing the utility of a representative consumer (so that one necessarily has $\frac{\partial D_i}{\partial p_j} = \frac{\partial D_j}{\partial p_i}$).

[26] Some insight into the role of convexity of the cost function for the profit function to have increasing differences can be gleaned from Dierker and Dierker (1999).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

It is worth noting that Conditions (3.12) and (3.13) are not necessary for the conclusion that the reaction curve is increasing. To see this, consider the following example from Amir and Grilo (2003):

Example 6 Consider the following demand function: $D_1(p_1,p_2) = 10e^{p_2-p_1} + 2 - p_1$, and assume production costs are zero. Then

$$D^1_{p_1} = -10e^{p_2-p_1} - 1, D^1_{p_2} = 10e^{p_2-p_1} \text{ and } D^1_{p_1p_2} = -10e^{p_2-p_1}$$

It is easy to verify that neither Condition (3.12) nor (3.13) holds globally for this demand function. More specifically, we have

$$\frac{\partial D_1}{\partial p_2} + p_1 \frac{\partial^2 D_1}{\partial p_1 \partial p_2} = 10(p_1 - 1)e^{p_2-p_1} \geq 0 \text{ if and only if } p_1 \leq 1$$

and

$$D_1 \frac{\partial^2 D_1}{\partial p_1 \partial p_2} - \frac{\partial D_1}{\partial p_2}\frac{\partial D_1}{\partial p_1} = 10(1 - p_1)e^{p_2-p_1} \geq 0 \text{ if and only if } p_1 \geq 1.$$

Therefore, according to the two propositions above, we cannot conclude that the reaction curve here is upward-sloping.

Next consider the monotone transformation $h(x) = \log(ax + b)$, applied to $\Pi_1(p_1,p_2) = p_1D_1(p_1,p_2)$. By direct calculation, we have that

$$\frac{\partial^2 h \circ \Pi_1(p_1,p_2)}{\partial p_1 \partial p_2} \text{ has the same sign as } 10e^{p_2-p_1}\left[(p_1+1)(p_1-1)^2\right], \text{ which is } \geq 0.$$

Therefore, by Topkis's theorem applied to $h \circ \Pi_1(p_1,p_2)$, we conclude that $h \circ \Pi_1(p_1,p_2)$ has increasing differences, and therefore that $\Pi_1(p_1,p_2)$ has the single-crossing property.

From the supermodularity of the Bertrand game, one obtains at once that all the general properties of such games are satisfied by price competition, including in particular the partial coincidence of the usual solution concepts in terms of extremal bounds, the convergence of a wide class of adaptive learning algorithms to the same bounds, and the clear-cut comparative statics properties of Nash equilibrium strategies.

As to uniqueness, a standard dominant-diagonal argument on each firm's profit function follows directly from the approach pioneered by Rosen (1965). The result is as follows (see e.g., Vives, 1985, and Milgrom and Roberts, 1990, for more on this approach):

Proposition 12 *In addition to Assumptions **B1–B4**, assume that each cost function C_i is linear. Then there exists a unique Bertrand equilibrium.*

This proposition was used for instance in Amir, Encaoua, and Lefouili (2014), as a preparation for a comparative statics exercise on prices and profit as the common unit cost increases exogenously.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

5.2 The Case of Complementary Goods

This subsection provides a summary of the analogous results concerning the case of complementary products. The duality between the two cases, initially brought to light in Singh and Vives (1984) for the special case of linear demand, is discussed in some detail.

As usual, we say that two goods i and j are gross complements, or simply complements, when the condition $\frac{\partial D_i}{\partial p_j} < 0$ holds, i.e., when an increase in the price of one good leads to a decrease in the demand for the other good.

For complementary goods, the typical configuration for price competition is that the associated oligopoly game has strategic substitutes. The following result is the natural dual of the two propositions in the previous subsection:

Proposition 13 *When goods are complements, the Bertrand game has strategic substitutes if either of the following conditions holds:*

$$(i)\quad \frac{\partial D_i}{\partial p_j} + \left(p_i - C_i'\right)\frac{\partial^2 D_i}{\partial p_j \partial p_i} - C_i''\frac{\partial D_i}{\partial p_i}\frac{\partial D_i}{\partial p_j} \leq 0, \text{ over the set } S_i;$$

$$(ii)\quad D_i\frac{\partial^2 D_i}{\partial p_j \partial p_i} - \frac{\partial D_i}{\partial p_j}\frac{\partial D_i}{\partial p_i} \leq 0, \text{ over the set } S_i.$$

Combining with the previous subsection, this suggests that, in the typical situation, a Bertrand oligopoly game will have strategic complements (resp., substitutes) when the goods are substitutes (resp., complements).[27] For the special case of linear demand, this association and its dual version hold perfectly well, as highlighted in Singh and Vives (1984) and others.

However, for general non-linear demands, this connection need not hold. In other words, it is possible for the game to have strategic complements (i.e., Conditions (3.12) or (3.13) to hold) when goods are complements, and for the previous result to hold for goods that are substitutes. Clearly, the conditions needed then become quite restrictive, but robust classes of such can easily be found. Here is an illustrative example of such, somewhat restrictive but nevertheless robust, cases:

Example 7 Consider the following two-good demand system:

$$D_1(p_1,p_2) = \frac{1}{(p_1+1)^2} + (p_2+1)e^{-p_1} \text{ and } D_2(p_1,p_2) = a - e^{-p_1} - p_2, a > 0.$$

It is easily verified that $\partial D_1(p_1,p_2)/\partial p_2 = \partial D_2(p_1,p_2)/\partial p_1$, so that this demand can be derived via utility maximization (of a representative consumer). Partial differentiation yields that

$$D_1\frac{\partial^2 D_1}{\partial p_1 \partial p_2} - \frac{\partial D_1}{\partial p_2}\frac{\partial D_1}{\partial p_1} = \frac{-e^{-2p_1}}{(p_1+1)^3} < 0$$

[27] The opposite holds for Cournot oligopoly with differentiated products, namely strategic substitutes goes with substitutes and strategic complements goes with complements.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

and

$$D_2 \frac{\partial^2 D_2}{\partial p_1 \partial p_2} - \frac{\partial D_2}{\partial p_2} \frac{\partial D_2}{\partial p_1} = e^{-p_1} > 0.$$

From the signs of the two expressions and the above results, we conclude that the reaction curve is downward-sloping although, for demand D_1, the two goods are substitutes. On the other hand, for demand D_2, the two goods are complements but the reaction curve is upward-sloping.

6 CONCLUSION

This survey provides a thorough account of the literature on the existence and comparative statics of equilibria in Cournot and Bertrand oligopoly. The main focus is on the more recent literature on these topics, which is largely based on the methodology of supermodular optimization and games, also often called the lattice-theoretic approach.

The first strand of literature covered in some detail is that dealing with the existence of Cournot equilibrium in general asymmetric settings for the homogeneous good case. Due to the fact that Cournot oligopoly is an aggregative game, as understood by the literature in the 1970s already, existence of Cournot equilibrium hinges completely on the central property of strategic substitutes.

The second strand of literature deals with the comparative statics of symmetric Cournot equilibria with respect to the entry of new firms, and includes its own results regarding the existence of symmetric equilibrium. The latter are distinct from the general existence results, and are governed by the same key assumption, of limited scale economies, as the comparative statics of entry. For this part, the key property that delineates the two relevant cases of analysis is neither strategic substitutes nor strategic complements.

As to Bertrand competition with differentiated products, the literature reviewed covers much fewer studies and is far more recent. The main question deals with natural conditions on the demand and cost functions that lead such Bertrand games to be games of strategic complements. Existence of Bertrand equilibrium then follows directly, as do all the powerful properties of supermodular games. The relationship between the strategic complementarity of the game and the gross substitute/complements properties of the products for non-linear demand functions is also clarified.

REFERENCES

Acemoglu, D. and M.K. Jensen (2009), Aggregate comparative statics, *Games and Economic Behavior*, 81, 27–49.
Amir, R. (1996a), Cournot oligopoly and the theory of supermodular games, *Games and Economic Behavior*, 15, 132–148.
Amir, R. (1996b), Sensitivity analysis in multisector optimal economic dynamics, *Journal of Mathematical Economics*, 25, 123–141.
Amir, R. (1996c), Continuous stochastic games of capital accumulation with convex transitions, *Games and Economic Behavior*, 15, 111–131.
Amir, R. (2005), Supermodularity and complementarity in economics: An elementary survey, *Southern Economic Journal*, 71, 636–660.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Amir, R. and I. Grilo (2003), On complementarity conditions in Bertrand oligopoly, *Economic Theory*, 22, 227–232.
Amir, R. and V. Lambson (2000), On the effects of entry in Cournot markets, *Review of Economic Studies*, 67, 235–254.
Amir, R. and N. Lazzati (2011), Network effects, market structure and industry performance, *Journal of Economic Theory*, 146: 2389–2419.
Amir, R., D. Encaoua, and Y. Lefouili (2014), Optimal licensing of uncertain patents in the shadow of litigation, *Games and Economic Behavior* 88, 320–338.
Amir, R., F. Garcia, and M. Knauff (2010), Symmetry-breaking in two-player games via strategic substitutes and diagonal nonconcavity: A synthesis, *Journal of Economic Theory*, 145, 1968–1986.
Amir, R., L. Koutsougeras, and L. de Castro (2014), Free entry versus socially optimal entry, *Journal of Economic Theory*, 154, 112–125.
Anderson, S. and R. Renault (2003), Efficiency and surplus bounds in Cournot competition, *Journal of Economic Theory*, 113, 253–264.
Anderson, S., N. Erkal, and D. Piccinin (2014), Aggregative oligopoly games with entry, mimeo.
Bamon, R. and J. Fraysse (1985), Existence of Cournot equilibrium in large markets, *Econometrica*, 53, 587–598.
Baumol, W., J. Panzar, and R. Willig (1982), *Contestable Markets and the Theory of Industry Structure*, New York: Harcourt Brace Jovanovich, Inc.
Bertrand, J. (1883), Book review of "Theorie mathematique de la richesse sociale" and of "Recherches sur les principles mathematiques de la theorie des richesses", *Journal des Savants* 67: 499–508.
Bulow, J., J. Geanakoplos, and P. Klemperer (1985), Multimarket oligopoly: Strategic substitutes and complements, *Journal of Political Economy*, 93, 488–511.
Caplin, A. and B. Nalebuff (1991), Aggregation and imperfect competition: On the existence of equilibrium, *Econometrica*, 59, 25–59.
Corchón, L. (1994), Comparative statics for aggregative games: The strong concavity case, *Mathematical Social Sciences*, 28, 151–165.
Cornes, R. and R. Hartley (2012), Fully aggregative games, *Economics Letters*, 116, 631–633.
Cournot, A. (1838), *Recherches sur les Principes Mathematiques de la Theorie des Richesses*, Paris: Hachette.
De Meza, D. (1985), A stable Cournot-Nash industry need not be quasi-competitive, *Bulletin of Economic Research*, 37, 153–156.
Deneckere, R. and D. Kovenock (1996), Bertrand-Edgeworth duopoly with unit cost asymmetry, *Economic Theory*, 8, 1–25.
Dierker, E. and H. Dierker (1999), Product differentiation and market power, *Advances in Mathematical Economics*, 1, 39–67.
Dierker, E. and K. Podczek (1992), The distribution of consumers' tastes and the quasiconcavity of the profit function, *Vienna Economics Papers* 9208, University of Vienna.
Dixit, A. (1986), Comparative statics for oligopoly, *International Economic Review*, 107–122.
Dubey, P., O. Haimanko, and A. Zapechelnyuk (2006), Strategic complements and substitutes, and potential games, *Games and Economic Behavior*, 54, 77–94.
Echenique, F. (2002), Comparative statics by adaptive dynamics and the correspondence principle, *Econometrica*, 70, 833–844.
Edlin, A.S. and F. Echenique (2004), Mixed equilibria are unstable in games of strategic complements, *Journal of Economic Theory*, 118, 61–79.
Edlin, A. and C. Shannon (1998), Strict monotonicity in comparative statics, *Journal of Economic Theory*, 89, 201–219.
Ewerhart, C. (2014), Cournot games with bi-concave demand, *Games and Economic Behavior*, 85, 37–47.
Frank, C.R. Jr. (1965), Entry in a Cournot market, *Review of Economic Studies*, 32, 245–250.
Frank, C.R. Jr. and R.E. Quandt (1963), On the existence of Cournot equilibrium, *International Economic Review*, IV, 92–96.
Friedman, J. (1977), *Oligopoly and the Theory of Games*, Amsterdam, North-Holland.
Friedman, J. (1982), Oligopoly theory, in: K. Arrow and M. Intrilligator, eds, *Handbook of Mathematical Economics*, Amsterdam: North Holland, Chapter 11.
Fudenberg, D. and J. Tirole (1984), The fat cat effect, the puppy dog ploy and the lean and hungry look. *American Economic Review Papers and Proceedings*, 74, 361–368.
Gaudet, G. and S. Salant (1991), Uniqueness of Cournot equilibrium: New results from old methods, *Review of Economic Studies*, 58, 399–404.
Hahn, F. (1962), The stability of the Cournot oligopoly solution, *Review of Economic Studies,* 29, 329–331.
Hoernig, S.H. (2003), Existence of equilibrium and comparative statics in differentiated goods Cournot oligopolies, *International Journal of Industrial Organization*, 21, 989–1019.
Jensen, M.K. (2010), Aggregative games and best-reply potentials, *Economic Theory*, 43, 45–66.
Kolstad, C. and L. Mathiesen (1987), Necessary and sufficient conditions for uniqueness of a Cournot equilibrium, *Review of Economic Studies*, 54, 681–690.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Kukushkin, N.S. (1994), A fixed-point theorem for decreasing mappings, *Economics Letters*, 46, 23–26.
Mankiw, N.G. and M.D. Whinston (1986), Free entry and social inefficiency, *Rand Journal of Economics*, 17, 48–58.
McManus, M. (1962), Number and size in Cournot equilibrium, *Yorkshire Bulletin of Economic and Social Research*, 14, 14–22.
McManus, M. (1964), Equilibrium, number and size in Cournot equilibrium, *Yorkshire Bulletin of Economic and Social Research,* 16, 68–75.
Milgrom, P. and J. Roberts (1990), Rationalizability, learning, and equilibrium in games with strategic complementarities, *Econometrica*, 58, 1255–1278.
Milgrom, P. and J. Roberts (1994), Comparing equilibria, *American Economic Review*, 84, 441–459.
Milgrom, P. and C. Shannon (1994), Monotone comparative statics, *Econometrica*, 62, 157–180.
Novshek, W. (1980), Cournot equilibrium with free entry, *Review of Economic Studies*, 47, 473–486.
Novshek, W. (1984), Finding all n-firm Cournot equilibria, *International Economic Review*, 25, 61–70.
Novshek, W. (1985), On the existence of Cournot equilibrium, *Review of Economic Studies*, 52, 85–98.
Okuguchi, K. (1973), Quasi-competitiveness and Cournot oligopoly, *Review of Economic Studies*, 40, 145–148.
Quah, J.K.-H. and B. Strulovici (2012), Aggregating the single crossing property, *Econometrica*, 80, 2333–2348.
Quirmbach, H.C. (1988), Comparative statics for oligopoly: Demand shift effects, *International Economic Review*, 29, 451–459.
Roberts, J. and H. Sonnenschein (1976), On the existence of Cournot equilibrium without concave profit functions, *Journal of Economic Theory*, 13, 112–117.
Roberts, J. and H. Sonnenschein (1977), On the foundations of the theory of monopolistic competition, *Econometrica*, 45, 101–113.
Rosen, J.B. (1965), Existence and uniqueness of equilibrium points for concave n-person games, *Econometrica*, 3, 520–534.
Roy, S. and T. Sabarwal (2012), Characterizing stability properties in games with strategic substitutes, *Games and Economic Behavior*, 75, 337–353.
Ruffin, R.J. (1971), Cournot oligopoly and competitive behavior, *Review of Economic Studies*, 38, 493–502.
Samuelson, P. (1947), *Foundations of Economic Analysis, Harvard Economic Studies, Vol 80*, Cambridge, MA: Harvard University Press.
Seade, J. (1980), On the effects of entry, *Econometrica*, 48, 479–489.
Selten, R. (1970), *Preispolitik der Mehrproduktenunternehmung in der Statischen Theorie*, Berlin: Springer Verlag.
Shapiro, C. (1989), Theories of oligopoly behavior, in R. Schmalensee and R. Willig, eds, *Handbook of Industrial Organization*, Amsterdam: North-Holland.
Shapley, L.S. (1971), Cores of convex games, *International Journal of Game Theory*, 1, 11–26.
Singh, N. and X. Vives (1984), Price and quantity competition in a differentiated duopoly, *The RAND Journal of Economics*, 15, 546–554.
Suzumura, K. and K. Kiyono (1987), Entry barriers and economic welfare, *Review of Economic Studies*, 54, 157–167.
Szidarovszky, F. (1970), On the oligopoly game, *Working Paper* DM70-1, Karl Marx University of Economics, Budapest.
Szidarovszky, F. and K. Okuguchi (2008), On the existence and uniqueness of pure Nash equilibrium in rent-seeking games, *Games and Economic Behavior*, 18, 135–140.
Szidarovszky, F. and S. Yakowitz (1977), A new proof of the existence and uniqueness of the Cournot equilibrium, *International Economic Review*, 18, 787–789.
Szidarovszky, F. and S. Yakowitz (1982), Contributions to Cournot oligopoly theory, *Journal of Economic Theory*, 28, 51–70.
Tarski, A. (1955), A lattice-theoretic fixed point theorem and its applications, *Pacific Journal of Mathematics*, 5, 285–309.
Topkis, D. (1978), Minimizing a submodular function on a lattice, *Operations Research*, 26, 305–321.
Topkis, D. (1979), Equilibrium points in nonzero-sum *n*-person submodular games, *SIAM Journal of Control and Optimization*, 17, 773–787.
Topkis, D. (1998), *Submodularity and Complementarity*, Princeton NJ: Princeton University Press.
Vives, X. (1985), On the efficiency of Bertrand and Cournot equilibria with product differentiation, *Journal of Economic Theory*, 36, 166–175.
Vives, X. (1990), Nash equilibrium with strategic complementarities, *Journal of Mathematical Economics*, 19, 305–321.
Vives, X. (1999), *Oligopoly Pricing: Old Ideas and New Tools*, Cambridge, MA: MIT Press.
Vives, X. (2005a), Complementarities and games: New developments, *Journal of Economic Literature*, 43, 437–479.
Vives, X. (2005b), Games with strategic complementarities: New applications to industrial organization, *International Journal of Industrial Organization* 23, 625–637.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

APPENDIX

Proofs

We provide the proof of Proposition 2, omitted in the text.

Proof of Proposition 2 We need to show that for any $q_i' > q_i$ and $Q'_{-i} > Q_{-i}$,

$$q_i'P(q_i' + Q_{-i}) - C_i(q_i') \leq q_iP(q_i + Q_{-i}) - C_i(q_i) \Rightarrow$$
$$q_i'P(q_i' + Q'_{-i}) - C_i(q_i') < q_iP(q_i + Q'_{-i}) - C_i(q_i). \tag{3.14}$$

Log-concavity of $P(\cdot)$ is equivalent to log-submodularity of $P(q_i + Q_{-i})$ in (q_i, Q_{-i}), i.e.,

$$\log P(q_i' + Q'_{-i}) + \log P(q_i + Q_{-i}) \leq \log P(q_i' + Q_{-i}) + \log P(q_i + Q'_{-i}),$$

or

$$P(q_i' + Q'_{-i})P(q_i + Q_{-i}) \leq P(q_i' + Q_{-i})P(q_i + Q'_{-i}). \tag{3.15}$$

With the LHS of (3.14) as starting point,

$$q_i'P(q_i' + Q_{-i}) - C_i(q_i') \leq q_iP(q_i + Q_{-i}) - C_i(q_i) \leq q_i\frac{P(q_i' + Q_{-i})P(q_i + Q'_{-i})}{P(q_i' + Q'_{-i})} - C_i(q_i) \tag{3.16}$$

multiply across by $P(q_i' + Q'_{-i})/P(q_i' + Q_{-i})$,

$$q_i'P(q_i' + Q'_{-i}) - \frac{P(q_i' + Q'_{-i})}{P(q_i' + Q_{-i})}C_i(q_i') \leq q_iP(q_i + q'_{-i}) - \frac{P(q_i' + Q'_{-i})}{P(q_i' + Q_{-i})}C_i(q_i).$$

Since $C_i(q_i') > C_i(q_i)$ and $P(q_i' + Q'_{-i}) < P(q_i' + Q_{-i})$, it follows that

$$q_i'P(q_i' + Q'_{-i}) - C_i(q_i') < q_iP(q_i + Q'_{-i}) - C_i(q_i). \tag{3.17}$$

Since (3.16) implies (3.17), (3.14) holds. ■

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

4. Aggregative games

*Martin Kaae Jensen**

1 INTRODUCTION AND OVERVIEW

The machinery of aggregative games has a proven track-record of applications in game theory and industrial organization (IO). This chapter's aim is to survey the literature with a firm focus on usefulness: to show *when* the methods apply, *what* they can be used to establish, and *why* they are usually simpler and more powerful than any alternative (when an alternative exists, which is not always the case).

To begin with the *when*, consider what is perhaps the oldest and most familiar model from industrial organization, namely the Cournot model. An oligopolist i has profit function $\pi_i(s_1, \ldots, s_I) = s_i P(s_i + \sum_{j \neq i} s_j) - C_i(s_i)$ where P is the inverse demand function and C_i the cost function. The objective is to maximize profits with respect to the firm's output s_i, taking the other firms' outputs $(s_j)_{j \neq i}$ as given. When all firms $i \in \{1, \ldots, I\}$ maximize profits given other firms' outputs, we have a Cournot equilibrium. What makes this game aggregative is the fact that profit functions can be expressed as a function of firms' own outputs and an *aggregate*, here aggregate output $g(s) = \sum_j s_j$:

$$\pi_i(s_1, \ldots, s_i) = s_i P(g(s)) - c_i(s_i).$$

As we shall see when we define generalized and quasi-aggregative games in Section 2, this definition generalizes substantially so that, for example, g may take other functional forms than the linear sum. Examples of aggregative games abound in the literature: public good provision games, tournaments, teamwork games, contests, patent races, Bertrand oligopoly, many network games, etc., etc. are all aggregative games.

As for *what* aggregative games are useful for, the answer is that aggregative games methodology is useful for addressing nearly all of the questions one typically seeks to answer in applied work. Here most of the focus will be on pure-strategy Nash equilibria and their comparative statics (Section 4), existence (Section 3), and uniqueness (Section 6). But issues such as stability (Section 7.5), computation and algorithms (Sections 3.2, 7.2), entry (Section 7.3) will be briefly surveyed as well. Extensions, including games with a continuum of agents (large games/mean-field games, Section 7.1) and alternative equilibrium concepts (*e.g.*, evolutionary stable states; Section 7.4) will also be introduced.

Since the main focus is on existence, comparative statics, and uniqueness, a brief discussion is called for. Beginning with existence, one issue that "plagued" industrial economics until well into the 1980s (Novshek 1985) is that to prove existence of a Cournot equilibrium by the usual route via Kakutani's fixed-point theorem, one will have to assume that profit functions are concave or at least quasi-concave. This may sit poorly with the spirit of oligopolistic

* I would like to thank Dario Bauso and Parise Francesca for helpful comments and suggestions. All remaining errors are my responsibility.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

competition where, arguably, the fact that firms are large enough relative to the market to constitute an oligopoly in the first place could be due to non-convexities such as entry costs, or increasing returns. Luckily, quasi-concavity is in fact *not* necessary for existence when the game exhibits strategic substitutes, *i.e.*, when increased output of other firms leads to decreased production for the firm at hand. And this is hardly an assumption that will raise any eyebrows in the Cournot model (as well as in many other models of imperfect competition). Formulating and generalizing this result will occupy Section 3. Note that to this day, the aggregative games methodology is the only known way to establish such existence without quasi-concavity in games with strategic substitutes.

While existence is of course critical, *comparative statics* is the backbone of predictions. In aggregative games, the study of comparative statics was initiated by Corchón (1994), and the topic has recently received a lot of attention, leading to a kind of "merger" with the monotone comparative statics literature of Topkis (1978), Milgrom and Roberts (1994), Milgrom and Shannon (1994), and Quah (2007), among others. Aggregative games are able to offer general and robust comparative statics results when all other methods fail – including the monotone methods just mentioned – an example being games of strategic substitutes with payoff functions that are not quasi-concave (*cf.*, the Cournot model discussed in the previous paragraph). There are several approaches: one can work directly with so-called backward-response correspondences, or one can use general comparative statics theorems. Either way, the simplicity far outpaces any direct application on the implicit function theorem even when the implicit function theorem applies. This is the topic of Section 4.

Concerning uniqueness of equilibrium, the situation mirrors that of comparative statics. One is able to derive very powerful uniqueness results by use of the aggregative games methodology. And importantly, the results are straight forward to apply. This should be contrasted with a direct approach to uniqueness, which is, in general, extremely challenging (see *e.g.*, Vives, 2000 for an overview).

Finally, a caveat. The reader should be aware that this survey has somewhat of a bias towards my own work and interests. In an attempt to counter this, quite a long section (Section 7) surveys applications and extensions in a number of different directions, which will, at the very least, allow interested readers to read in other directions.

2 BASIC DEFINITIONS

Aggregative games form a subset of pure-strategy non-cooperative games. This section defines three classes of aggregative games, each of which increases generality a notch: linearly aggregative, generalized aggregative, and quasi-aggregative games respectively. We consider games with scalar aggregates and a finite set of players $I \in \mathbb{N}$ players and postpone discussion of games with a continuum of agents and multi-dimensional aggregates for later (Section 7.1). Throughout, the notation is fixed as follows: Agent $i \in \{1, \dots, I\}$ has strategy set S_i with typical element s_i. A (joint) strategy is denoted $s = (s_1, \dots, s_I) \in S \equiv \prod_{i=1}^{I} S_i$, and for a fixed player i, a vector of opponents' strategies is denoted $s_{-i} = (s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_I) \in S_{-i} \equiv \prod_{j \neq i} S_j$. Agent i's payoff function is $\Pi_i : S \to \mathbb{R}$. Finally, s^* is a (pure-strategy Nash) equilibrium if:

$$\Pi_i(s^*) \geq \Pi_i(s_i, s^*_{-i}) \text{ for all } s_i \in S_i,\ i = 1, \dots, I. \tag{4.1}$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Before we begin, a "warning": this section is entirely "assumption free". This simplifies the exposition tremendously but, of course, it means that in all cases, the objects defined may be empty sets and/or encompass all games (depending on interpretation and the specific situation). In later sections when results are presented, assumptions are introduced as they are needed.

2.1 Linearly Aggregative Games

The original and simplest definition of an aggregative game is due to Selten (1970). It will here be called a *linearly* aggregative game to avoid confusion. We have already seen an example in the Introduction, namely the Cournot oligopoly.

Definition 1 (Linearly Aggregative Games) *A non-cooperative game* $(S_i, \Pi_i)_{i=1,\ldots,I}$ *where* $S_i \subseteq \mathbb{R}$ *for all* i *is linearly aggregative if for every agent* $i \in \{1,\ldots,I\}$ *there exists a function* $\pi_i : S_i \times \mathbb{R} \to \mathbb{R}$ *(the reduced payoff function) such that:*

$$\Pi_i(s) = \pi_i\left(s_i, \sum_j s_j\right) \quad \textit{for all } s \in S. \tag{4.2}$$

The function $g(s) = \sum_j s_j$ *is called the aggregator and a value in the range of* g, $\sum_j s_j \in X \equiv \left\{\sum_j s_j : s_j \in S_j \textit{ for } j = 1,\ldots,I\right\}$ *is called an aggregate. If* $(s_1^*,\ldots,s_i^*)$ *is a (pure-strategy Nash) equilibrium for the game,* $\sum_j s_j^*$ *is called an equilibrium aggregate.*

For a long list of specific examples, see *e.g.*, Alos-Ferrer and Ania (2005). In the Cournot model where firm i's profit function is $\Pi_i(s) = s_i P\left(\sum_j s_j\right) - C_i(s_i)$, (4.2) will be satisfied since $\pi_i(s_i, Q) = s_i P(Q) - c_i(s_i)$. Note the "$Q$-notation" used here (and repeatedly in what follows): whenever the reader sees a Q, this denotes an aggregate. Replacing Q with the *aggregator*, here $g(s) = \sum_j s_j$, then leads to the satisfaction of the relevant conditions (here (4.2)). Getting used to this notation will serve us well later on.

2.2 Generalized Aggregative Games

The extension from linearly to generalized aggregative games proceeds by relaxing the requirement that the aggregator $g(s) = \sum_j s_j$ from Definition 1 is linear. Instead, g is allowed to be any additively separable function (Gorman, 1968) which means that $g(s) = H(\sum_{i=1}^I h_i(s_i))$, $s \in S$ where $H : \mathbb{R} \to \mathbb{R}$ and $h_i : S_i \to \mathbb{R}$, $i = 1,\ldots,I$ are strictly increasing functions. Note that while a linearly aggregative game must have real-valued strategy sets ($S_i \subseteq \mathbb{R}$ all i), a generalized aggregative game allows for multi-dimensional strategy sets although the aggregator must still be scalar valued. Thus the generalization from a linear aggregator to an additively separable aggregator leads to a substantial generalization in the dimensionality of the game.[1]

[1] As discussed in Section 7.1 there is nothing in the way of allowing strategy sets to be subsets of more general spaces than $\mathbb{R}^n$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Definition 2 (Generalized Aggregative Games) *A non-cooperative game* $(S_i, \Pi_i)_{i=1,\ldots,I}$ *where* $S_i \subseteq \mathbb{R}^n$ *for all* i *is generalized aggregative if there exists an additively separable function* $g : S \to \mathbb{R}$ *(the aggregator) and functions* $\pi_i : S_i \times \mathbb{R} \to \mathbb{R}$ *(the reduced payoff functions) such that:*

$$\Pi_i(s) = \pi_i(s_i, g(s)) \text{ for all } s \in S,\ i = 1, \ldots, I. \tag{4.3}$$

A value in the range of g, $g(s) \in X \equiv \{g(s) : s_j \in S_j \text{ for } j = 1, \ldots, I\}$ *is called an aggregate, and an equilibrium aggregate if* s^* *is a (pure-strategy) Nash equilibrium for the game.*

Of course, $g(s) = \sum_i s_i$ is additively separable. Hence any linearly aggregative game is generalized aggregative. The mean $g(s) = I^{-1} \sum_{i=1}^{I} s_i$ is an equally obvious example, and in fact all of the standard means, including the harmonic mean, the geometric mean, and the power means are additive separable functions (Jensen, 2010, Section 2.3.2). Two other useful examples are $g(s) = (\alpha_1 s_1^{\beta} + \ldots + \alpha_I s_I^{\beta})^{1/\beta}$, $S \subseteq \mathbb{R}_+^N$, and $g(s) = \prod_{i \in \mathscr{I}} s_i^{\alpha_i}$, $S \subseteq \mathbb{R}_{++}^N$, where $\beta, \alpha_1, \ldots, \alpha_I > 0$, which are, respectively, a constant elasticity of substitution (CES) function and a Cobb-Douglas function.[2]

Examples of generalized aggregative games abound in the literature (see, for example, Jensen, 2010 for an extensive list). One such game that has received much attention is the *contests* (see the Chapters 6 and 7 in Volume II of this *Handbook* as well as Section 7.6 below). Let agent (contender) i have payoff function,

$$\Pi_i\left(s_i, s_{-i}\right) = V_i \cdot \frac{h_i\left(s_i\right)}{R + \sum_{j=1}^{I} h_j\left(s_j\right)} - c_i\left(s_i\right), \tag{4.4}$$

where $s_i \in S_i \subseteq \mathbb{R}_+$ denotes *effort*, $c_i : \mathbb{R}_+ \to \mathbb{R}_+$ is the agent's *cost function*, and $V_i > 0$ the *valuation* of the prize. The strictly increasing functions $h_1, \ldots, h_I : \mathbb{R}_+ \to \mathbb{R}_+$ and the constant $R \geq 0$ determine the *contest success functions* $\frac{h_i(s_i)}{R + \sum_{j=1}^{I} h_j(s_j)}$, $i \in \mathscr{I}$, which map a given profile of efforts $s = (s_1, \ldots, s_I)$ into the agents' winning probabilities. It is clear that if we define,

$$g(s) = \sum_{j=1}^{I} h_j\left(s_j\right),$$

then this is a generalized aggregative game with aggregator g and reduced payoff functions:

$$\pi_i\left(s_i, Q\right) = V_i \cdot \frac{h_i\left(s_i\right)}{R + Q} - c_i\left(s_i\right), i = 1, \ldots, I. \tag{4.5}$$

Note that here we have again used the "Q-notation" mentioned at the end of the previous subsection.

[2] In the first case $h_i(s_i) = \alpha_i s_i^{\beta}$ (with $s_i \geq 0$) and $H(z) = z^{1/\beta}$. In the second $h_i(s_i) = \alpha_i \log(s_i)$ and $H(z) = \exp(z)$ (with $s_i > 0$).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

2.3 Quasi-aggregative Games

To ease our way into quasi-aggregative games, consider the following special case that directly relates to linearly and generalized aggregative games:

Definition 3 (Quasi-aggregative Games I) *A non-cooperative game* $(S_i, \Pi_i)_{i=1,\ldots,I}$ *where* $S_i \subseteq \mathbb{R}^n$ *for all* i *is quasi-aggregative with aggregator* $g : S \to \mathbb{R}$ *if there exist functions* $\pi_i : S_i \times \mathbb{R} \to \mathbb{R}$ *(the reduced payoff functions) such that*

$$\Pi_i(s) = \pi_i(s_i, g(s)) \text{ for all } s \in S, \ i = 1, \ldots, I, \tag{4.6}$$

as well as continuous functions $F_i : \mathbb{R} \times S_i \to \mathbb{R}$ *(the shift functions), and* $\sigma_i : S_{-i} \to X_{-i} \subseteq \mathbb{R}$, $i \in \mathscr{I}$ *(the interaction functions), such that:*

$$g(s) = F_i(s_i, \sigma_i(s_{-i})), \text{ for all } s \in S \text{ and all } i \in \mathscr{I}. \tag{4.7}$$

Any generalized aggregative game is quasi-aggregative since a function g is additively separable if and only if the functional equations (4.7) hold for strictly increasing functions F_i and σ_i, $i = 1, \ldots, I$ (Gorman, 1968). Of course, the converse is false: if F_i and/or σ_i is not strictly increasing for some i, then g will not be additively separable and the quasi-aggregative game will therefore not be generalized aggregative.[3] The simplest example of a game that satisfies Definition 3 but is not generalized aggregative is that of a game with aggregator $g(s) = \prod_j s_j$ (here $F_i(s_i, x_{-i}) = s_i \cdot x_{-i}$ and $\sigma_i(s_{-i}) = \prod_{j \neq i} s_j$) when $S_i \subseteq \mathbb{R}_+$. Indeed, g is *not* strictly increasing (take $s_i = 0$ for some i), hence g is not additively separable.

Now, if one "plugs" (4.7) into (4.6),

$$\Pi_i(s) = \pi_i(s_i, g(s)) = \pi_i(s_i, F_i(s_i, \sigma_i(s_{-i}))),$$

it is seen that the payoff of player i always depends on the player's own strategy and the interaction term $\sigma_i(s_{-i})$ determined by the interaction function $\sigma_i : S_{-i} \to \mathbb{R}$. The standard definition of a quasi-aggregative game allows for this increased level of generality:

Definition 4 (Quasi-aggregative Games) *The game* $(S_i, \Pi_i)_{i \in \mathscr{I}}$ *is said to be a quasi-aggregative game with aggregator* $g : S \to \mathbb{R}$, *if there exist continuous functions* $F_i : \mathbb{R} \times S_i \to \mathbb{R}$ *(the shift functions), and* $\sigma_i : S_{-i} \to X_{-i} \subseteq \mathbb{R}$, $i \in \mathscr{I}$ *(the interaction functions) such that each of the payoff functions* $i \in \mathscr{I}$ *can be written:*

$$\Pi_i(s) = \pi_i(s_i, \sigma_i(s_{-i})), \tag{4.8}$$

where $\pi_i : X_{-i} \times S_i \to \mathbb{R}$, *and:*

$$g(s) = F_i(s_i, \sigma_i(s_{-i})), \text{ for all } s \in S \text{ and all } i \in \mathscr{I}. \tag{4.9}$$

[3] Needless to say, one must assume *something* about the functional forms involved so as to not "drown" in the generality. See Section 3. The example that follows satisfies all of the assumptions of that section.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

One further generalization turns out to be useful:

Definition 5 (Generalized Quasi-aggregative Games) *The game* $(S_i, \Pi_i)_{i \in \mathscr{I}}$ *is said to be a generalized quasi-aggregative game with aggregator* $g : S \to \mathbb{R}$ *if it satisfies Definition* 4 *with* (4.9) *replaced by the weaker set of conditions,*

$$g(s) = F_i(s_i, \sigma_i(s_{-i})) + v_i(s_{-i}), \text{ for all } s \in S \text{ and all } i \in \mathscr{I}, \tag{4.10}$$

where $v_1, \dots, v_I$ *may be arbitrary real-valued functions.*

As is clear from the definition, a quasi-aggregative game is an *ordinal* concept, *i.e.*, all functions involved are only determined up to a strictly monotonic transformation (which may be either strictly increasing or strictly decreasing).[4] We shall return to some non-trivial examples of (generalized) quasi-aggregative games in section 3 when we discuss network games and potentials.

It was already mentioned above that (4.9) will be satisfied for strictly monotonic functions F_i and σ_i provided that g is *additively separable* (Gorman, 1968). If one's starting point is payoff functions in the form (4.2) where g is additively separable, the game is consequently quasi-aggregative. The converse is true as well: if (4.9) holds for strictly monotonic functions F_i and σ_i, then g will be additively separable. It is important to stress, then, that standard results on (generalized) quasi-aggregative games do not require that F_i and σ_i are monotonic (see Section 3).

3 EQUILIBRIUM AGGREGATES AND THE BACKWARD-RESPONSE CORRESPONDENCE

In this section, attention is restricted to generalized aggregative games. The analysis of such games can be reduced to studying fixed points of a correspondence from $\mathbb{R}$ into $\mathbb{R}$ called the aggregate backward-response correspondence. This reduced the dimensionality from a fixed-point problem of dimension $I \times n$ (where I is the number of players and n the dimension of the strategy sets) to dimension 1 – which very substantially simplifies the analysis and allows for a wealth of strong results on existence, uniqueness, comparative statics, etc.

3.1 The Backward-response Correspondence

To begin we must define the so-called *backward-response function*, or more generally, *the backward-response correspondence*. This function/correspondence is the most basic tool in the analysis of aggregative games up until the generality afforded by generalized aggregative

[4] Hence the term *quasi*-aggregative. The observation is trivial for the interaction functions. It is also clear that (4.9) will hold for functions g and F_i, $i \in \mathscr{I}$ if and only if it holds for $\tilde{g} = h \circ g$ and $\tilde{F}_i = h \circ F_i$, $i \in \mathscr{I}$ where $h : \mathbb{R} \to \mathbb{R}$ is either strictly increasing or strictly decreasing.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

games, after which potentials take over as the main tool (Section 3). The construction dates back, once again, to Selten (1970).[5]

The definition of the backward-response correspondence is particularly simple in the case of linearly aggregative games, so we shall begin with this case. When a game is linearly aggregative, player i's *best-response correspondence* $r_i : S \to 2^{S_i} \cup \emptyset$ is given by,

$$r_i(s_{-i}) = \arg\max_{s_i \in S_i} \Pi_i(s_i, \sum_j s_j) = \arg\max_{s_i \in S_i} \Pi_i(s_i, s_i + \sum_{j \neq i} s_j). \tag{4.11}$$

Regardless of the exact structure of the (reduced) payoff function, agent i's best responses can, as seen, only ever depend on the sum of the opponents' strategies, $\sum_{j \neq i} s_j$. For this reason, we may define the *reduced best-response correspondence*, $R_i : \mathbb{R} \to 2^{S_i} \cup \emptyset$, which gives the best responses for any value of the sum of the opponents' strategies $x_{-i} \in X_{-i} = \left\{\sum_{j \neq i} s_j : s_j \in S_j \text{ for all } j \neq i\right\} \subseteq \mathbb{R}$,

$$R_i(x_{-i}) = \arg\max_{s_i \in S_i} \Pi_i(s_i, s_i + x_{-i}). \tag{4.12}$$

Now fix a real number $Q \in X \equiv \left\{\sum_j s_j : s_j \in S_j \text{ all } j\right\} \subseteq \mathbb{R}$, which should be interpreted as a value that the sum of all agents' strategies (*i.e.*, the aggregate) could take. Given Q as well as (4.12) now go "one step backwards" and ask which best responses are compatible with Q being the aggregate in the first place:

$$B_i(Q) = \{s_i \in S_i : s_i \in R_i(Q - s_i)\}. \tag{4.13}$$

Definition 6 (Backward-response Correspondence, Linearly Aggregative Games) *In a linearly aggregative game, the backward-response correspondence of an agent i is the set-valued mapping* $B_i : X \to 2^{S_i} \cup \emptyset$ *defined in* (4.13)*. The aggregate backward-response correspondence* $Z : \mathbb{R} \to 2^{\mathbb{R}} \cup \emptyset$ *is,*

$$Z(Q) = \left\{\sum_j s_j : s_j \in B_j(Q) \textit{ for all } j = 1, \ldots, I\right\}. \tag{4.14}$$

The previous construction goes through step-by-step for generalized aggregative games, the only added complexity involving notation. Since any generalized aggregative game (Definition 2) can be summarized by the tuple $\{(\pi_i, S_i)_{i=1,\ldots,I}, g\}$ the *best-response correspondence* of a player i is given by,

$$r_i(s_{-i}) = \arg\max_{s_i \in S_i} \Pi_i(s_i, g(s)) = \arg\max_{s_i \in S_i} \Pi_i\left(s_i, H\left(h_i(s_i) + \sum_{j \neq i} h_j(s_j)\right)\right). \tag{4.15}$$

[5] Selten called it the "Einpassungsfunktion", which roughly translates into the "fitting-in function", which is an alternative name sometimes encountered (*e.g.*, Phlips, 1995). Yet another name for it is "replacement function" (Cornes and Hartley, 2005). The term "backward- (best-) response function" is due to Novshek (1985).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Comparing with (4.11), we see that the appropriate aggregate of opponents' strategies is now $\sum_{j\neq i} h_j(s_j)$ and once we realize this we can define the reduced best-response correspondence:

$$R_i(x_{-i}) = \arg\max_{s_i \in S_i} \Pi_i \left(s_i, H\left(h_i(s_i) + x_{-i}\right)\right). \tag{4.16}$$

Note that, by construction, $r_i(s_{-i}) = R_i\left(\sum_{j\neq i} h_j(s_j)\right)$ for all $s \in S$. Now fix an *aggregate* $Q \in X \equiv \{g(s) : s \in S\}$; and note that $Q = g(s) \Leftrightarrow \sum_{j\neq i} h_j(s_j) = H^{-1}(Q) - h_i(s_i)$. Hence we have:

Definition 7 (Backward-response Correspondence, Generalized Aggregative Game) *Let* $\{(\Pi_i, S_i)_{i=1,\ldots,I}, g\}$ *be a generalized aggregative game:*

- *The backward-response correspondence of agent i,* $B_i : X \to 2^{S_i} \cup \emptyset$*, is:*

$$B_i(Q) \equiv \{s_i \in S_i : s_i \in R_i(H^{-1}(Q) - h_i(s_i))\}, \tag{4.17}$$

where R_i *was given in* (4.16).
- *The aggregate backwards-response correspondence,* $Z : X \to 2^X \cup \emptyset$ *is,*

$$Z(Q) \equiv \{g(s) \in X : s_i \in B_i(Q) \textit{ for all } i \in \mathscr{I}\} \tag{4.18}$$

Proposition 1 *An aggregate* $Q \in X$ *is an equilibrium aggregate if and only if it is a fixed point of the aggregate backward-response correspondence,*

$$Q \in Z(Q). \tag{4.19}$$

Note that if Q is an equilibrium aggregate then any $(s_1^*, \ldots, s_I^*) \in B_1(Q) \times \ldots \times B_I(Q)$ with $g(s^*) = Q$ must necessarily be a (pure-strategy) Nash equilibrium. Thus we have established a direct correspondence between the set of equilibria on the one hand, and the set of equilibrium aggregates on the other.

To investigate a generalized aggregative game's equilibria we may therefore proceed by computing Z and solve the simple one-dimensional fixed-point problem (4.19). By Proposition 1 we would be assured that any fixed-point Q determines a pure-strategy Nash equilibrium. This allows us to shift our attention entirely to the one-dimensional problem expressed via Z, which reduces the complexity enormously. To this should be added that, in many situations, the aggregate is actually more interesting from an applied perspective than the individual strategies. We return to this theme in Section 4.

3.2 Nice Aggregative Games, Computation

Quite a large literature, of which Selten (1970) is most definitely a part, directly compute (or at least characterize) the aggregate backward-response correspondence and then employ Proposition 1 to address various questions. For an illustration, see Cornes and Hartley 2005 who study asymmetric contests and in the case of linear technologies directly compute the

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

backward-reply function (Equation (6), p. 927) which they then use to establish uniqueness of equilibrium and put bounds on rent dissipation (Theorems 1–2, pp. 929–930). For a more abstract illustration, see Novshek (1985) who characterizes the aggregate backward-response correspondence in a linearly aggregative game and uses this to prove existence of an equilibrium when the game exhibits strategic substitutes. We return to this in Section 3. Like Cornes and Hartley, (2005) but unlike Novshek (1985), the current section focuses on cases where the backward-response correspondence can be computed directly via first-order conditions.

In order to be able to compute backward-response correspondences directly and without too much fuss, one needs first-order conditions to be necessary and sufficient for optimality. Conditions for this are, of course, well known and indeed are imposed in studies such as Selten (1970) and Cornes and Hartley (2005) mentioned a moment ago. Nonetheless, let us for later purposes single out a fairly general class where first-order conditions convey all information about agents' optimal behavior. To this end, recall first that a differentiable function π_i is pseudo-concave (Mangasarian, 1965) in s_i if for all $s_i, s_i' \in S_i$:

$$(s_i' - s_i)^T D_{s_i}\pi_i(s_i, s_{-i}) \leq 0 \;\Rightarrow\; \pi_i(s_i', s_{-i}) \leq \pi_i(s_i, s_{-i}).$$

Naturally, any concave function is pseudo-concave.[6]

Definition 8 (Nice Aggregative Games) *A generalized aggregative game* $\left((\Pi_i, S_i)_{i=1}^I, g\right)$ *is a nice aggregative game if:*

1. *the aggregator g is twice continuously differentiable;*
2. *each strategy set S_i is compact and convex, and every payoff function $\pi_i(s,t) = \Pi_i(s_i, g(s))$ is twice continuously differentiable, and pseudo-concave in the player's own strategies;*
3. *for each player, the first-order conditions hold whenever a boundary strategy is a (local) best response, i.e., $D_{s_i}\Pi_i(s_i, g(s)) = 0$ whenever $s_i \in \partial S_i$ and $(v - s_i)^T D_{s_i}\Pi_i(s_i, g(s)) \leq 0$ for all $v \in S_i$.*

Condition (3) ensures that if a best response lies at the boundary of the strategy set, then it is captured by the first-order conditions too. It can be replaced with an Inada condition (since under Inada conditions *no* optimal strategy lies on the boundary). Alternatively, one could work with Kuhn-Tucker conditions or even more advanced tools from optimization theory/convex analysis and modify what follows accordingly. None of this poses any real difficulties, so here we stick to the simplest case. Also, in many situations (3) can be removed altogether (see Section 4) so generalizing it here is somewhat of a wasteful effort.

[6] A quasi-concave function is not necessarily pseudo-concave. Not surprisingly so, since first-order conditions are *not* sufficient for an interior optimum of a quasi-concave function (to see this just consider the function $f(x) = x^3$).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Write the marginal payoff for player i in terms of the reduced payoff function:

$$D_{s_i}\pi_i(s,t) = D_1\Pi_i(s_i, g(s)) + D_2\Pi_i(s_i, g(s))H'(H^{-1}(g(s))Dh_i(s_i), \tag{4.20}$$

where $D_m\Pi_i(s_i, g(s)) \equiv D_{x_m}\Pi_i(x_1, x_2)_{|(x_1,x_2)=(s_i,g(s))}$, $m = 1, 2$.

The thing to note is that the marginal payoff is a function of the player's own strategy s_i and the aggregate $g(s)$ *only*. A game with this feature is called *fully aggregative* in Cornes and Hartley (2012), who go on to show that when $I \geq 3$ and all involved functions are at least twice continuously differentiable, (4.20) implies that g *must* be additively separable (Cornes and Hartley, 2012, Proposition 1). To put it differently, Cornes and Hartley (2012) show that, under the mentioned conditions, a game is fully aggregative if and only if it is generalized aggregative. This result is very significant because it shows, in effect, that any analysis based on backward-response correspondences is "immune" to generalizations beyond generalized aggregative games (in particular, it cannot apply to quasi-aggregative games in general).

After this detour, let us return to (4.20). As mentioned, payoffs depend on own strategies and the aggregate only. We may define a function $\Psi_i : S_i \times X \to \mathbb{R}^N$ that makes this feature explicit,[7]

$$\Psi_i(s_i, Q) \equiv D_1\Pi_i(s_i, Q) + D_2\Pi_i(s_i, Q)H'(H^{-1}(Q))Dh_i(s_i). \tag{4.21}$$

In linearly aggregative games, Ψ_i is precisely the function used by Corchón (1994) to establish comparative statics results in linearly aggregative games (Corchon denotes this function by T_i). And we shall return to Ψ_i when we consider comparative statics in Section 4 and also when we turn to uniqueness in Section 6.

Proposition 2 *In a nice aggregative game,*

$$s_i \in B_i(Q) \Leftrightarrow \Psi_i(s_i, Q) = 0\,. \tag{4.22}$$

In particular, one can for any given $Q \in X = \left\{\sum_j s_j : s_j \in S_j \text{ for } j = 1, \ldots, I\right\}$ *compute* $B_i(Q)$ *by solving the n equations* $\Psi_i(s_i, Q) = 0$.[8]

Throughout this survey, we shall again and again encounter applications of the relationship in (4.22). But already considering what has been said so far, it is clear that using Proposition 2 one can compute every backward-response correspondence, and thus the aggregate backward-response correspondence Z. It is in this step, then, that the dimensionality of the problem reduces from $n \times I$ to 1. Then one can use Proposition 1 to conclude that the fixed points of Z correspond to the equilibria of the original game.

---

7 Here the variables s_i and Q are *independent* arguments in Ψ_i, so that Q is kept *fixed* when taking the derivative of Ψ_i. Hence, *e.g.*, $D_{s_i}\Psi_i(s_i, Q) = \frac{\partial \Psi_i(x_1,x_2)}{\partial x_1}_{|(x_1,x_2)=(s_i,Q)}$ when $N = 1$.

8 Recall here that n is the dimension of S_i, in particular there is just a single equation to solve if S_i is one-dimensional.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

4 COMPARATIVE STATICS

We are going to focus in this section on the class of generalized aggregative games (Section 2.2). To set the stage, consider a generalized aggregative game just as in Section 2.2 but allow now for the explicit consideration of exogenous parameters:

$$\Gamma_t \equiv \{(\Pi_i, S_i)_{i=1,\ldots,I}, g, t\}$$

So the game is Γ_t and it explicitly depends on the parameter t, which we are going to assume lives in a set $T \subseteq \mathbb{R}$. While $S_i \subseteq \mathbb{R}^n$ (agent i's strategy set) and the additively separable aggregator $g : S \to X \subseteq \mathbb{R}$ are just as before, we now allow the reduced payoff functions to depend on a parameter $\Pi_i : S_i \times X \times \{t\} \to \mathbb{R}$, $i = 1, \ldots, I$.[9] In the Cournot model, say, we might have $\Pi_i(s_i, g(s), t) = s_i P\left(\sum_j s_j\right) - ts_i$, which is to say that the parameter t is the unit cost of production.

Let $E(t)$ denote the (at this point possibly empty) set of equilibria:

$$E(t) = \{s^* \in S : \Pi_i(s_i^*, g(s^*), t) \geq \Pi_i(s_i, g(s_i, s_{-i}^*), t) \text{ for all } s_i \in S_i; i = 1, \ldots, I\} \tag{4.23}$$

The fundamental equilibrium comparative statics question, stated here as abstractly (but also as generally) as possibly, is this:

Imagine that the parameter t is changed from t' to t'' say. Then how is the set of equilibria going to change, i.e., how does $E(t'')$ compare to $E(t')$?

This question is difficult because it involves the comparative statics of equilibria. One situation where we *can* deal with it very generally is when a game exhibits *strategic complementarities* (Bulow, Geanakoplos and Klemperer 1985, Vives, 1990), *i.e.*, if agents' strategies are non-decreasing in opponents' strategies.[10] An example is a Bertrand oligopoly where each price-setting firm will, under standard assumptions, choose to *raise* its price if one or more of the competitors raise their prices. Roughly speaking, one can in this situation conclude that if best responses are increasing in t, then the set of equilibria $E(t)$ will also be increasing in t; where *increasing* in both instances is in the set-valued sense of the strong set order (Topkis, 1998).[11] Such results carry over *ipso facto* if a game is also aggregative.

As we shall see below in Section 4.2, the aggregative games framework allows us to also derive general results in games of strategic substitutes.[12] But as we will see to begin with, neither of the two (strategic complements or substitutes) are required if we are within the setting of the nice games of Section 3.2. See also Roy and Saberwal (2010, 2012).

---

[9] Note that there is really no problem extending existing results to allow strategy sets to depend on the exogenous parameter also. We are not going to pursue the issue here though.

[10] For a lengthy treatment of this case, the reader is referred to Vives (2000). See also Milgrom and Roberts (1990).

[11] If best responses are always unique (strong concavity), and $E(t)$ is always a singleton (uniqueness of equilibrium), the statement reduces to saying that if t increases, $t' < t''$, all strategies in $E(t'')$ will be coordinate-wise greater than or equal to the strategies in $E(t')$.

[12] In Section 3 we shall see that a similar statement applies to existence of equilibrium where strategic complementarities also allow for very general results.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

4.1 Nice Aggregative Games

A game is now a tuple $\left((\Pi_i, S_i)_{i=1}^I, g, t\right)$ where $t \in T \subseteq \mathbb{R}$ is an exogenous parameter. The definition of a nice game (Definition 8) carries over to the parameterized setting as long as we require that it holds for all $t \in T$. This shall be the framework of the rest of this section. Acemoglu and Jensen (2013) (Theorem 5, p. 36) establish the existence of equilibrium in nice aggregative games, *i.e.*, that $E(t) \neq \emptyset$ in terms of the current notation. Given an equilibrium $s(t) \in E(t)$, let $Q(t) = g(s(t))$ denote the associated equilibrium aggregate (*cf.* Definition 2 and Proposition 1). The theorem in Acemoglu and Jensen just referred to, goes on to show that there exist a *smallest* and a *largest* equilibrium aggregate, $Q_*(t)$ and $Q^*(t)$ and that the functions of t thus defined are, respectively, lower semi-continuous and upper semi-continuous.[13] Note that, intuitively, all equilibrium aggregates "live" in the interval $[Q_*(t), Q^*(t)]$ and so if both $Q_*(t)$ and $Q^*(t)$ increase with t, the set of equilibria increases up to the ambiguity of multiplicity. If the equilibrium is unique (Section 6), such ambiguity disappears of course, since then $Q_*(t) = Q^*(t)$.

Next, recall the function Ψ_i defined in equation (4.24), which when an exogenous parameter t explicitly is allowed for reads,

$$\Psi_i(s_i, Q, t) \equiv D_1\Pi_i(s_i, Q, t) + D_2\Pi_i(s_i, Q, t)H'(H^{-1}(Q))Dh_i(s_i). \tag{4.24}$$

If we fix Q and t and differentiate Ψ_i with respect to s_i, we get an $N \times N$ matrix

$$D_{s_i}\Psi_i(s_i, Q, t) \in \mathbb{R}^{N\times N}.$$

The determinant of this matrix is denoted by $|D_{s_i}\Psi_i(s_i, Q, t)| \in \mathbb{R}$. If strategy sets are one-dimensional, the determinant coincides with the derivative: $|D_{s_i}\Psi_i(s_i, Q, t)| = D_{s_i}\Psi_i(s_i, Q, t) \in \mathbb{R}$.

Consider now the following conditions, which, as one easily sees, become increasingly stringent:

Definition 9 (Local Solvability) *Player $i \in \mathscr{I}$ is said to satisfy the local solvability condition if $\Psi_i(s_i, Q, t) = 0 \Rightarrow |D_{s_i}\Psi_i(s_i, Q, t)| \neq 0$ for all $s_i \in S_i$, $Q \in X$, and $t \in T$.*

Definition 10 (Uniform Local Solvability) *When $S_i \subseteq \mathbb{R}$, player $i \in \mathscr{I}$ is said to satisfy the uniform local solvability condition if $\Psi_i(s_i, Q, t) = 0 \Rightarrow D_{s_i}\Psi_i(s_i, Q, t) < 0$ for all $s_i \in S_i$, $Q \in X$, and $t \in T$.*

Definition 11 (Strong Concavity) *When $S_i \subseteq \mathbb{R}$, player $i \in \mathscr{I}$ is said to satisfy the strong concavity condition if $D_{s_i}\Psi_i(s_i, Q, t) < 0$ and $D_Q\Psi_i(s_i, Q, t) < 0$ for all $s_i \in S_i$, $Q \in X$, and $t \in T$.*

---

13 That $Q_*(t)$ is the smallest equilibrium aggregate simply means that for any $s(t) \in E(t)$, we have $g(s(t)) \geq Q_*(t)$. The largest equilibrium aggregate is defined similarly by reversing the inequality.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Local solvability and uniform local solvability were introduced in Acemoglu and Jensen (2013), while strong concavity is due to Corchón (1994). As we shall see in a moment, either of these conditions provide robust comparative statics results. Before getting to that an example is in order. Note that this example is taken directly from Acemoglu and Jensen (2013) where more examples can also be found:

Example 1 *Take the Cournot model where* $\Pi_i(s) = s_i P(\sum_j s_j) - c_i(s_i)$ *and so* $\Psi_i(s_i, Q) = P(Q) + s_i P'(Q) - c_i'(s_i)$ *(suppressing here exogenous parameters). Hence the local solvability condition will hold if either* $D_{s_i}\Psi_i(s_i, Q) = P'(Q) - c_i''(s_i) < 0$ *or* $D_{s_i}\Psi_i(s_i, Q) = P'(Q) - c_i''(s_i) > 0$ *whenever* $P(Q) + s_i P'(Q) - c_i'(s_i) = 0$. *If the first of the two holds whenever* $P(Q) + s_i P'(Q) - c_i'(s_i) = 0$, *the uniform local solvability condition is satisfied. For example, this will be the case when costs are convex and inverse demand is strictly decreasing (these conditions are clearly not necessary).*

Before we get to the results, we need to consider conditions on how the exogenous parameter enters the payoff function. If strategy sets are one-dimensional ($S_i \subseteq \mathbb{R}$ for all i), and more generally if strategy sets are lattices and payoff functions supermodular in own strategies, we can define *positive shocks* in the standard way known from games with strategic complementarities (see *e.g.*, Vives, 2000). Recall again that if $t \in T \subseteq \mathbb{R}^M$, $M > 1$, then an increase in t means that at least one of t's coordinates increases.

Definition 12 (Positive Shocks) *Consider the payoff functions* $\pi_i(s, t) = \Pi_i(s_i, g(s), t)$. *Then an increase in t is a positive shock if each* S_i *is a lattice, and* π_i *is supermodular in* s_i *and exhibits increasing differences in* s_i *and t. In particular, if* $S_i \subseteq \mathbb{R}$ *for all i, then t is a positive shock if each* π_i *exhibits increasing differences in* s_i *and t.*

Note that in a linearly aggregative game where $\pi_i(s, t) = \Pi_i(s_i, \sum_j s_j, t)$, an increase in t is a positive shock if $D^2_{s_i t}\Pi_i(s_i, Q, t) + D^2_{Qt}\Pi_i(s_i, Q, t) \geq 0$:

Theorem 1 (Aggregate Comparative Statics) *Consider a nice aggregative game where each player's payoff function satisfies the local solvability condition. Then a positive shock* $t \in T$ *leads to an increase in the smallest and largest equilibrium aggregates, i.e., the functions* $Q_*(t)$ *and* $Q^*(t)$ *will be increasing in t. When strategy sets are one-dimensional and each player's payoff function satisfies the uniform local solvability condition, the result remains valid without imposing the boundary condition (3) of Definition 8.*

In industrial organization, it is often of great interest what happens when additional agents (typically, additional firms) enter a market. The next result, also from Acemoglu and Jensen (2013) answers this question. For a related, albeit weaker result that instead is based on strong concavity, see Corchón (1994).

Theorem 2 (Entry) *Under the conditions of Theorem 1 entry of an additional player increases the smallest and largest equilibrium aggregates, i.e., if* $Q_*(I)$ *and* $Q^*(I)$ *denote the smallest and largest equilibrium aggregates in a game with* $I \in \mathbb{N}$ *players then* $Q_*(I) \leq Q_*(I+1)$ *and* $Q^*(I) \leq Q^*(I+1)$ *for all* $I \in \mathbb{N}$. *The previous inequalities will be strict if the entrant does not choose the "inaction" strategy* $\inf S_{I+1}$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

The previous results both predict how the aggregate changes in equilibrium. Acemoglu and Jensen (2013) also consider a different type of shock that roughly speaking "hits the aggregator directly" and derives parallel results. In many situations, results that predict how the aggregate changes with exogenous parameters are sufficient. In some situations (such as when the conditions below do not hold) it may also be that there are simply no more results to be found; in other words, the equilibrium aggregate may respond predictably to a change in an exogenous parameter, but nothing can be said in general about individual strategies.[14]

Theorem 3 (Individual Comparative Statics) *Let the conditions of Theorem 1 be satisfied and consider player i's equilibrium strategy $s_i^*(t)$ associated with the smallest (or largest) equilibrium aggregate at some equilibrium $s^* = s^*(t)$ given $t \in T$. Assume that the equilibrium s^* lies in the interior of S and that t is a positive shock. Then the following results hold:*

- *$s_i^*(t)$ is (coordinate-wise) locally increasing in t provided that*

$$-[D_{s_i}\Psi_i(s_i^*, g(s^*), t)]^{-1} D_Q \Psi_i(s_i^*, g(s^*), t) \geq 0.$$

- *Suppose that the shock t does not directly affect player i (i.e., $\pi_i = \pi_i(s)$). Then the sign of each element of the vector $D_t s_i^*(t)$ is equal to the sign of each element of the vector $-[D_{s_i}\Psi_i(s_i^*, g(s^*))]^{-1} D_Q \Psi_i(s_i^*, g(s^*))$. In particular, $s_i^*(t)$ will be (coordinate-wise) locally decreasing in t whenever:*

$$-[D_{s_i}\Psi_i(s_i^*, g(s^*))]^{-1} D_Q \Psi_i(s_i^*, g(s^*)) \leq 0.$$

4.2 Strategic Substitutes

In this section, the following assumption will be in force throughout:

Assumption 1 *The component functions of the aggregator $H, h_1, \ldots, h_I$ are all continuous, and for each agent $i = 1, \ldots, I$, the strategy set $S_i \subseteq \mathbb{R}$ is compact and the reduced payoff function Π is continuous on $S_i \times X \times T$.*

Note that the restriction to one-dimensional strategy sets is for convenience only.[15]

Definition 13 (Strategic Substitutes) *Consider the payoff functions $\pi_i(s, t) = \Pi_i(s_i, g(s), t)$, $i = 1, \ldots, I$. The game is a game of strategic substitutes if each π_i exhibits decreasing differences in s_i and s_j for all $j \neq i$.*

---

[14] See Acemoglu and Jensen (2015) for a lengthy discussion of this issue and its economic significance.

[15] With multi-dimensional strategy sets, strategy sets must be lattices and payoff functions must be supermodular in own strategies. All results then go through if $S_i \subseteq \mathbb{R}^n$ (see Acemoglu and Jensen, 2013).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

It should be mentioned that all that is required to get results is that the least and greatest selections from the reduced best-response correspondence $R_i(\cdot, t)$ are non-increasing. This will hold if π_i satisfies the reverse single-crossing property in s_i and x_{-i} (Milgrom and Shannon, 1994). Decreasing differences has the advantage of being simple to characterize when Π_i is sufficiently smooth, and this will in turn allow us to compare various results in the existing literature. To be precise, if Π_i and g are twice differentiable, decreasing differences in s_i and s_j holds if and only if:

$$\frac{\partial^2 \Pi_i(s_i, g(s), t)}{\partial s_i \partial s_j} \leq 0. \tag{4.25}$$

If the game is linearly aggregative ($g(s) = \sum_j s_j$), this holds if and only if for all $s_i \in S_i$, $Q = \sum_j s_j$ and $t \in T$,

$$D_Q \Psi_i(s_i, Q, t) = D^2_{12}\Pi_i(s_i, Q, t) + D^2_{22}\Pi_i(s_i, Q, t) \leq 0. \tag{4.26}$$

Note that (4.26) is implied by strong concavity (Definition 11). Strong concavity thus implies uniform local solvability (Definition 10) as well as strategic substitutes. It is of some interest to note, then, that strong concavity is equivalent to the very familiar Hahn (1962) conditions of the Cournot model (see Corchón, 1994, p. 156, and also Vives, 2000).

We say that an increase in t is a positive idiosyncratic shock to agent i, if the shock is a positive shock (Definition 12) and the parameter/variable t affects *only* the payoff function of player i, *i.e.*,

$$\Pi_i = \Pi_i(s_i, g(s), t) \text{ and } \Pi_j = \Pi_j(s_j, g(s)) \text{ for all } j \neq i. \tag{4.27}$$

Note that the literature usually denotes a parameter that affects only one of the agents by the agent's index. So we would normally denote the parameter t by t_i and write $\Pi_i(s_i, g(s), t_i)$ in place of $\Pi_i(s_i, g(s), t)$ in (4.27).

Theorem 4 (The Comparative Statics of Idiosyncratic Shocks) *A positive idiosyncratic shock to player i will increase the smallest and largest equilibrium strategies for player i, and decrease the associated aggregates of the remaining players (which are, respectively, the largest and smallest such aggregates).*

Corollary 1 (Payoff Effects) *Assume in addition to the conditions of Theorem 4 that all payoff functions are decreasing [respectively, increasing] in opponents' strategies and that player i's payoff function is increasing [respectively, decreasing] in the idiosyncratic shock* t_i*. Then an increase in* t_i *increases [respectively, decreases] player i's payoff in equilibrium and decreases [respectively, increases] the payoff of at least one other player.*

Theorem 5 (The Comparative Statics of Entry) *In an aggregative game with strategic substitutes, entry of an additional player leads to a decrease in the smallest and largest aggregates of the existing players in equilibrium.*

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

5 EXISTENCE AND BEST-RESPONSE POTENTIALS

When faced with an aggregative game, whether linearly, generalized, or quasi-aggregative, the standard routes to existence of an equilibrium apply:

- If payoff functions are quasi-concave and upper semi-continuous in own strategies, continuous in opponents' strategies, and strategy sets are compact and convex, existence can be proved by a standard application of Kakutani's fixed-point theorem. In particular, any nice game (Definition 8) has an equilibrium.
- If the game exhibits strategic complementarities (see the beginning of Section 4), one can instead prove existence of equilibrium via Tarski's (1955) fixed-point theorem or a set-valued version of this result. For details see Topkis (1998) or Vives (2000).

Now, as discussed in the introduction, neither of the two cases above necessarily sit particularly well with certain applications — not least if we are thinking of applications to industrial organization. To see why, it suffices to look at the Cournot model, which *does not* exhibit strategic complementarities except for a few rather special cases (Amir, 2005), and where it may *not* be satisfactory to assume quasi-concavity because monopolistic firms arguably could be characterized by profit functions that violate quasi-concavity due to returns to scale, fixed costs, etc. The Cournot model *does* exhibit strategic substitutes under fairly weak conditions, however, and this can be exploited in aggregative games.

In light of what we know about comparative statics, the gap is equally evident: we have existence as well as robust comparative statics results for nice games and games of strategic complementarities. We also have robust comparative statics results for games of strategic substitutes. The purpose of what follows is to establish existence at a level of generality sufficient to include the setting of Section 4.2, *i.e.*, without assuming convexity of strategy sets and/or quasi-concavity of payoff functions.

The presentation chosen here is different from that of the other sections where simplicity was always favored over generality. Here we go in the opposite direction and firstly present a very general existence result. This is then used to state existence in a number of special cases as corollaries. For even more general conditions – in fact, conditions that are likely to exhaust the topic – the interested reader is referred to Kukushkin (2016). The following result is taken from Jensen (2010), which in turn builds on earlier contributions by Dubey et al. (2006) and Kukushkin (2005):

Proposition 3 *A generalized quasi-aggregative game (Definition 5) has a pure-strategy Nash equilibrium under the following conditions:*

1. *For each agent $i = 1, \dots, I$, the strategy set $S_i \subseteq \mathbb{R}$ is compact and the reduced payoff function $\Pi = \pi_i(s_i, x_{-i})$ is upper semi-continuous in $s_i \in S_i$ and continuous in $x_{-i} \in X_{-i}$.*
2. *For all $i = 1, \dots, I$, the interaction function $\sigma_i : S_{-i} \to X_{-i}$ is continuous, and the shift-function $F_i : S_i \times X_{-i} \to \mathbb{R}$ is continuously differentiable and (possibly after a monotone transformation) exhibits strictly increasing differences in s_i and x_{-i}.*[16]

[16] Since $X_{-i} \subseteq \mathbb{R}$ and F_i is continuously differentiable, the strictly increasing differences condition will hold if and only if $D_{x_{-i}} F_i(s_i, x_{-i})$ is (strictly) increasing in every coordinate of s_i. When $S_i \subseteq \mathbb{R}$ and F_i is twice differentiable,

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

3. *Each of the reduced best-response correspondences $R_i : X_{-i} \to 2^{S_i}$ has a decreasing selection, i.e., for every $i = 1, \ldots, I$, there exists a function $\hat{r}_i : X_{-i} \to S_i$ with $\hat{r}_i(x_{-i}) \in R_i(x_{-i})$ for all $x_{-i} \in X_{-i}$ such that $\hat{r}_i(\tilde{x}_{-i}) \geq \hat{r}_i(x_{-i})$ whenever $\tilde{x}_{-i} \leq x_{-i}$.*

It should be emphasized that Proposition 3 does not assume that strategy sets are convex and/or that payoff functions are quasi-concave in own strategies. The reader who is interested in an application that uses the previous proposition directly is referred to Shirai (2010) who establishes the existence of Cournot-Walras equilibria in the standard model of monopolistic competition. Here we are going to focus on a number of special cases.

Corollary 2 (Novshek, 1985, Kukushkin, 1994) *Let $(\Pi_i, S_i)_{i=1}^I$ be a linearly aggregative game of strategic substitutes that satisfies (1) of Proposition 3. Then there exists an equilibrium.*

Proof The game is generalized quasi-aggregative with $\nu_i = 0$, $F_i(s_i, x_{-i}) = s_i + x_{-i}$, and $\sigma_i(s_{-i}) = \sum_{j \neq i} s_j$, for $i = 1, \ldots, I$. All conditions are immediately satisfied with the exception of F_i exhibiting strictly increasing differences (in fact it does not!). A monotone transformation of F_i does exhibit strictly increasing differences, namely $h \circ F_i(s_i, x_{-i}) = \exp(F_i(s_i, x_{-i}))$. ■

Corollary 3 *Let $(\Pi_i, S_i)_{i=1}^I$ be a generalized aggregative game with a continuous aggregator g that exhibits strategic substitutes and satisfies Condition (1) of Proposition 3. Then there exists an equilibrium.*

Proof The game is generalized quasi-aggregative with $\nu_i = 0$, $F_i(s_i, x_{-i}) = H(h_i(s_i) + x_{-i})$ and $\sigma_i = h_i$ where $H, h_1, \ldots, h_I$ are the strictly increasing component functions of the additively separable function g. All conditions are immediately satisfied, in particular, the transformation $h \circ F_i = \exp \circ H^{-1} \circ F_i$ exhibits strictly increasing differences. ■

Definition 14 (Reciprocal Interactions) *A game is said to be a game with reciprocal interactions, if $\Pi_i(s) = \pi_i(s_i, \sigma_i(s_{-i}))$ for all i, where the σ_i's are real-valued, continuously differentiable interaction functions that satisfy:*

$$\frac{\partial \sigma_i(s_{-i})}{\partial s_j} = \frac{\partial \sigma_j(s_{-j})}{\partial s_i}, \text{ for all } i, j \in \mathscr{I} \text{ and all } s \in \hat{S}, \tag{4.28}$$

where $\hat{S}$ is an open, convex subset of $\mathbb{R}^I$ that contains S.

Specific instances of games with reciprocal interactions are studied in Dubey et al. (2006) and the general class was defined by Kukushkin (2005) (the name is due to Kukushkin). A particularly interesting example of a game with reciprocal interactions is a linearly aggregative game where players' payoff functions only depend on subsets of opponents' strategies, *i.e.*,

increasing differences is equivalent to having $D^2_{s_i x_{-i}} F_i(s_i, x_{-i}) > 0$. Note that since monotone transformations of F_i are allowed, each of the previous statements actually apply to such a monotone transformation (if one is needed), but the notation gets too messy to justify its writing!

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

where agents are positioned in a symmetric network. Then we have $\Pi_i(s) = \pi_i(s_i, \sum_{j \in N(i)} s_j)$ for all i where $N(i) \subseteq \{1, \ldots, I\}$ and $j \in N(i) \Leftrightarrow i \in N(j)$ (symmetry). For example, we might have four firms who established along a circle and are only influenced by their neighbors in an otherwise standard (linearly) aggregative game: $\Pi_1(s) = \pi_1(s_1, s_2 + s_4)$, $\Pi_2(s) = \pi_2(s_2, s_1 + s_3)$, $\Pi_3(s) = \pi_3(s_3, s_2 + s_4)$, and $\Pi_4(s) = \pi_4(s_4, s_1 + s_3)$. It is easily seen that (4.28) is satisfied.

Corollary 4 (Dubey et al., 2006, Kukushkin, 2005) *Let $(\Pi_i, S_i)_{i=1}^I$ be a game with reciprocal interactions where each σ_i is either strictly increasing or strictly decreasing in $s_{-i} \in S_{-i}$ and let (1) and (3) of Proposition 3 be satisfied. Then the game is generalized quasi-aggregative and there exists an equilibrium.*[17]

Proof The proof consists in showing that the game is generalized aggregative with associated functions g, F_i, ν_i, $i = 1, \ldots, I$ that satisfy condition (2) of Proposition 3 (since then Proposition 2 applies). The interested reader is referred to Jensen, 2010, Section 2.3.2. ■

Let $\perp_i = \min_{s_{-i} \in S_{-i}} \sigma_i(s_{-i})$, $\top_i = \max_{s_{-i} \in S_{-i}} \sigma_i(s_{-i})$. Since F_i is continuously differentiable, the following function is well-defined:[18]

$$P^R(s_i, s_{-i}) = \sum_i \left[\int_{\perp_i}^{\top_i} \min\{D_2 F_i(s_i, \tau), D_2 F_i(\hat{r}_i(\tau), \tau)\}\, d\tau + F_i(s_i, \perp_i) \right] - g(s_i, s_{-i}). \tag{4.29}$$

Since P^R is a continuous function, $\nu = \max_{s \in S} P^R(s) - \min_{s \in S} P^R(s) + 1 > 1$ is well-defined. Now define:

$$P(s) = P^R(s) + \sum_i \chi_i(s_i), \tag{4.30}$$

where $\chi_i(s_i) = \nu$ if $s_i \in \{s_i : s_i \in R_i(\sigma_i(s_{-i}))$ for some $s_{-i} \in S_{-i}\}$ and $\chi_i(s_i) = 0$ otherwise. It is proved in Jensen, 2010 that P is an upper semi-continuous pseudo best-response potential under the conditions of Proposition 3.[19] That P is a pseudo best-response potential means that if we maximize it with respect to $s_i \in S_i$ for any i, then we get best responses as optimizers:

$$\arg\max_{s_i \in S_i} P(s) \subseteq \arg\max_{s_i \in S_i} \pi_i(s) \text{ for } i = 1, \ldots, I. \tag{4.31}$$

[17] To be accurate, Dubey et al. (2006) and Kukushkin (2005) prove existence. The statement on the quasi-aggregative structure of such games is found in Jensen, 2010.

[18] Note that we are skipping a technicality here related to the domain of $\hat{r}_i$ which may – since we have not assumed that strategy sets are convex – have to be extended from a subset of $[\perp_i, \top_i]$ to the entire set. See Jensen, 2010, proof of Theorem 1 for details.

[19] There is one case when the potential P^R applies directly (and in particular, the potential will then be a continuous function since so is P^R). This is when the selection $\hat{r}_i$ of Proposition 3 is continuous. Such a selection will map onto *an interval*, which allows one to extend the domain of the reduced best-reply selections and take $\hat{S}_i = S_i$ for all i above (see Dubey et al., 2006, p. 82, l.6-8 for details). This clearly works whether the aggregator is linear, as in that treatment, or not.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

It is clear then that any global maximum for P, *i.e.*, any,

$$s^* \in \arg\max_{s \in S} P(s),$$

must be an equilibrium (since at a global maximum all agents must be choosing best responses given opponents' strategies). Existence then follows from upper hemi-continuity of P, compactness of S, and Weierstrass's theorem.

If Condition (3) of Proposition 3 is strengthened to require that *every* selection from R_i is decreasing, then the game will in fact be a best-response potential game in the sense of Voorneveld (2000), which means that we get to put equality signs in (4.31):

$$\arg\max_{s_i \in S_i} P(s) = \arg\max_{s_i \in S_i} \pi_i(s) \text{ for } i = 1, \ldots, I. \tag{4.32}$$

Note that potentials are important for two reasons. First, the proof of existence just described is *substantially* simpler than using backward-response correspondences to prove existence. And, second, backward-response correspondences exist only up until the generality afforded by generalized aggregative games as discussed in Section 3.1. So whether for existence, as considered here, or for any other question of interest, one cannot use backward-response correspondences if one has, say, a network as the one described after Definition 14, or a teamwork game as described in Dubey et al. (2006) or Jensen, 2010.

And to be sure, it would be possible to address comparative statics as well as uniqueness (Section 6) using potentials. For example, if (4.32) holds and P is strictly concave, the equilibrium must be unique. This belongs to the realm of future research, however. Any contribution in this direction, whether to address uniqueness or comparative statics, would greatly expand the applicability of the aggregative games methodology.

6 UNIQUENESS

The state of the art on uniqueness in aggregative games exploits the backward-reply correspondence of Section 3.1 and so applies only to generalized aggregative games (see the discussion following (4.20)). As discussed at the end of Section 3, there is reason to believe that parallel results can be developed for more general classes of aggregative games using best-response potentials instead of backward-response correspondences. But currently this remains an open question.

In order to proceed we must first of all ensure that the aggregate backward-response correspondence is single-valued (a function). To the best of this author's knowledge, the most general known result that ensures this outcome is the following:[20]

[20] As discussed in Section 3.1, the conditions of nice games can of course be replaced with anything that ensures that first-order conditions characterize optimality. One can also move over to Kuhn-Tucker or related conditions in which case best responses are characterized by an equivalence of the type $s_i \in B_i(Q) \Leftrightarrow \Psi_i(s_i, Q) \geq 0$. Such extensions will not be pursued here.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Proposition 4 *Consider a nice (generalized) aggregative game* $\{(\Pi_i, S_i)_{i=1,\dots,I}, g\}$ *with one-dimensional strategy sets,* $S_i \subseteq \mathbb{R}$ *all i. Then the backward-reply correspondences are single-valued if the uniform local solvability condition (Definition 10) holds for all agents.*

Proof The uniform local solvability condition says that $\Psi_i(s_i, Q) = 0 \Rightarrow D_{s_i}\Psi_i(s_i, Q) < 0$ for all $s_i \in S_i$ and $Q \in X$. Since $s_i \in B_i(Q) \Leftrightarrow \Psi_i(s_i, Q) = 0$ in a nice game, the result follows directly from the continuity of Ψ_i (if the continuous function $\Psi_i(\cdot, Q)$ is strictly decreasing whenever it crosses the first axis, it can cross it at most once). ■

When the backward response correspondence is in fact a function, denote it by $b_i : X \to S_i$, and from it the aggregate backward-reply function:

$$z(Q) = g\left((b_i(Q))_{i=1,\dots,I}\right). \tag{4.33}$$

It follows from Proposition 1 in Section 3.1 that Q^* is an equilibrium aggregate if and only if

$$Q^* = z(Q^*). \tag{4.34}$$

If (4.34) has a unique solution, the game necessarily has a unique equilibrium, namely $s^* = (b_1(Q^*), \dots, b_I(Q^*))$. The uniqueness question thus reduces to a particularly simple one-dimensional question. In specific applications, it is definitely possible to attack this question straight on by computing, or at least characterizing z. But even then, an additional step has turned out to be exceptionally useful in applications. Define the *share function* (see *e.g.*, Cornes and Hartley, 2005a) by dividing the backward-response function with the aggregate:

$$\rho(Q) = \frac{z(Q)}{Q}. \tag{4.35}$$

Note that the share function is only defined for $Q \neq 0$. Say that an equilibrium aggregate is *non-trivial* if it is different from zero. Obviously, Q is then a non-trivial equilibrium aggregate if and only if $\rho(Q) = 1$ (to see this simply divide (4.34) through with Q and call on Proposition 1). In most applications, the fact that the share function cannot be defined for $Q = 0$ is not a major problem since we have $S_i \subseteq \mathbb{R}_+$ and $g(s) = 0 \Leftrightarrow s = 0$ (note that because g is additively separable, this in turn implies that $g : S \to \mathbb{R}_+$). One must of course take care then to either rule out that $s = 0$ is an equilibrium, or alternatively, be content with interpreting the following result as telling us that there is a unique "non-trivial" equilibrium:

Theorem 6 *Consider a generalized aggregative game* $\{(\Pi_i, S_i)_{i=1,\dots,I}, g\}$ *with* $S_i \subseteq \mathbb{R}_+$ *and* $g : S \to \mathbb{R}_+$. *Then if the share function* ρ *is a continuous function that is strictly decreasing in the neighborhood of any point where it equals* 1*, there exists at most one non-trivial equilibrium. In particular, there is a unique equilibrium if an equilibrium exists and* $s = 0$ *is not an equilibrium.*

Proof Since ρ is continuous, the fact that it strictly decreases when $\rho(Q) = 1, Q > 0$ implies that $\rho(\tilde{Q}) < 1$ for all $\tilde{Q} > Q$ since for there to exist a $\tilde{Q} > Q$ with $\rho(\tilde{Q}) = 1$ we would necessarily have ρ not strictly decreasing at $\tilde{Q}$. ■

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Just to be clear, Theorem 6 is a fairly straightforward extension of the method developed and exploited in *e.g.*, Cornes and Hartley, (2005). For an application that uses the previous result's full generality, see Jensen (2016) who proves uniqueness of an equilibrium in a general class of contests that, in particular, includes patent races (Jensen, 2016, Theorem 1). In many applications, the share function is everywhere strictly decreasing (and the proof of Theorem 6 is then trivial). An example is Cornes and Hartley (2005) (Proposition 3) who show that the share function must be strictly decreasing in asymmetric contests.

7 MISCELLANEOUS APPLICATIONS AND EXTENSIONS

7.1 Large Aggregative Games, Mean-field Games

In the previous treatment, it was assumed that the set of players is finite. A large and rapidly growing literature considers so-called *large aggregative games* that are populated by a non-atomic measure space of agents. Large aggregative games are closely related to so-called *mean-field games* (Lasry and Lions, 2007, Huang, Caines, and Malhame, 2007), which have attracted great interest from the mathematics and engineering community (see below).

Now, if we are considering the definitions of Section 2, there is in each case nothing in the way of allowing the set of agents to be, say, the unit interval with the Lebesgue measure and Borel algebra. Also, there is no inherent difficulty in allowing strategies to be chosen from more general sets than subsets of $\mathbb{R}^n$. For example, strategy sets could be compact metric spaces, which is general enough to include Bayesian games. In both cases, the question that arises is how to suitably extend the notion of an aggregator. With $i \in [0,1]$, say, linearly aggregative games have a direct extension to a large aggregative game taking,

$$g(s) = \int_0^1 s(i)\, di.$$

More generally, one can easily allow for additively separable aggregators also in the infinite dimensional case (see Vind and Grodal, 2003 for the relevant definitions). When strategies are random variables so that the integral defining the aggregator would "integrate across random variables", such integrals must be carefully defined with reference to some law of large numbers. If all of this is done with care, one effectively recovers the deterministic aggregates considered everywhere above. For details of the more technical issues and further references, the reader is referred to Acemoglu and Jensen (2010) who establish robust comparative statics results similar in spirit to the results of Section 4. Notably, Acemoglu and Jensen, 2010 also allow for multi-dimensional aggregates, *i.e.*, situations where g maps joint strategies into $\mathbb{R}^N$ where N may be greater than 1.

When an aggregative game is large, agents' influence on the aggregate becomes infinitesimal, which simplifies the analysis very considerably.[21] To put it in simple terms, agents will take the aggregate as given so agents do not have to consider their own effect on the aggregate

[21] This statement comes with a caveat of course, namely that one instead faces issues with integration and the law of large numbers.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

when choosing strategies.[22] At the same time, large games can be seen as approximating games with a finite but large set of players (see again Acemoglu and Jensen, 2010 as well as, for example, Adlakha and Johari, 2013). Mean-field games mentioned above take precisely this route: a mean-field game is a linearly aggregative game populated by a finite set of agents but one that is large enough for the large game limit to be a good approximation. Generally, mean-field games allow for multi-dimensional strategy sets that with linear aggregators implies multi-dimensional aggregates, $g(s) = \sum_i s_i$ where $s_i \in S_i \subseteq \mathbb{R}^n$. As is clear, there are more similarities than differences between large aggregative games and mean-field games (or, at least, the large limit of mean-field games), and future research will no doubt lead to further integration.

To mention a few studies within this literature, Camacho, Kamihigashi, and Saglam (2016) establish robust comparative statics results in large aggregative games that extend the result of Acemoglu and Jensen (2010) to changes in distributions of parameters across agents. This allows them to study inequality in a simple model of income distribution (for the reader who is interested in the comparative statics of changing distributions more generally, see also Jensen, 2015). Babichenko's (2013) studies best-response dynamics in large aggregative games. As for the literature on mean-field games, this literature is as mentioned, very large. Parise et al. (2015b) and Grammatico et al. (2015) study control and convergence in mean-field games. Adlakha and Johari (2013) study stochastic mean-field games with strategic complementarities where agents react to the long-run average state of other players. The interested reader may consult these papers for further references.

7.2 Wireless Networks, Smart Grids, Games on Graphs, and Algorithms

There is a growing literature not just in economics but also in engineering and physics that uses aggregative games to model various problems related to network interaction between optimizing agents. Da Costa et al. (2009) study wireless spectrum sharing among so-called C-cells, which are autonomous units capable of configuring their spectrum allocation. This leads to the study of Nash equilibrium in an aggregative game of strategic substitutes (see Sections 4.2 and 3). The Nash equilibrium in Da Costa et al. (2009) is the fully distributed competitive solution to the spectrum-sharing problem. Chen et al. (2014) take an explicit aggregative games approach to the study of demand-side management in a smart grid, *i.e.*, an electricity network where consumers are able to schedule their energy consumption profiles. Consumers' objective is to minimize costs, which leads to a standard strategic form aggregative game (see Section II.A in Chen et al., 2014). Koshal, Nedik, and Shanbhag (2012) study aggregative games on graphs and so consider an explicit (abstract) network structure. The focus is on a gossip-based algorithm for the distribution across the network (for computational algorithms see also Kukushkin, 2014). Given the importance for practical (real-world) applications of such algorithmic approaches, these certainly deserve to be studied more within the setting of aggregative games. As discussed before and after Proposition 4 in Section 3, games with reciprocal interactions and generalized quasi-aggregative games more generally produce explicit potential functions for the modeling of aggregative games

[22] See Section 7.4 for a setting where agents likewise take the aggregate as given – without there being a continuum of agents.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

played on graphs/networks. Not only does this lead to existence proofs, it also points the way on how to deal with uniqueness, comparative statics, stability, and other topics. See also Parise et al. (2015a) who study networks in quasi-aggregative games and establish distributed algorithms whose implementation steer best responses to Nash equilibrium (see also Parise et al., 2015c).[23] Closer to economics and industrial organization are of course studies that look at (imperfect) competition between firms with a specific location structure. An example here is Harks and Klimm (2015) who study what they call aggregative location games where oligopolistic firms each can supply only a single market out of a set of markets. Firms must then choose quantities as well as location, which leads to some very interesting observations.

7.3 Entry and Free-entry Games

It has been known for many years that *entry* of additional players (typically firms in an IO setting) is a topic particularly amendable to aggregative games methodology. See, for example, Corchón (1994) (Proposition 1, p. 158) who presents a result on entry in games of strategic substitutes (that result is generalized by Proposition 5 in Section 4). See also Seade (1980) for an early treatment of entry. Aggregative games are particularly well suited for the study of entry (and exit) because "adding an additional agent" has a very clear and natural interpretation. One basically adds a player to the game by adding her strategy to the existing aggregate – so there is no ambiguity about how new players' strategies enter existing players' payoff function.[24] It is this feature that also makes aggregative games a particularly natural setting for evolutionary game theory, as returned to in Section 7.4. And to be sure, some of the papers described in that section (*e.g.*, Alos-Ferrer and Ania, 2005) are very closely related to IO studies of entry. This feature is also what makes a statement such as "a large aggregative game is the limit of a finite agent aggregative game as $I \to \infty$" (where I is the number of players) meaningful. See Section 7.1 above.

Now, there are settings where adding more players will not naturally lead to a large game limit with "aggregate-taking behavior" but rather, will lead to the endogenous determination of certain *number* of agents, all of whom still have a non-infinitesimal influence on the aggregate. We are speaking, of course, of monopolistic competition with *free entry*. In the most basic example, firms have start-up costs and upon entry earn (supernormal) profits. As the number of firms increase, profits decrease. At a free-entry equilibrium, the number of firms has adjusted so that firms earn zero profits, which then discourages additional entry. Okumura (2015) establishes existence of a subgame perfect equilibrium in a free-entry aggregative game and also provides an algorithm to compute the equilibrium. Anderson, Erkan, and Piccinin (2015) study aggregative oligopoly games with entry. That paper is particularly relevant for IO, providing an aggregative games "toolkit" for IO, and establishing an array of interesting results related to mergers, market structure, and surplus/welfare.

---

[23] Note that these papers allow for multi-dimensional aggregates and so are only quasi-aggregative in the sense of Section 2.3 when strategy sets are one-dimensional.

[24] In a game that is not aggregative, it is not at all clear what it means to add an additional player. In particular, one will in general have to replace all players' payoff functions with new ones with different domains if the set of agent changes.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

7.4 Evolutionary Games

Aggregative games have come to play an increasingly important role in evolutionary game theory and a number of results have been established. The reason aggregative games are well suited for studies on evolution is that one can very naturally think of adding agents, removing agents, and selecting subsets of agents once an aggregator has been fixed (see the related discussion in Section 7.3). Many evolutionary games are also either games of strategic substitutes or complements, making the results presented in the previous sections particularly relevant.

Possajennikov (2015) studies the evolutionary stability of players' *beliefs* about the aggregate in both (infinite population) random matching and finite population playing-the-field. He shows, in particular, that evolutionary stability in aggregative games dictates that agents must believe that their actions have no effect on the aggregate. This is reminiscent of price-taking/Walrasian behavior, studied in aggregative games in Alos-Ferrer and Ania (2005) and Possajennikov (2003) among others. Basically, these papers show that "aggregate-taking behavior" – already encountered in large games in Section 7.1 – is evolutionarily stable. Note that much of this literature overlaps with basic IO questions. See the previous papers for further references, as well as Schipper (2004).

7.5 Dynamics and Stability

Aggregative games that admit a best response potential (Section 3) have very nice stability properties under iterative best-response dynamics where players move in turn and to a best response. Note that such best-response improvement dynamics, is nothing but the Cournot tatonnement taught in first-year modules extended to many players and general payoff functions. Jensen, 2010 shows that if best responses are single-valued, any generalized quasi-aggregative game that satisfies the conditions of Proposition 3 have Cournot tatonnements that converge to equilibria. The method of proof in Jensen, 2010 owes everything to early contributions on this topic by Nikolai Kukushkin (2004, 2005); see also Kukushkin (2015, 2016) for more recent treatments. For different stability concepts than Cournot tatonnement see Dindosa and Mezzetti (2006) as well as Paccagnan, Kamgarpour, and Lygeros (2016).[25] While aggregative games that admit best-response potentials are particularly amenable to stability analysis, the structure of aggregative games naturally lends itself to stability analysis more generally (in which should be included algorithmic computation of Nash equilibrium). Babichenko's (2013) analysis of stability in large aggregative games was already mentioned above (Section 7.1), and several of the referred papers in that section as well as Section 7.2 specifically deal with stability in one form or another.

7.6 Contests and Public Good Provision Games

Contests as well as patent races are generalized aggregative games and so the full aggregative games machinery applies to such games. The study dates back to Loury (1979) and Tullock (1980) and the literature is both voluminous and active. For a recent survey see Corchón

[25] The latter of the two studies aggregative as well as mean-field games (Section 7.1) and consider "gradient-like" updates using tools from distributed optimization and variational inequalities.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

(2007). Since Chapter 6 by Corchón and Serena, and Chapter 7 by Hoffman and Rota-Preziosi, Volume II of this *Handbook* are devoted to contests, a few comments will suffice here.

Contests (as well as patent races) are neither games of strategic substitutes or complements. However, they satisfy the uniform local solvability condition (Definition 10) under standard conditions and so uniqueness and comparative statics results can be established at a very high level of generality using the results of Section 4.1 and 6 (see Acemoglu and Jensen, 2013, Jensen, 2016).[26] See also Szidarovszky and Okuguchi (1997) and Cornes and Hartley (2005). For some recent contributions that make active use of the aggregative games methodology to push the setting forward see Kelsey and Melkonyan (2014) who study ambiguity in contest, and Dietl et al. (2015) who investigate how promotion and relegation in sports leagues affect the strength of the divisions.

While contests are neither games of strategic substitutes nor complements, *public good provision* models are aggregative and if the public good is a normal good, also a game of strategic substitutes (Acemoglu and Jensen, 2013, Section 5.1). Just like in the case of contests, there is a large literature on public good provision. Examples that explicitly use an aggregative games approach include Cornes and Hartley (2007), Kotchen (2007), and the study of clubs in public good provision of Al-Nowaihi and Fraser (2012).

REFERENCES

Acemoglu, D. and M.K. Jensen (2010): "Robust Comparative Statics in Large Static Games", *IEEE Proceedings on Decision and Control* CDC 2010, 3133–3139.

Acemoglu, D. and M.K. Jensen (2013): "Aggregate Comparative Statics", *Games and Economic Behavior* 81, 27–49.

Acemoglu, D. and M.K. Jensen (2015): "Robust Comparative Statics in Large Dynamic Economies", *Journal of Political Economy* 123, 587–640.

Adlakha, S. and R. Johari (2013): "Mean Field Equilibrium in Dynamic Games with Strategic Complementarities", *Operations Research* 61, 971–989.

Al-Nowaihi, A. and C.D. Fraser (2012): "Does the Public Sector Over-provide Club Goods? A General Result", *Economics Letters* 117, 397–400.

Alos-Ferrer, C. and A.B. Ania (2005): "The Evolutionary Logic of Perfectly Competitive Behaviour", *Economic Theory* 26, 497–516.

Amir, R. (2005): "Ordinal versus Cardinal Complementarity: The Case of Cournot Oligopoly", *Games and Economic Behavior* 53, 1–14.

Anderson, S.P., N. Erkal, and D. Piccinin (2015): "Aggregative Oligopoly Games with Entry", Working Paper.

Babichenko, Y. (2013): "Best Reply Dynamics in Large Aggregative Games", Working Paper, accessed August 16, 2017 at SSRN: http://ssrn.com/abstract=2210080.

Bulow, J.I., J.D. Geanakoplos, and P.D. Klemperer (1985): "Multimarket Oligopoly: Strategic Substitutes and Complements", *Journal of Political Economy* 93, 488–511.

Camacho, C., T. Kamihigashi, and C. Saglam (2016): "Robust Comparative Statics for Non-monotone Shocks in Large Aggregative Games", Working Paper.

Chen, H., Y. Li, R.H.Y. Louie, and B. Vucetic (2014): "Autonomous Demand Side Management Based on Energy Consumption Scheduling and Instantaneous Load Billing: An Aggregative Game Approach", *IEEE Transactions on Smart Grid* 5, 1744–1754.

Corchón, L. (1994): "Comparative Statics for Aggregative Games. The Strong Concavity Case", *Mathematical Social Sciences* 28, 151–165.

Corchón, L. (2007): "The Theory of Contests: A Survey", *Review of Economic Design* 11, 69–100.

Cornes, R. and R. Hartley (2005): "Asymmetric Contests with General Technologies", *Economic Theory* 26, 923–946.

[26] Strictly speaking contests are not nice games (Definition 8) since payoffs are not continuous at the boundary. There is an easy fix to this problem, however, which basically reinstates the entire theory word for word (see Jensen, 2016).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Cornes, R. and R. Hartley (2007): "Weak Links, Good Shots and Other Public Good Games: Building on BBV", *Journal of Public Economics* 91, 1684–1707.
Cornes, R. and R. Hartley (2012): "Fully Aggregative Games", *Economics Letters* 116, 631–633.
Da Costa, G.W.O., A.F. Cattoni, I.Z. Kovacs, and P.E. Mogensen (2009): "A Scalable Spectrum-sharing Mechanism for Local Area Network Deployment", *IEEE Transactions on Vehicular Technology* 59, 1630–1645.
Dietl, H., M. Grossmann, A. Hefti and M. Lang (2015): "Spillovers in Sports Leagues with Promotion and Relegation", *Scottish Journal of Political Economy* 62, 59–74.
Dindosa, M. and C. Mezzetti (2006): "Better-reply Dynamics and Global Convergence to Nash Equilibrium in Aggregative Games", *Games and Economic Behavior* 54, 261–292.
Dubey, P., O. Haimanko, and A. Zapechelnyuk (2006): "Strategic Complements and Substitutes, and Potential Games", *Games and Economic Behavior* 54, 77–94.
Gorman, W.M. (1968): "The Structure of Utility Functions", *Review of Economic Studies* 35, 367–390.
Grammatico, S., F. Parise, M. Colombino, and J. Lygeros (2015): "Decentralized Convergence to Nash Equilibria in Constrained Deterministic Mean Field Control", *IEEE Transactions on Automatic Control* 99, 1–1.
Hahn, F. (1962): "The Stability of the Cournot Oligopoly Solution Concept", *Review of Economic Studies* 29, 329–331.
Harks, T. and M. Klimm (2015): "Equilibria in a Class of Aggregative Location Games", *Journal of Mathematical Economics* 61, 211–220.
Huang, M., P.E. Caines, and R.P. Malhame (2007): "Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ϵ-Nash equilibria", *IEEE Transactions on Automatic Control* 52, 1560–1571.
Jensen, M.K. (2010): "Aggregative Games and Best-reply Potentials", *Economic Theory* 43, 45–66.
Jensen, M.K. (2015): "Distributional Comparative Statics", *University of Leicester Discussion Papers in Economics* 15/14.
Jensen, M.K. (2016): "Existence, Uniqueness, and Comparative Statics in Contests", in P. von Mouche and F. Quartieri, eds, *Contributions on Equilibrium Theory for Cournot and Related Games: Essays in Honour of Koji Okuguchi*, Berlin: Springer Verlag.
Kelsey, D. and T. Melkonyan (2014): "Contests with Ambiguity", Working paper, accessed August 16, 2017 at SSRN: http://ssrn.com/abstract=2458196.
Koshal, J., A. Nedic, and U.V. Shanbhag (2012): "A Gossip Algorithm for Aggregative Games on Graphs", *Decision and Control (CDC) 2012 IEEE 51st Annual Conference*, 4840–4845.
Kotchen, M.J. (2007): "Equilibrium Existence and Uniqueness in Impure Public Good Models", *Economics Letters* 97, 91–96.
Kukushkin, N.S. (1994): "A Fixed-point Theorem for Decreasing Mappings", *Economics Letters* 46, 23–26.
Kukushkin N.S. (2004): "Best response dynamics in finite games with additive aggregation", *Games and Economic Behavior* 48, 94–110.
Kukushkin, N.S. (2005): "'Strategic supplements' in games with polylinear interactions", *Econ WPA Paper* 0411008, accessed August 14, 2017 at http://ideas.repec.org/p/wpa/wuwpga/0411008.html
Kukushkin, N.S. (2014): "Novshek's Trick as a Polynomial Algorithm", *Discrete Applied Mathematics* 162, 28–31.
Kukushkin, N.S. (2015): "Cournot Tatonnement and Potentials", *Journal of Mathematical Economics* 59, 117–127.
Kukushkin, N.S. (2016): "Cournot Tatonnement in Aggregative Games with Monotone Best Responses", in P. von Mouche and F. Quartieri, eds, *Contributions on Equilibrium Theory for Cournot and Related Games: Essays in Honour of Koji Okuguchi* Berlin: Springer Verlag.
Lasry, J.-M. and P.-L. Lions (2007): "Mean Field Games", *Japanese Journal of Mathematics* 2, 229–260.
Loury, G.C. (1979): "Market Structure and Innovation", *Quarterly Journal of Economics* 93, 395–410.
Mangasarian, O.L. (1965): "Pseudo-convex Functions", *SIAM Journal of Control* 3, 281–290.
Milgrom, P. and J. Roberts (1990): "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities", *Econometrica* 58, 1255–1277.
Milgrom, P. and J. Roberts (1994): "Comparing Equilibria", *American Economic Review* 84, 441–459.
Milgrom, P. and C. Shannon (1994): "Monotone Comparative Statics", *Econometrica* 62, 157–180.
Novshek, W. (1985): "On the Existence of Cournot Equilibrium", *Review of Economic Studies* 52, 85–98.
Okumura, Y. (2015): "Existence of Free Entry Equilibrium in Aggregative Games with Asymmetric Agents", *Economics Letters* 127, 14–16.
Paccagnan, D., M. Kamgarpour, and J. Lygeros (2016): "On Aggregative and Mean Field Games with Applications to Electricity Markets", *Proceedings of the ECC* (June).
Parise, F., B. Gentile, S. Grammatico, and J. Lygeros (2015a): "Network Aggregative Games: Distributed Convergence to Nash Equilibria", *IEEE Proceedings on Decision and Control CDC 2015*, 2295–2300.
Parise, F., S. Grammatico, M. Colombino, and J. Lygeros (2015b): "On Constrained Mean Field Control for Large Populations of Heterogeneous Agents: Decentralized Convergence to Nash Equilibria", *Control Conference* (*ECC 2015, European*), 3316–3321.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Parise, F., S. Grammatico, B. Gentile, and J. Lygeros (2015c): "Network Aggregative Games and Distributed Mean Field Control via Consensus Theory", Working Paper., accessed August 16, 2017 at http://arxiv.org/pdf/1506.07719.pdf,2015.

Phlips, L. (1995): *Competition Policy: A Game-theoretic Perspective*, New York: Cambridge University Press.

Possajennikov, A. (2003): "Evolutionary Foundation of Aggregative-taking Behavior", *Economic Theory* 21, 921–928.

Possajennikov, A. (2015): "Economic Conjectural Variations in Aggregative Games: An Evolutionary Perspective", *Mathematical Social Science* 77, 55–61.

Quah, J.K.-H. (2007): "The Comparative Statics of Constrained Optimization Problems", *Econometrica* 75, 401–431.

Roy, S. and T. Sabarwal (2010): "Monotone Comparative Statics for Games with Strategic Substitutes", *Journal of Mathematical Economics* 46, 793–806.

Roy, S. and T. Sabarwal (2012): "Characterizing Stability Properties in Games with Strategic Substitutes", *Games and Economic Behavior* 75, 337–353.

Schipper, B. (2004): "Submodularity and the Evolution of Walrasian Behavior", *International Journal of Game Theory* 32, 471–477.

Seade, J. (1980): "On the Effects of Entry", *Econometrica* 48, 256–278.

Selten, R. (1970): *Preispolitik der Mehrproduktenunternehmung in der Statischen Theorie*, Berlin: Springer Verlag.

Shirai, K. (2010): "An Existence Theorem for Cournot-Walras Equilibria in a Monopolistically Competitive Economy", *Journal of Mathematical Economics* 46, 1093–1102.

Szidarovszky, F. and K. Okuguchi (1997): "On the Existence and Uniqueness of Pure Nash Equilibrium in Rent-seeking Games", *Games and Economic Behavior* 18, 135–140.

Tarski, A. (1955): "A Lattice-theoretical Fixpoint Theorem and its Applications", *Pacific Journal of Mathematics* 5, 285–309.

Topkis, D.M. (1978): "Minimizing a Submodular Function on a Lattice", *Operations Research*, 26, 305–321.

Topkis, D.M. (1998): *Supermodularity and Complementarity*, Princeton, NJ: Princeton University Press.

Tullock, G. (1980): "Efficient Rent-seeking" in J.M. Buchanan, R.D. Tollison, and G. Tullock, eds, *Toward a Theory of the Rent-Seeking Society*, College Station, TX: Texas A&M University Press.

Vind, K. and B. Grodal (2003): *Independence, Additivity, Uncertainty (Studies in Economic Theory, Vol. 14)*, Berlin: Springer Verlag.

Vives, X. (1990): "Nash Equilibrium with Strategic Complementarities", *Journal of Mathematical Economics* 19, 305–321.

Vives, X. (2000): *Oligopoly Pricing: Old Ideas and New Tools*, Cambridge, MA: MIT Press.

Voorneveld, M. (2000): "Best-response Potential Games", *Economics Letters* 66, 289–295.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

5. Monopolistic competition without apology

Jacques-François Thisse and Philip Ushchev *

1 INTRODUCTION

The absence of a general equilibrium model of oligopolistic competition unintentionally paved the way for the success of the constant elasticity of substitution (CES) model of monopolistic competition. This model, developed by Dixit and Stiglitz (1977), has been used in so many economic fields that a large number of scholars view it as virtually *the* model of monopolistic competition. The main thrust of this chapter is that monopolistic competition is a market structure in its own right, which encompasses a much broader set-up than what most economists believe it to be.

According to Chamberlin (1933), monopolistic competition is defined as a market environment in which a firm has no impact on its competitors (as in perfect competition) but is free to choose the output (or price) that maximizes its profits (as a monopolist). In other words, although one firm is negligible to the market, it is endowed with market power because it sells a differentiated product. For this to hold true, each firm must compete against the market as a whole or, to use Triffin's (1940) formulation, the cross-elasticity between any two varieties has to be negligible. According to the "folk theorem of competitive markets," perfect competition almost holds when firms are small relative to the size of the market. Hence, for a long time, economists debated heatedly whether Chamberlin's assumptions make sense. We will make no attempt to summarize this debate. Nevertheless, a few contributors raised fundamental questions that will be discussed later on.

We choose to focus on the two main approaches that have been developed to study monopolistic competition and explore the conditions under which these approaches lead to similar results. In the first, we consider an oligopolistic game in which firms compete in quantity (Cournot) or price (Bertrand). We then ask whether the sequence of Nash equilibria of these games converges to a competitive outcome when the number of firms grows indefinitely. If not, *monopolistic competition may be viewed as approximating a market in which strategic interactions among firms are weak.*

The second approach builds on Aumann (1964) who shows that the distribution of agents must be non-atomic for each agent to be negligible to the market. The same idea is applied to firms to account for Chamberlin's idea that a firm's action has no impact on its competitors. In other words, the supply side of the market is described by a continuum of firms whose mass is pinned down by the zero-profit condition. The next step is to check whether the Nash equilibria of these non-atomic games are identical to the competitive equilibria. When the

* We are especially grateful to a referee, and readers C. Arkolakis and S. Redding for an impressive list of stimulating comments and suggestions. We thank M. Parenti for long and fruitful discussions on the subject matter. We also thank S. Anderson, D. Connolly, L. Corchón, F. Di Comite, P. Fajgelbaum, A. Osharin, F. Robert-Nicoud, A. Tarasov, H. Vandenbussche, and X. Vives for insightful discussions and remarks. The study has been funded by the Russian Academic Excellence Project '5–100'.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

answer is negative, monopolistic competition may be considered as a market structure per se. To put it differently, *monopolistic competition is the equilibrium outcome of a non-atomic game with an endogenous mass of players*.

Modeling monopolistic competition as a non-atomic game yields a framework easier to handle than standard oligopoly models while coping with general equilibrium effects, a task that is hard to accomplish in oligopoly theory (Hart, 1985a). Furthermore, even though firms do not compete strategically, *general models of monopolistic competition are able to mimic oligopolistic markets with free entry within a general equilibrium framework*. This is in accordance with Mas-Colell (1984, p. 19) for whom "the theory of monopolistic competition derives its theoretical importance not from being a realistic complication of the theory of perfect competition, but from being a simplified, tractable limit of oligopoly theory."

How product differentiation and consumer preferences are modeled has far-reaching implications for what is meant by monopolistic competition. In an influential review of Chamberlin's book, Kaldor (1935) objects to the idea that each firm is able to compete directly with all the others. According to Kaldor, firms are rooted in specific places. As a consequence, they have competitors that are close while the others are remote. Regardless of the number of firms in total, the number of firms competing for any particular consumer is small, so a decision made by one firm has a sizable impact only on the neighboring firms. Under these circumstances, monopolistic competition would not make sense. This argument is similar to the ideas developed by Hotelling (1929) and later on by Beckmann (1972) and Salop (1979). For these authors, *competition is localized*, meaning that a firm faces a limited number of direct competitors that operate in its vicinity. Building on Hotelling's intuition, Lancaster (1979) puts forward the same idea in the context of a characteristics space where products are positioned while consumers have their own ideal varieties forming a constellation of points that belong to the very same space. These various strands of literature have given rise to a model of spatial competition with free entry. This model remains in the tradition of oligopoly theory: firm behavior is strategic because competition is localized while its global impact is diffused among firms through chain effects that link any two firms belonging to the same industry.

In contrast, if consumers have a love of variety, Kaldor's criticism ceases to be relevant. In this context, firms all compete together as they all strive to attract the entire population of consumers. This is why Chamberlin's model of monopolistic competition is henceforth associated with consumers who aim to consume many varieties, rather than consuming their ideal variety. After having attracted a great deal of attention in the 1930s, Chamberlin's ideas languished until Spence (1976) and, above all, Dixit and Stiglitz (1977) brought them back onto the scientific stage by proposing a model capable of being used in various economic fields. Spence developed a partial equilibrium setting, whereas the Dixit-Stiglitz model places itself in a general equilibrium context. Both modeling strategies are used in the literature. The former is more popular in industrial organization, whereas the latter is the workhorse of new trade and growth theories. This justifies our choice not to take a stance on choosing one particular strategy, but rather to deal with both.

On the production side, Chamberlin remained in the Marshallian tradition by assuming that firms face the U-shaped average cost curves. Since firms face downward-sloping demand schedules and profits vanish under free entry, each firm produces at the tangency point of the demand and average cost curves. As a result, the equilibrium output level is smaller than the one that minimizes its unit costs, a claim dubbed the "excess capacity theorem." Under

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

the severe conditions of the Great Depression, this was viewed as evidence that competition may generate a waste of resources. However, this argument overlooks the fact that, when consumers value product differentiation, a wider product range generates welfare gains that must be taken into account when assessing the (in)efficiency of monopolistic competition. Under these circumstances, there is a trade-off between scale economies associated with the production of varieties and the range of available varieties. This suggests the following question: *does the market over- or under-provide variety?*

This chapter reflects those various lines of research. However, its main emphasis will be on the models whose origin lies in the pioneering work of Dixit and Stiglitz (1977). Modeling monopolistic competition as a non-atomic game makes the corresponding market structure different from those studied in industrial organization. The upshot of the matter is that monopolistic competition encapsulates increasing returns and imperfect competition in a general equilibrium setting. Such a combination leads to a wide range of findings that may differ greatly from those obtained in a general competitive analysis, while permitting the study of issues that are hard to tackle within an oligopoly framework (Matsuyama, 1995).

The remainder of the chapter is organized as follows. First of all, there is a lot to learn from early contributions that are often disregarded in the modern literature. For this reason, Section 2 is devoted to those contributions, but we make no attempt to provide a detailed survey of what has been accomplished. In doing this, we follow the tradition of oligopoly theory and focus on partial equilibrium. Section 3 highlights the role of the negligibility hypothesis in the CES and linear-quadratic (LQ) models. Being negligible to the market, each firm treats parametrically market aggregates, which relaxes substantially the technical difficulties of working with imperfect competition in general equilibrium. In Section 4, we discuss a general set-up under the negligibility hypothesis and the "heroic assumption" that both demand and cost curves are symmetric (Chamberlin, 1933, p. 82). The focus is now on a *variable elasticity of substitution* (VES), which depends upon the individual consumption and mass of varieties. Under these circumstances, the VES model encompasses the whole family of models with symmetric preferences. Furthermore, the VES model of monopolistic competition is able to mimic key results of oligopoly theory. To a certain extent, we therefore find the dichotomy between oligopolistic and monopolistic competition unwarranted.

In Section 5, we make less heroic assumptions by recognizing that firms are heterogeneous. The literature on heterogeneous firms is huge and therefore we are content to provide an overview of the main findings (Redding, 2011). In the spirit of the preceding sections, we depart from the CES, which has taken center stage ever since Melitz's (2003) pioneering contribution. Section 6 is devoted to the classical question: does the market provide too many or too few varieties? As anticipated by Spence (1976), the numerous effects at work leave little hope of coming up with robust results, the reason being that the answer depends on the demand-side properties. As a consequence, there is no need to discuss this question at length. Note, however, that the variety of welfare results casts some doubt on prescriptions derived from quantitative models that use CES preferences. Section 7 concludes and proposes a short research agenda.

A final comment is in order. This chapter is about the *theory* of monopolistic competition. This does not reflect any prejudice on our part, but dealing with econometric and applied issues would take us way beyond the scope of this chapter. We refer to De Loecker and Goldberg (2014) for a detailed survey of this literature.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

2 MONOPOLISTIC COMPETITION AS THE LIMIT OF OLIGOPOLISTIC COMPETITION

There are (at least) three ways to model preferences for differentiated products. In the first, consumers are endowed with a utility $U(\mathbf{x})$ defined on the set X of *potential* varieties, which is continuous and strictly quasi-concave in $\mathbf{x}$ (see, e.g., Vives, 1999). It is well known that the convexity of preferences describes *variety-seeking* behavior. When preferences are symmetric, the convexity of preferences implies that a consumer has a love for variety, that is, she strictly prefers to consume the whole range of available varieties than any subset.

In the second approach, every consumer has one *ideal variety* and different consumers have different ideal varieties. In the spatial metaphor proposed by Hotelling (1929), a consumer's ideal variety is represented by her location in some geographical space (Main Street), while the variety provided by a firm is the location of this firm in the same space. Formally, the set X of varieties is defined by a metric space, such as a compact interval or a circle. Using a metric space allows one to measure the "distance" between any two locations, while the utility loss incurred by a consumer for not consuming her ideal variety is interpreted as the transport cost this consumer must bear to visit a firm, which increases with distance. Regardless of the number of available varieties, a consumer purchases a single variety. In this event, preferences are no longer convex, making it problematic to prove the existence of an equilibrium. However, ever since Hotelling (1929), it is well known that this difficulty may be obviated when there is a large number (formally, a continuum) of heterogeneous consumers.

A third approach was developed to account for taste heterogeneity, as in spatial models, but in a set-up that shares some basic features of symmetric models. This is achieved using the random utility model developed in psychology and applied to econometrics by McFadden (1974). Interestingly, this approach looks at first sight like the second approach, but is isomorphic at the aggregate level to the first approach. Although discrete choice models have not been developed to study monopolistic competition per se, the results obtained under oligopoly can be used to study the market outcome when the number of firms is arbitrarily large.

The literature is diverse and, therefore, difficult to integrate within a single framework. In addition, some papers are technically difficult. In what follows, we use simple models to discuss under which conditions each of these three approaches leads to perfect or monopolistic competition when the number of firms grows indefinitely.

2.1 Variety-seeking Consumers

2.1.1 Additive aggregate

There are two goods, a differentiated good and a homogeneous good. The homogeneous good x_0 is unproduced and used as the numéraire. The differentiated good is made available under the form of a finite number $n \geq 2$ of varieties, which are strong gross substitutes. Throughout this chapter, unless stated otherwise, each variety is produced by a single firm because firms seek to avoid the negative consequences of face-to-face competition, while each firm produces a single variety because there are no scope economies. Producing x_i units of variety $i = 1, \ldots, n$ requires cx_i units of the numéraire where the marginal cost $c > 0$ is constant.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

There is a unit mass of identical consumers or, equivalently, a representative consumer who are each endowed with one unit of the numéraire. Like in mainstream oligopoly theory, consumers have quasi-linear preferences given by

$$U(\mathbf{x}) = \varphi\left(X(\mathbf{x})\right) + x_0, \tag{5.1}$$

where φ is twice continuously differentiable, strictly increasing, strictly concave over $\mathbb{R}_+$, and such that $\varphi(0) = 0$, while the sub-utility $X(\mathbf{x})$ maps the consumption profile $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n_+$ into $\mathbb{R}_+$. The utility φ measures the desirability of the differentiated good relative to the numéraire. The concavity of $\varphi(\cdot)$ implies that the marginal utility of X decreases, and thus the marginal rate of substitution between X and x_0 decreases with X.

The sub-utility $X(\mathbf{x})$ is supposed to be symmetric and additive:

$$X(\mathbf{x}) \equiv \sum_{i=1}^{n} u(x_i), \tag{5.2}$$

where u is thrice continuously differentiable, strictly increasing, strictly concave over $\mathbb{R}_+$, and $u(0) = 0$. The concavity of $u(\cdot)$ amounts to assuming that consumers are variety-seekers: rather than concentrating their consumption on a small mass of varieties, they prefer to spread it over the whole range of available varieties. As a consequence, the elasticity of the sub-utility with respect to the per variety consumption level does not exceed one: $\mathcal{E}_{x_i}(u) \equiv x_i u'(x_i)/u(x_i) \leq 1$. The behavior of this elasticity plays a major role in shaping the welfare properties of monopolistic competition (see Dhingra and Morrow, 2018, and Section 6 of this chapter). Furthermore, it should be clear that the symmetry of lower-tier utility (14.27) means that the utility level is unaffected if varieties are renumbered.

Following Zhelobodko et al. (2012), we define the *relative love for variety* (RLV) as follows:

$$r_u(x) \equiv -\frac{xu''(x)}{u'(x)},$$

which is strictly positive for all $x > 0$. Very much like the Arrow-Pratt relative risk aversion, the RLV is a local measure of love for variety. Consumers do not care about variety when $u(x_i) = x_i$, which means $r_u(x) \equiv 0$ for all $x > 0$. As the value of $r_u(x)$ grows, the consumer has a stronger love for variety. Therefore, how the RLV changes with the per variety consumption is crucial for the analysis of the equilibrium. Under the CES, we have $u(x) = x^{(\sigma-1)/\sigma}$ where σ, the elasticity of substitution between any two varieties, is a constant larger than 1; the RLV is given by $1/\sigma$. Other examples of additive preferences include the constant absolute risk aversion (CARA) (Behrens and Murata, 2007) and the addilog (Simonovska, 2015).

Let $\mathbf{p} = (p_1, \ldots, p_n)$ be a price vector. Utility maximization yields the inverse demand for variety i:

$$p_i\left(x_i, \mathbf{x}_{-i}\right) = \varphi'(X(\mathbf{x})) \cdot u'(x_i). \tag{5.3}$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

The Marshallian demands $x_i(\mathbf{p})$ are obtained by solving the following system of equations:

$$p_i = \varphi'(X(\mathbf{x})) \cdot u'(x_i), \qquad i = 1, \ldots, n. \tag{5.4}$$

Combining (14.27) and (5.4) yields the Marshallian demand for variety i:

$$x_i(p_i, \mathbf{p}_{-i}) = \xi\left(\frac{p_i}{P(\mathbf{p})}\right), \tag{5.5}$$

where $\xi(\cdot) \equiv (u')^{-1}(\cdot)$, while $P(\mathbf{p})$ is the unique solution to the equation:

$$P = \varphi'\left[\sum_{j=1}^{n} u\left(\xi\left(\frac{p_j}{P}\right)\right)\right]. \tag{5.6}$$

Clearly, a price cut by firm i draws demand equally from all the other firms, which reflects the symmetry of preferences, while P plays the role of the Lagrange multiplier when the budget constraint is binding.

Bertrand competition We consider a non-cooperative game in which the players are firms. The strategy of firm i is given by its price p_i and its payoff by its profits given by

$$\Pi_i^B(\mathbf{p}) = (p_i - c)x_i(\mathbf{p}) = (p_i - c)\xi\left(\frac{p_i}{P(\mathbf{p})}\right), \quad i = 1, \ldots, n. \tag{5.7}$$

A Nash equilibrium $\mathbf{p}^* = (p_1^*, \ldots, p_n^*)$ of this game is called a *Bertrand equilibrium*, which is symmetric if $p_i^* = p^B(n)$ for all $i = 1, \ldots, n$. It follows from (5.7) that $\Pi_i^B(\mathbf{p})$ is a function of p_i and $P(\mathbf{p})$ only. Therefore, the Bertrand game under additive preferences is an *aggregative game* in which $P(\mathbf{p})$ is the market statistic.

In the remainder of this chapter, we denote by $\mathcal{E}_z(f)$ the elasticity of a function $f(z)$ with respect to z. Differentiating (5.5) with respect to p_i and using (5.4) yields the price elasticity of the demand for variety i:

$$\mathcal{E}_{p_i}(x_i) = \frac{1 - \mathcal{E}_{p_i}(P)}{r_u\left[\xi(p_i/P)\right]}. \tag{5.8}$$

Since firm i's profit-maximizing markup is given by $m_i^B = 1/\mathcal{E}_{p_i}(x_i)$, (5.8) implies that *firm i's Bertrand-markup* may be written as follows:

$$m_i^B = \frac{1}{1 - \mathcal{E}_{p_i}(P)} \cdot r_u\left[\xi\left(\frac{p_i}{P}\right)\right], \tag{5.9}$$

where $0 < \mathcal{E}_{p_i}(P) < 1$ is shown to hold in Appendix.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Assume that firms treat P parametrically, so that $\mathcal{E}_{p_i}(P) = 0$. In this case, (5.9) boils down to $m_i^B = r_u(x_i) > 0$. Hence, even when firms are not aware that they can manipulate P, they price above marginal cost because their varieties are differentiated. When firms understand that they can manipulate P ($\mathcal{E}_{p_i}(P) > 0$), we have $1/\left[1 - \mathcal{E}_{p_i}(P)\right] > 1$. This new effect stems from the strategic interactions among firms through the market statistic P, which allows them to hold back their sales and to raise their profit. In summary, (5.9) highlights the existence of *two* sources of market power: *monopoly* power ($r_u(x) > 0$) and *strategic* power ($\mathcal{E}_{p_i}(P) > 0$).

We now show that the strategic power of firms vanishes as the number of firms unboundedly grows. Consider a symmetric Bertrand equilibrium $p_i^* = p^B(n)$ and find the equilibrium consumption $x^B(n)$, which is the unique solution to:

$$\varphi'\left[nu(x)\right] \cdot u'(x) = p^B(n).$$

The expression (A.1) in the Appendix implies that

$$\lim_{n \to \infty} \mathcal{E}_{p_i}(P)\Big|_{p_i = p^B(n)} = 0,$$

which means that *strategic interactions vanish at the limit.*

How does the monopoly term $r_u\left[x^B(n)\right]$ behave when n grows unboundedly? To check this, note that the budget constraint, together with $p^B(n) \geq c$, implies that $x^B(n) \leq 1/cn$. Therefore, when n tends to infinity, $x^B(n)$ converges to zero. Combining this with (5.9), we obtain:

$$\lim_{n \to \infty} m^B(N) = \lim_{n \to \infty} \frac{1}{1 - \mathcal{E}_{p_i}(P)\Big|_{p_i = p^B(n)}} \cdot \lim_{n \to \infty} r_u\left[x^B(n)\right] = r_u(0).$$

Cournot competition Firm i's profit function is now given by

$$\Pi_i^C(\mathbf{x}) = \left[p_i(x_i, \mathbf{x}_{-i}) - c\right] x_i = \left[\varphi'(X(\mathbf{x})) \cdot u'(x_i) - c\right] x_i,$$

A *Cournot equilibrium* is a vector $\mathbf{x}^* = (x_1^*, \ldots, x_n^*)$ such that each strategy x_i^* is firm i's best reply to the strategies $\mathbf{x}_{-i}^*$ chosen by the other firms. This equilibrium is *symmetric* if $x_i^* = x^C(n)$ for all $i = 1, \ldots, n$. Using (5.3), we may restate *firm i's Cournot-markup* as follows:

$$m_i^C = r_u(x_i) + r_\varphi(X)\mathcal{E}_{x_i}(X), \tag{5.10}$$

where X is the market statistic (14.27).

As in the Bertrand game, there are two sources of market power, that is, strategic power and monopoly power. Under Cournot, the decomposition is additive, whereas the decomposition is multiplicative under Bertrand (see (5.9)). Despite this difference, both (5.9) and (5.10) show the importance of product differentiation for consumers through the value of the RLV.

Assume for simplicity that $r_\varphi(X)$ is bounded from above by a positive constant K; this property holds for the logarithmic and power functions. Since

$$\mathcal{E}_{x_i}(X) = \mathcal{E}_{x_i}(u) \cdot \frac{u(x_i)}{X},$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

and $0 < \mathcal{E}_{x_i}(u) < 1$ for all $x_i > 0$, it must be that

$$\lim_{n\to\infty} r_\varphi(X)\mathcal{E}_{x_i}(X)\big|_{x_i=x^C(n)} \leq \lim_{n\to\infty} \frac{K}{n} = 0.$$

Therefore, as in the Bertrand game, strategic power is diluted in an ocean of small firms selling differentiated varieties. As for the monopoly term in (5.10), the argument developed in the Bertrand case applies. It then follows from (5.10) that

$$\lim_{n\to\infty} m^C(n) = \lim_{n\to\infty} r_u\left[x^C(n)\right] + \lim_{n\to\infty} r_\varphi(X)\mathcal{E}_{x_i}(X)\big|_{x_i=x^C(n)} = r_u(0).$$

The limit of Bertrand and Cournot competition The following proposition comprises a summary:

Proposition 1 *If there is $n_0 \geq 2$ such that a symmetric equilibrium exists under Cournot and Bertrand for all $n > n_0$, then*

$$\lim_{n\to\infty} m^B(n) = \lim_{n\to\infty} m^C(n) = r_u(0).$$

Since the strategic terms $\mathcal{E}_{p_i}(P)$ and $\mathcal{E}_{x_i}(X)$ converge to 0 when n goes to infinity, whether the limit of Bertrand and Cournot competition is perfectly competitive or monopolistically competitive is the same under both regimes and hinges on the value of $r_u(0)$. When $r_u(0) > 0$, a very large number of firms whose size is small relative to the market is consistent with the idea that firms retain enough market power to have a positive markup. To be precise, even when individuals face a very large number of varieties and consume very little of each variety, they still value diversity. It then follows from (5.8), that the price elasticity of a firm's demand is finite, which allows firms to retain monopoly power and to sustain a positive markup. On the other hand, when $r_u(0) = 0$, a growing number of firms always leads to a perfectly competitive outcome. Since both sources of market power vanish at the limit, the price elasticity of a firm's demand is infinite. Intuitively, consumers no longer care about diversity because their per variety consumption is too low. In brief, *the love for variety must be sufficiently strong for monopolistic competition to emerge*.

2.1.2 Linear-quadratic preferences

In the case of two varieties, the LQ utility is given by:

$$U(x_1, x_2) = \alpha(x_1 + x_2) - \frac{\beta}{2}(x_1^2 + x_2^2) - \gamma x_1 x_2 + x_0, \tag{5.11}$$

where α, β and γ are three positive constants such that $\gamma < \beta$. In the case of $n > 2$ varieties, there are at least two different specifications of the LQ utility, which each reduces to (5.11) when $n = 2$:

$$U(\mathbf{x}) = \alpha \sum_{i=1}^{n} x_i - \frac{\beta}{2}\sum_{i=1}^{n} x_i^2 - \gamma \sum_{i=1}^{n}\sum_{j\neq i}^{n} x_i x_j + x_0, \tag{5.12}$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

or

$$U(\mathbf{x}) = \alpha \sum_{i=1}^{n} x_i - \frac{\beta}{2} \sum_{i=1}^{n} x_i^2 - \frac{\gamma}{n-1} \sum_{i=1}^{n} \sum_{j \neq i}^{n} x_i x_j + x_0, \tag{5.13}$$

where $\Sigma_i x_i^2$ is the Herfindahl-Hirschman index measuring the dispersion of the consumption profile $\mathbf{x}$, so that β measures the intensity of love for variety; α is the willingness-to-pay for the differentiated product, while γ is an inverse measure of the degree of differentiation across varieties.

It is readily verified that under (5.12) the equilibrium markup tends to 0 when n goes to infinity, whereas the equilibrium markup is constant and positive under (5.13). In other words, (5.12) leads to perfect competition and (5.13) to monopolistic competition.

To sum up, *whether the limit of oligopolistic competition is monopolistic or perfect competition hinges on preferences* (see also Vives, 1985, Proposition 3). For example, under CARA or the LQ (5.12), we have $r_u(0) = 0$, and thus the limit of oligopolistic competition is perfect competition. In contrast, under the CES, we have $r_u(0) = 1/\sigma > 0$; the limit of the CES oligopoly model may thus be viewed as a "true" model of monopolistic competition.

2.2 Heterogeneous Consumers: The Spatial Approach

In his review of Chamberlin's book, Kaldor (1935) argues forcefully that product locations in characteristics space, or firms' locations in the geographical space, mold market competition in a very specific way: whatever the total number of firms in the industry, each one competes more vigorously with its immediate neighbors than with more distant firms. Or, in the words of Kaldor (1935, p. 390): "the different producers' products will never possess the same degree of substitutability in relation to any particular product. Any particular producer will always be faced with rivals who are nearer to him, and others who are farther off. In fact, he should be able to class his rivals, from his own point of view, in a certain order, according to the influence of their prices upon his own demand."

To develop the idea that some firms are close whereas others are distant, Kaldor used Hotelling's (1929) spatial metaphor. In spatial models of product differentiation, a consumer is identified by her "ideal" variety $s \in S \subset \mathbb{R}^n$, while the variety provided by firm i is denoted by $s_i \in S$. Hotelling (1929) uses the following spatial metaphor: firms and consumers are located in a metric space $S \subset \mathbb{R}^n$ where $d(s, s')$ is the physical distance between any two locations s and $s' \in S$. Because moving from one place to another involves a cost, space is sufficient to render heterogeneous consumers who are otherwise homogeneous. In a characteristics space à la Gorman-Lancaster, $d(s, s')$ is the inverse measure of the degree of substitutability between the varieties s and s'. Besides the distance d, the other salient feature of the spatial model of product differentiation is the *transport rate* $t > 0$, which can be viewed as the intensity of preferences for the ideal variety, or the unit cost of travelling to a store. The taste mismatch of consumer s with variety s_i is now expressed by the weighted distance $td(s, s_i)$ between the consumer's ideal variety and firm i's variety. Even though the individual purchase decision is discontinuous – a consumer buys from a single firm – Hotelling (1929) finds it reasonable to suppose that firms' aggregate demands are continuous. Supposing that consumers are continuously distributed across locations solves the apparent contradiction between discontinuity at the individual level and continuity at the aggregate level.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Spatial models of product differentiation attracted a lot of attention in 1970s and 1980s. In this set-up, a consumer can purchase from any firm, provided she is willing to pay the transport cost, and thus the boundaries between firms are endogenous to firms' prices and locations. One of the earliest contributions is Beckmann (1972) who studied how firms equidistantly distributed over S compete to attract consumers who are uniformly distributed over the same space with a unit density. Each consumer buys one unit of the good up to a given reservation price, while transport costs are linear in the Euclidean distance. Accordingly, competition is *localized*, whereas it is *global* in models with symmetric preferences such as those discussed above. In the geographical space, the goods sold by any two stores are physically identical but differentiated by the places where they are made available. As a consequence, a consumer buys from the firm with the lowest full price, which is defined as the posted (mill) price plus the transport cost to the corresponding firm.

Assume that S is given by a one-dimensional market without boundary, e.g., the real line or a circle. In this case, firm i has only two neighbors located at a distance Δ on either side of s_i. When t takes on a high value, firm i is a local monopoly because it is too expensive for consumers located near the midpoint between firms $i-1$ and i to make any purchase. On the contrary, when t is sufficiently low, each firm competes with its two neighbors for the consumers located between them. As argued by Kaldor in the above quotation, the market power of a firm is restrained by the actions of neighboring firms. In other words, their (geographic) isolation avails them only local monopoly power, for firm i's demand depends upon the prices set by the neighboring firms $i-1$ and $i+1$:

$$x_i(p_{i-1}, p_i, p_{+1}) = \max\left\{0, \frac{p_{i-1} - 2p_i + p_{i+1} + 2t\Delta}{2t}\right\}. \tag{5.14}$$

If n firms are symmetrically located along a circle C of length L ($\Delta = L/n$), the equilibrium price is given by

$$p^*(n) = c + t\Delta = c + \frac{tL}{n}. \tag{5.15}$$

Hence, $p^*(n)$ decreases with the transport rate t because firms benefit less from their geographical separation. At the limit, when $t = 0$, distance does not matter anymore, implying that firms price at marginal cost. Thus, *the limit of the spatial model of monopolistic competition is perfect competition*. When fixed costs are taken into account, the free-entry equilibrium price can be shown to decrease when the market size is expanded by raising the consumer density over S.

Beckmann's paper went unnoticed outside the field of regional science. It is also worth mentioning the contributions made by Eaton and Lipsey (1977) who build a theory of market competition in a spatial economy. Again, despite the quality of the work, Eaton and Lipsey's contributions attracted a limited amount of attention. It was not until Salop (1979), who used the circular city model, that scholars in industrial organization started paying attention to spatial competition models, more or less at the same time as they came across Hotelling's (1929) potential for new applications.[1]

[1] This is not yet the end of the story. Several results obtained using the spatial competition model in industrial organization were anticipated by Vickrey in his *Microstatics* published in 1964.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Spatial models have proven to be very powerful tools because they account explicitly for the product specification chosen by firms, whereas the Chamberlinian and discrete choice models provide no basis for a theory of product choice and product design. Spatial models are appealing in two more respects. First, they capture consumer heterogeneity by means of a simple and suggestive metaphor, which has been used extensively to describe heterogeneous agents in several economic fields as well as in political science. Second, the spatial model of monopolistic competition is well suited for studying various facets of the market process, for example, by assuming that firms have a base product that is associated with the core competence of the firm. This product may be redesigned to match consumer requirements if the corresponding firm is willing to incur a cost that grows as the customized product becomes more differentiated from the base product. In this set-up, firms are multi-product and each variety is produced at a specific marginal cost. This problem may be studied by replacing Hotelling-like *shopping* models with *shipping* models, where firms deliver the product and take advantage of the fact that customer locations are observable to price discriminate across space (Macleod, Norman and Thisse, 1988; Eaton and Schmitt, 1994; Norman and Thisse, 1999; Vogel, 2011).

Unfortunately, spatial models become quickly intractable when they are cast in a general framework involving a non-uniform distribution of consumers and price-sensitive consumption of a variety. In those cases, showing the existence of a Nash equilibrium in pure strategies is problematic, especially when the location pattern is asymmetric.

2.3 Heterogeneous Consumers: The Discrete Choice Approach

There is a continuum of consumer types $\theta \in \mathbb{R}$. When n varieties are available, a consumer of type θ is described by a type-specific vector $\mathbf{e}^\theta = (e_1^\theta, \dots, e_n^\theta) \in \mathbb{R}^n$ of match values with the varieties, which can also be viewed as the consumer's transport costs she bears to reach the varieties. Each consumer buys one unit of a single variety. More specifically, the indirect utility from consuming variety i by a θ-type consumer is given by

$$V_i^\theta = y - p_i + e_i^\theta, \tag{5.16}$$

where y is the consumer's income. Given prices, a consumer chooses her "best buy," that is, the variety that gives her the highest surplus net of its price.

We assume that each type θ is distributed according to the same continuous density $f(\cdot)$ and cumulative distribution function $F(\cdot)$. In this case, the market demand for variety i is given by

$$x_i(\mathbf{p}) = \int_{-\infty}^{\infty} f(\theta) \prod_{k \neq i} F(p_k - p_i + \theta) \mathrm{d}\theta, \tag{5.17}$$

where, for any given type θ, $\Pi_{k\neq i} F(p_k - p_i + \theta)$ is the density of consumers who choose variety i at the price vector $\mathbf{p}$. The probability of indifference between two varieties being zero, each consumer buys a single variety. Because it embodies symmetry across varieties, (5.17) has a Chamberlinian flavor. However, even though consumer types obey the same distribution, they have heterogeneous tastes, for the probability that two types of consumers have the same match values is zero.

The Bertrand game associated with the demand system (5.17) can be studied along the lines of Section 2.1. Anderson, De Palma and Nesterov (1995) show that a price equilibrium

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

always exists if the density $f(\theta)$ is log-concave, while Perloff and Salop (1985) show the following result:

Proposition 2 *If either the support of the density $f(\theta)$ is bounded from above or $\lim_{\theta\to\infty} f'(\theta)/f(\theta) = -\infty$ then Bertrand competition converges to perfect competition as $n \to \infty$.*

Proposition 2 holds that the upper tail of the density of types is not "too" fat.[2] Under this condition, as new varieties enter the market, it becomes more likely that two varieties are very close substitutes, implying the two producers get trapped into a price war. This in turn pulls down all prices close to the marginal cost. This can be illustrated in the case of a normal distribution where $f'(\theta)/f(\theta) = -\theta$, so that $p^B(n)$ tends to c. Otherwise, the tail is fat enough for a growing number of varieties to enter the market while remaining distant enough from each other, thus allowing firms to price above the marginal cost even when n is arbitrarily large. For example, when the match values are drawn from the Gumbel distribution, we obtain (Anderson, De Palma and Thisse, 1992, ch. 7):

$$\lim_{n\to\infty} p^B(n) = \lim_{n\to\infty} \left(c + \frac{n}{n-1}\varkappa \right) = c + \varkappa, \tag{5.18}$$

where $\varkappa$ is the standard deviation of the Gumbel distribution up to the coefficient $\sqrt{6}/\pi$. Since $\varkappa > 0$, the limit of Bertrand competition is thus monopolistic competition. In addition, as a higher μ signals a more dispersed distribution of tastes across consumers, (5.18) implies that *a more heterogeneous population of consumers allows firms to charge a higher price*. Alternatively, we may say that each variety has a growing number of consumers prepared to buy it even at a premium. Observe also that $p^B(n) = c$ for all n when consumers are homogeneous because $\varkappa = 0$, as in the standard Bertrand duopoly. Last, Proposition 2 is the mirror image of Proposition 1. The former means that individual preferences must be sufficiently dispersed for monopolistic competition to be the equilibrium outcome in a large economy, while the latter shows that a strong love for variety is needed for monopolistic competition to arise.

The model (5.16) can be easily extended to cope with the case where consumers have ideal varieties and a variable consumption level. This can be achieved by assuming that

$$V_i^\theta = \psi(y) - \phi(p_i) + e_i^\theta,$$

where both ψ and ϕ are increasing. A natural candidate investigated by Sattinger (1984) is obtained when $\psi(y) = \ln y$ and $\phi(p) = \ln p$. In this case, individual consumptions are variable and determined by price ratios, rather than price differences. Under the assumptions of Proposition 2, it is readily verified that $\lim_{n\to\infty} p^B(n) = c$. However, this ceases to hold under the Gumbel distribution where

$$\lim_{n\to\infty} p^B(n) = \lim_{n\to\infty} c\left(1 + \frac{n}{n-1}\varkappa\right) = c(1+\varkappa).$$

2 Gabaix et al. (2016) show that markups are asymptotically determined by the tail behavior of the distribution of tastes.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Again, taste dispersion allows firms to set prices higher than marginal cost.

Note, finally, that the symmetry of preferences may be relaxed by assuming the vector of match values is drawn from a multivariate distribution $F(x_1, \ldots, x_n)$, such as the probit where the covariance measures the substitutability between the corresponding two varieties. Though empirically relevant, it is hard to characterize the market equilibrium at this level of generality.

The above models are related to, but differ from, Hart (1985a). As in discrete choice models, Hart focuses on consumers who have heterogeneous tastes. However, unlike these models where consumers can switch between varieties, individual choices are restricted to a *given* and *finite* set of desirable varieties, which is consumer specific. In equilibrium, every consumer chooses the quantity (which can be zero) to consume of each desirable variety. Hart (1985b) then shows that, in a large economy, a monopolistically competitive equilibrium exists if the taste distribution is sufficiently dispersed.

2.4 Where Do We Stand?

2.4.1 Summary

We have discussed three different families of models that describe preferences over differentiated products. In each case, the same conclusion emerges: the limit of Cournot or Bertrand competition may be monopolistic competition. Unlike what Robinson, Kaldor, Stigler and others have argued, a large number of firms need not imply perfect competition. As anticipated by Chamberlin, *when firms are many, their strategic power vanishes*. Nevertheless, *product differentiation may allow every firm to retain monopoly power over the demand for its variety* in an environment in which strategic considerations are banned.

Whereas the spatial models are very intuitive, the symmetric representative consumer models display a high degree of versatility. They both seem to belong to different worlds. This need not be so, however. The two families of models can generate the same market outcome. For this to happen, the market space of any variety must share a border with the market space of any other variety, while the distance between any two varieties must be the same. More specifically, if the number of characteristics is equal to $n - 1$, where n is the number of varieties, each firm competes directly with every other firm (Anderson, De Palma, and Thisse, 1992, ch. 4). To put it differently, a reconciliation between discrete choice theory, the representative consumer approach, and the spatial models of product differentiation is possible when *the number of product characteristics is sufficiently large relative to the number of product varieties*.

2.4.2 Syntheses

Two attempts at providing a synthesis of spatial and symmetric models are worth mentioning.[3] First, Chen and Riordan (2007) developed an ingenious synthesis of the spatial and variety-seeking models using a *spokes network*. There are $\mathcal{N}$ potential varieties and a unit mass of consumers uniformly distributed over $n \leq \mathcal{N}$ spokes connected at the center $x = 0$ of the plane. Each spoke has the same length $\Delta/2$ and a single store is located at the endpoint $x = \Delta/2$. The distance between any two stores is thus equal to Δ. A consumer's ideal variety is described by her location along a particular spoke. Consumer variety-seeking behavior is captured by assuming that each consumer may purchase her second most-preferred variety

[3] Other attempts include Hart (1985c), Deneckere and Rotschild (1992) and Ushchev and Zenou (2015).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

chosen by nature with a probability equal to $1/(\mathcal{N}-1)$, so that this variety need not be available. When $n \leq \mathcal{N}$ varieties are available, the demand for variety $i = 1, \ldots, n$ is formed by consumers whose ideal variety is i and those who choose i as a second choice. Each firm has some monopoly power on its spokes, but competes symmetrically with the other firms. Hence, the model combines the above two transport geographies.

Assuming that all consumers buy their most and second most-preferred varieties, the equilibrium price is given by

$$p^*(n) = c + t\Delta \frac{2\mathcal{N} - n - 1}{n - 1}.$$

As the number n of varieties/spokes grows and reaches the value $\mathcal{N}$, the equilibrium price decreases toward $c + t\Delta > c$. Hence, regardless of the value of $\mathcal{N}$ *the limit of the spokes model is monopolistic competition*, whereas the limit of the circular model is perfect competition ($\Delta = 0$). This echoes what we have seen in the foregoing.

Second, Anderson and De Palma (2000) developed an integrative framework that links spatial and symmetric models. A consumer buys a fixed number $\bar{x}$ of units of the differentiated product (e.g., a given number of restaurant dinners per month) and has an entropy-like utility:

$$U(\mathbf{x}) = \sum_{i=1}^{n} x_i - \varkappa \sum_{i=1}^{n} x_i \log x_i + x_0 \qquad \text{s.t.} \sum_{i=1}^{n} x_i = \bar{x},$$

where the parameter $\varkappa > 0$ is a measure of the degree of differentiation across varieties. This specification corresponds to a special case of (5.1) in which $u(x) = x - \varkappa \times \log x$ and $\varphi(X) = X$. The *entropy* of a consumption profile $\mathbf{x}$ may be viewed as a measure of its dispersion. Therefore, the impact of the entropy term on the consumer's utility level tells us how differentiated varieties are from the consumer's point of view.

Assume that identical consumers are uniformly distributed over the real line, while firms are equidistantly located over the set of integers $i = 0, \pm 1, \ldots$ let $t > 0$ be the unit shopping cost. In this case, a consumer located at s has a *logit* demand given by

$$x_i(\mathbf{p}; s) = \bar{x} \frac{\exp\left[-(p_i + t\,|s - i|)/\varkappa\right]}{\sum_{k=-\infty}^{\infty} \exp\left[-(p_k + t\,|s - k|)/\varkappa\right]} > 0.$$

Competition is localized when $\varkappa = 0$. As $\varkappa$ rises from zero, market boundaries get blurred: a firm's spatial market is encroached on by its competitors; but this firm also captures customers from its rivals. At the limit, when $\varkappa \to \infty$ the market demand is equally spread across firms. For given prices, the individual demand for any variety is positive, as in Chamberlinian models, but decreases with the distance between the consumer and the variety supplier. The market price is given by

$$p^*(\varkappa, t) = c + \varkappa \frac{(1+\phi)^2 \ln \phi}{2\phi(1+\phi) - \ln(1-\phi)^2},$$

where $\phi \equiv \exp(-t/\varkappa)$ is a measure of the degree of global competition in the market. Differentiating p^* with respect to t for any given $\varkappa$, or with respect to $\varkappa$ for any given

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

t, shows that higher transport costs or a stronger love for variety lead to a higher price because the former weakens competition between neighboring firms while the latter means that varieties are more differentiated. When $\varkappa \to 0$, $p^*(\varkappa, t)$ boils down to the equilibrium price of the circular model, $p^*(0, t) = c + tL/n$, while $p^*(\varkappa, t)$ converges to $c + \varkappa n/(n-1)$ when $t \to 0$. Hence, when heterogeneity prevails along one dimension only, the equilibrium markups remain positive. However, $p^*(\varkappa, t)/\varkappa$ being homogeneous of degree zero, we have $p^*(0, 0) = c$ when heterogeneity completely vanishes.

3 THE NEGLIGIBILITY HYPOTHESIS IN MONOPOLISTIC COMPETITION

From now on, we assume that the supply side of the economy is described by a *continuum* of negligible firms whose *mass* is determined by free entry and exit. The *negligibility assumption* has several important implications. First, it captures the essence of the Chamberlinian idea of monopolistic competition, summarized in the following quote: "A price cut, for instance, which increases the sales of him who made it, draws inappreciable amounts from the markets of each of his many competitors, achieving a considerable result for the one who cut, but without making incursions upon the market of any single competitor sufficient to cause him to do anything he would not have done anyway" (Chamberlin, 1933, p. 83).

Second, because each firm treats the market as a given, it faces a given residual demand, very much like a monopolist. As a consequence, a firm can indifferently choose its profit-maximizing price and output (Vives, 1999, p. 168). In other words, the negligibility assumption makes monopolistic competition immune to the difficult choice to be made between Cournot and Bertrand. Third, ever since Gabszewicz and Vial (1972), it is well known that the choice of a good produced by oligopolistic firms as the numéraire affects the equilibrium. Under the negligibility hypothesis, the choice of any particular variety as the numéraire has no impact on the market outcome. Last, one of the typical assumptions of monopolistic competition is that of free entry and exit. The role of this assumption is worth stressing. Indeed, positive (or negative) profits would affect individual incomes, hence firm demands. This feedback effect is precisely one of the major difficulties encountered when one aims to introduce oligopolistic competition in general equilibrium.

In this section, we illustrate those ideas by discussing the CES and LQ models. Anderson, Erkal and Piccinin (2015) argue that these models can be viewed as aggregative oligopoly games in which "firms do not internalize the effects of their actions on the aggregate." In other words, the CES and LQ models of monopolistic competition may be viewed as sequential games in which a "Chamberlinian auctioneer" first chooses the value of the aggregate, while firms move second. As a result, the market outcome under monopolistic competition generates lower prices (or higher quantities) than those obtained under oligopolistic competition. Such an interpretation is accurate when each firm is negligible to the market. However, for the reasons put forward by Aumann (1964), it is inaccurate when there is a finite and discrete number of firms because firms are able to manipulate the aggregate (d'Aspremont, Dos Santos Ferreira and Gérard-Varet, 1996). This highlights the role of the negligibility hypothesis in monopolistic competition.

Note, finally, that the negligibility hypothesis keeps its relevance when firms are multi-product. Contrary to general belief, a firm supplying a finite number or a range of varieties

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

remains negligible to the market (Allanson and Montagna, 2005; Bernard, Redding and Schott, 2010, 2011; Dhingra, 2013; Mayer, Melitz and Ottaviano, 2014, 2016). Therefore, issues studied in industrial organization, such as the cannibalization and customerization effects, can be revisited in simpler settings.

3.1 The CES Model of Monopolistic Competition

Even economists with minimal exposure to monopolistic competition have probably heard of the constant elasticity of substution (CES) model. There is little doubt that this model has led to a wealth of new results (Matsuyama, 1995). For this reason, we find it useful to describe briefly how the CES model works.

3.1.1 The benchmark set-up

Firms and consumers Labor is the only factor of production and is inelastically supplied in a competitive market; labor is chosen as the numéraire. There are L consumers endowed with y efficiency units of labor. They share the same CES utility function:

$$U = \left(\int_0^N x_i^{\frac{\sigma-1}{\sigma}} \mathrm{d}i \right)^{\frac{\sigma}{\sigma-1}}, \tag{5.19}$$

where $\sigma > 1$ is the elasticity of substitution between any two varieties. Maximizing U subject to the budget constraint yields the individual demand for variety i:

$$x_i = \frac{p_i^{-\sigma}}{\int_0^N p_i^{-(\sigma-1)} \mathrm{d}i} y, \qquad i \in [0, N]. \tag{5.20}$$

This expression implies that the supply of an infinitesimal interval of new varieties increases the denominator and, consequently, leads to a reduction in the demand for the existing varieties so long as their prices remain unchanged. In other words, the entry of new varieties triggers the "fragmentation" of demand over a wider range of varieties.

Let

$$P \equiv \left(\int_0^N p_i^{-(\sigma-1)} \mathrm{d}i \right)^{\frac{-1}{\sigma-1}}$$

be the CES *price index* of the differentiated good. The price index, which is the power mean of prices, decreases with the mass of varieties. Indeed, if a non-negligible range of new varieties Δ is added to the incumbent ones, we get

$$P = \left(\int_0^N p_i^{-(\sigma-1)} \mathrm{d}i \right)^{\frac{-1}{\sigma-1}} > P_\Delta = \left(\int_0^{N+\Delta} p_i^{-(\sigma-1)} \mathrm{d}i \right)^{\frac{-1}{\sigma-1}}.$$

To put it differently, the price index falls as if competition among a larger mass of competitors were to lead to lower prices. In addition, as suggested by spatial models, the less differentiated the varieties, the lower the price index.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

The market demand functions Lx_i may then be rewritten as follows:

$$X_i = p_i^{-\sigma} P^{\sigma-1} Ly. \tag{5.21}$$

Thus, a firm's demand accounts for the aggregate behavior of its competitors via the sole price index, and the game is aggregative. Since firm i is negligible to the market, it treats P as a parameter; although firms are price-makers, they are *price index-takers*. As a consequence, Triffin's condition $\partial X_i/\partial p_k = 0$ holds for all $k \neq i$. Furthermore, (5.21) implies that market demands are *isoelastic*, the price elasticity being equal to the elasticity of substitution σ. Finally, the market demand is still given in (5.21) when individual incomes are redistributed because the demand X_i depends on the aggregate income only. As a consequence, the market demand is independent of the income distribution.

Firms share the same fixed cost F and the same constant marginal cost c. In other words, to produce q_i units of its variety, firm i needs $F + cq_i$ efficiency units of labor. Hence, firm i's profit is given by

$$\Pi_i(q_i) = (p_i - c)q_i - F. \tag{5.22}$$

Market equilibrium A *symmetric free-entry equilibrium* (SFE) is a four-tuple (x^*, q^*, p^*, N^*), which satisfies the following four conditions: (i) no firm can increase its profit by deviating from q^*; (ii) x^* maximizes a consumer's utility subject to his or her budget constraint; (iii) the product market clearing condition

$$q^* = Lx^*$$

holds; (iv) the mass of firms is pinned down by the zero-profit condition (ZPC). The Walras Law implies that the labor market balance

$$N^* \cdot (F + cq^*) = Ly$$

is satisfied.

The first-order condition (FOC) shows that the equilibrium price is given by (the second-order condition [SOC] is satisfied):

$$p^* = \frac{\sigma}{\sigma - 1}c,$$

which increases when varieties get more differentiated, as in the various models discussed in Section 2. The markup is constant and equal to

$$\frac{p^* - c}{p^*} = \frac{1}{\sigma}. \tag{5.23}$$

In other words, firm markups are the same in large/small/rich/poor countries, the reason being that firms' demands are isoelastic. In game-theoretic terms, this means that firms have a dominant strategy – the reaction functions are flat – a result that probably explains the lack of interest among researchers in industrial organization for the CES model of monopolistic

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

competition. A constant markup runs against the conventional wisdom that asserts that entry fosters lower market prices. The markup (5.23) is also independent of shocks on marginal cost and market size, which contradicts a growing number of empirical studies (De Loecker and Goldberg, 2014).

The above criticisms need qualification, however. Even if the equilibrium price remains unchanged when the mass of firms increases, the consumption of the differentiated good is fragmented over a wider range of varieties. This in turn implies that each firm's profits go down. In other words, we come back, albeit very indirectly, to a kind of competitive effect as the entry of new firms has a negative effect on the profitability of the incumbents. Note also that the Lerner index increases exogenously with the degree of differentiation across varieties, which also agrees with one of the main messages of industrial organization, that is, product differentiation relaxes competition.

To determine the equilibrium firm size, one could substitute the equilibrium price into the demand function (5.21). By plugging prices and quantities into the ZPC, one could obtain the equilibrium mass of firms/varieties. It is in fact simpler, but strictly equivalent, to proceed in the reverse order by determining first the volume of production thanks to the free-entry condition given by

$$\Pi_i = (p^* - c)q_i - F = \frac{c}{\sigma - 1}q_i - F = 0,$$

which yields

$$q^* = \frac{\sigma - 1}{c}F. \tag{5.24}$$

Thus, regardless of the mass of firms, firm size remains constant. This result, which is a direct consequence of a constant markup, is one of the major weaknesses of the CES model: there is no scale effect as q^* is independent of the market size L.

It follows immediately from the labor market balance that

$$N^* = \frac{Ly}{\sigma F}. \tag{5.25}$$

Hence, when varieties are less (more) differentiated, the mass of firms is smaller (larger), while a firm's output is larger (smaller) because the market demand is less (more) fragmented. Furthermore, a higher degree of increasing returns is associated with larger output and fewer but larger firms.

There is no question that the CES model of monopolistic competition captures some fundamental features of imperfect competition. But, and this is a big but, it is at odds with the main corpus of oligopoly theory. Despite (or, perhaps, because of) its great flexibility in applications and econometric estimations, the CES model brushes under the carpet several effects that may deeply affect the results it gives rise to. Therefore, although this model is a natural point of departure in studying issues where imperfect competition and increasing returns are crucial, we find it hard to maintain that it can serve as a cornerstone of any sound theory. For this, we need alternative or more general models to test the robustness of the results. Note, finally, that how appealing a model is depends on what questions one is

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

interested in and whether the features from which the CES model abstracts are important for the issue in question.

3.1.2 The weighted CES

Assume that the CES is modified as follows:

$$U = \left(\int_0^N (a_i x_i)^{\frac{\sigma-1}{\sigma}} \mathrm{d}i \right)^{\frac{\sigma}{\sigma-1}}, \tag{5.26}$$

where $a_i > 0$ are salience coefficients whose purpose is to account for asymmetries among varieties (Bernard et al., 2010, 2011). If $x_i = x_j$, then $a_i > a_j$ implies that, everything else being equal, the utility of consuming variety i exceeds that of variety j. However, the consumer is indifferent between consuming a_i/a_j units of variety i and one unit of variety j. Therefore, the preferences (5.26) can be made symmetric by changing the units in which the quantities of varieties are measured. Nevertheless, changing the units in which varieties are measured implies that firms that are otherwise symmetric now face different marginal costs. To be precise, firm i's marginal cost is equal to c/a_i. This implies that *a CES with asymmetric preferences and symmetric firms is isomorphic to a CES in which preferences are symmetric and firms heterogeneous*. Accordingly, to discriminate between cost heterogeneity and the salience coefficients a_i, one needs data on prices and sales because prices reflect the heterogeneity in costs while the salience coefficients act like demand shifters in the CES (Kugler and Verhoogen, 2012).

3.2 Monopolistic Competition Under Linear-quadratic Preferences

3.2.1 The benchmark set-up

We have seen that there are (at least) two versions of LQ preferences defined over a finite number of varieties, namely (5.12) and (5.13). Even though the former is not the limit of oligopolistic competition, it is associated with equilibrium values of the main variables that vary with the key parameters of the model (Ottaviano, Tabuchi and Thisse, 2002):[4]

$$U(\mathbf{x}) = \alpha \int_0^N x_i \mathrm{d}i - \frac{\beta}{2} \int_0^N x_i^2 \mathrm{d}i - \frac{\gamma}{2} \int_0^N \left(\int_0^N x_k \mathrm{d}k \right) x_i \mathrm{d}i + x_0. \tag{5.27}$$

One unit of labor is needed to produce one unit of the homogeneous good x_0, which is sold under perfect competition. This good is chosen as the numéraire so that the equilibrium wage is equal to 1. A consumer's budget constraint is as follows:

$$\int_0^N p_i x_i \mathrm{d}i + x_0 = 1 + \bar{x}_0, \tag{5.28}$$

where $\bar{x}_0$, the initial endowment in the numéraire, is supposed to be large enough for the consumption of this good to be strictly positive at the market outcome.

[4] Papers that use the LQ model include Belleflamme, Picard and Thisse (2000), Nocke (2006), Foster, Haltiwanger and Syverson (2008) and Dhingra (2013).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Solving (5.28) for the numéraire consumption, substituting the corresponding expression into (5.27) and solving the FOCs with respect to x_i yields the individual inverse demand for variety i:

$$p_i = \alpha - \beta x_i - \gamma X \qquad i \in [0, N] \tag{5.29}$$

where

$$X \equiv \int_0^N x_k \mathrm{d}k$$

is the total individual consumption of the differentiated product. Varieties interact through the value of X, which determines the demand intercept $\alpha - \gamma X$, so that a hike in X renders the inverse demand more elastic. As a consequence, *when choosing its output level each firm must guess what X will be*, meaning that the game is aggregative in nature.

Firm i's profit function is as follows:

$$\Pi_i = (p_i - c)q_i - F = L \cdot \left[(p_i - c)x_i - \frac{F}{L}\right],$$

so that maximizing Π_i with respect to q_i amounts to maximizing the bracketed term with respect to x_i. To ease the burden of notation, we assume that c is equal to zero, which amounts to rescaling the demand intercept.

The best-reply function

$$x^*(X) = \frac{\alpha - \gamma X}{2\beta}$$

shows how *each firm plays against the market* as x^* decreases with X. Since the equilibrium values of x and X must satisfy the condition $Nx = X$, for any given mass N of firms, the consumption $x^*(N)$ is given by

$$x^*(N) = \frac{\alpha}{2\beta + \gamma N}, \tag{5.30}$$

which decreases with the mass of competitors. Using (5.30), (5.29) yields the price $p^*(N)$:

$$p^*(N) = \frac{\alpha\beta}{2\beta + \gamma N} = \beta x^*(N), \tag{5.31}$$

which also decreases with N. Thus, unlike the CES, entry generates pro-competitive effects. In addition, as suggested by product differentiation theory, the market price rises when varieties get more differentiated (lower γ).

The ZPC implies that the equilibrium mass of firms is given by

$$N^* = \frac{1}{\gamma}\left(\alpha\sqrt{\frac{\beta L}{F}} - 2\beta\right). \tag{5.32}$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

It is readily verified that N^* increases at a decreasing rate with the market size (L), the consumer's willingness-to-pay for the differentiated product (α), the degree of product differentiation ($1/\gamma$), whereas it decreases with the marginal and fixed costs (c and F).

Substituting (5.32) in (5.30) and multiplying by L gives the equilibrium output:

$$q^* = Lx^* = \sqrt{\frac{FL}{\beta}},$$

which increases with L at a decreasing rate, while a stronger love for variety allows more firms to enter the market, but they all have a smaller size. Plugging (5.32) in (5.31) gives the equilibrium price:

$$p^* = \sqrt{\frac{\beta F}{L}},$$

which decreases with L at an increasing rate.

Thus, market size and cost parameters matter for all the equilibrium variables under free entry. This makes the linear model of monopolistic competition a good proxy of an oligopolistic market. Notwithstanding the absence of an income effect, the linear model performs fairly well in trade theory and economic geography (Ottaviano and Thisse, 2004; Melitz and Ottaviano, 2008).

4 THE VES MODEL OF MONOPOLISTIC COMPETITION

Choosing an appropriate framework for studying imperfect competition involves a trade-off between allowing firms to have sophisticated behavior and capturing basic general equilibrium effects. In this section, we discuss a model that aims to find a prudent balance between those two objectives. The CES and LQ models, as well as the translog developed by Feenstra (2003), are all special cases. Firms are still symmetric, which allows one to insulate the impact of preferences on the market outcome and to assess the limitations of specific models.

4.1 Firms and Consumers

Owing to their analytical simplicity, the CES and LQ models conceal a difficulty that is often ignored: working with a continuum of goods implies that we cannot use the standard tools of calculus. Rather, we must work in a functional space whose elements are functions, and not vectors.

Let $\mathbb{N}$, an arbitrarily large number, be the mass of *potential* varieties. As all potential varieties are not necessarily made available to consumers, we denote by $N \leq \mathbb{N}$ the endogenous mass of *available* varieties. A *consumption profile* $\mathbf{x} \geq 0$ is a Lebesgue-measurable mapping from the space of potential varieties $[0, \mathbb{N}]$ to $\mathbb{R}_+$, which is assumed to belong to $L_2([0, \mathbb{N}])$. Individual preferences are described by a *utility functional* $U(\mathbf{x})$ defined over the positive cone of $L_2([0, \mathbb{N}])$. In what follows, we assume that (i) U is symmetric over the range of potential varieties in the sense that any Lebesgue measure-preserving mapping from $[0, \mathbb{N}]$ into itself does not change the value of U, and (ii) U exhibits a love for variety. To

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

determine the inverse demand for a variety, Parenti, Ushchev and Thisse (2017) assume that there exists a unique function $D(x_i, \mathbf{x})$ from $\mathbb{R}_+ \times L_2$ to $\mathbb{R}_+$ such that, for any given N and for all $\mathbf{h} \in L_2$, the equality

$$U(\mathbf{x}+\mathbf{h}) = U(\mathbf{x}) + \int_0^N D(x_i, \mathbf{x})\, h_i \, \mathrm{d}i + \circ(||\mathbf{h}||_2) \tag{5.33}$$

holds, $||\cdot||_2$ being the L_2-norm.[5] The function $D(x_i, \mathbf{x})$, which is the marginal utility of variety i, is the same across varieties because preferences are symmetric. Parenti et al. (2017) focus on the utility functionals such that the marginal utility $D(x_i, \mathbf{x})$ is decreasing and twice differentiable with respect to x_i.

Maximizing the utility functional $U(\mathbf{x})$ subject to (i) the budget constraint

$$\int_0^N p_i x_i \mathrm{d}i = y,$$

and (ii) the availability constraint

$$x_i \geq 0 \text{ for all } i \in [0, N] \quad \text{and} \quad x_i = 0 \text{ for all } i \in [N, \mathbb{N}]$$

yields the inverse demand function for variety i:

$$p_i = \frac{D(x_i,\, \mathbf{x})}{\lambda} \quad \text{for all } i \in [0, N], \tag{5.34}$$

where λ is the Lagrange multiplier of the consumer's optimization problem. Expressing λ as a function of y and $\mathbf{x}$ yields

$$\lambda(y, \mathbf{x}) = \frac{\int_0^N x_i D(x_i, \mathbf{x})\, \mathrm{d}i}{y},$$

which is the marginal utility of income at the consumption profile $\mathbf{x}$ and income y.

Firm i maximizes (5.22) with respect to its output q_i subject to the inverse market demand function $p_i = LD/\lambda$, while the market outcome is given by a Nash equilibrium. Being negligible, each firm accurately treats the variables $\mathbf{x}$ and λ in (5.34) as parameters. Note the difference between the consumer and producer programs. The individual chooses a consumption level for all available varieties. By contrast, each firm selects an output level for a single variety. In other words, the consumer's choice variable $\mathbf{x}$ is defined on a non-zero measure set while firm i's choice variable q_i is defined on a zero-measure set. Thus, unlike in Aumann (1964), the key ingredient of monopolistic competition is the negligibility of firms rather than that of consumers.[6]

[5] Formally, this means that $U(\mathbf{x})$ is Fréchet-differentiable, which extends in a fairly natural way the standard concept of differentiability to L_2.

[6] We thank Kristian Behrens for having pointed out this difference to us.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Plugging (5.34) into (5.22) and using the product market clearing condition, the program of firm i may be rewritten as follows:

$$\max_{x_i} \Pi_i(x_i, \mathbf{x}) \equiv \left[\frac{D\,(x_i, \mathbf{x})}{\lambda} - c\right] L x_i - F.$$

Setting

$$D_i' \equiv \frac{\partial D(x_i, \mathbf{x})}{\partial x_i} \qquad D_i'' \equiv \frac{\partial D^2(x_i, \mathbf{x})}{\partial x_i^2},$$

the FOC for profit maximization is given by

$$D(x_i, \mathbf{x}) + x_i D_i' = [1 - \eta(x_i, \mathbf{x})]\, D(x_i, \mathbf{x}) = \lambda c, \tag{5.35}$$

where

$$\eta(x_i, \mathbf{x}) \equiv -\frac{x_i}{D(x_i, \mathbf{x})} \frac{\partial D(x_i, \mathbf{x})}{\partial x_i}$$

is the elasticity of the inverse demand for variety i. The right-hand side of (5.35) is variable, and thus each firm must guess what the equilibrium value of λ is to determine its profit-maximizing output. Parenti et al. (2017) show that the profit function Π_i is strictly quasi-concave in x_i for all admissible values of λc if and only if (**A**) *firm i's marginal revenue decreases in* x_i.

4.2 The Elasticity of Substitution

We have seen that the elasticity of substitution plays a central role in the CES model of monopolistic competition. Many would argue that this concept is relevant for such preferences only. This opinion is unwarranted. The argument goes as follows. It is well known that, when the number of goods exceeds 2, there are different definitions of the elasticity of substitution. Parenti et al. (2017) choose one of them, $\bar{\sigma}$, evaluated at $(x_i, \mathbf{x})$ where i is any arbitrary variety and $\mathbf{x} = xI_{[0,N]}$ but for i. Treating N parametrically, this amounts to considering $\bar{\sigma}\,(x_i, x; N)$, which depends only on two goods, that is, x_i and x. In this case, the choice of $\bar{\sigma}$ no longer matters because all definitions of the elasticity of substitution are equivalent. As a result, by setting

$$\sigma\,(x, N) \equiv \overline{\sigma}(x, x; N),$$

hence, along the diagonal, *the elasticity of substitution among varieties hinges only upon the individual consumption per variety and the total mass of available varieties*. Under the CES, $\sigma\,(x, N)$ is a constant.

To gain insights about the behavior of σ, we give below the elasticity of substitution for the two main families of preferences used in the literature: (i) when the utility is additive, we have:

$$\sigma\,(x, N) = \frac{1}{r_u(x)}, \tag{5.36}$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

where $r_u(x)$ is the relative love for variety (see Section 2.1). As implied by (5.36), σ depends only upon the individual per variety consumption; (ii) When preferences are homothetic, the elasticity of substitution evaluated at a symmetric consumption profile depends solely on the mass N of available varieties:

$$\sigma(x, N) = \frac{1}{\eta(1, I_{[0,N]})} \equiv \frac{1}{\mathcal{M}(N)}. \tag{5.37}$$

Using (5.36) and (5.37), we have the following: $\mathcal{E}_N(\sigma) = 0$ when preferences are additive, while $\mathcal{E}_x(\sigma) = 0$ when preferences are homothetic.

4.3 Market Equilibrium

Assume that (**A**) holds. Then, for any given $N \leq Ly/F$, Parenti et al. (2017) show that there exists a unique Nash equilibrium such that (i) no firm can increase its profit by changing its output; (ii) each consumer maximizes utility subject to her budget constraint; (iii) the product markets clear; (iv) the labor market balance holds. Furthermore, this equilibrium is symmetric and given by

$$x^*(N) = \frac{y}{cN} - \frac{F}{cL}, \qquad q^*(N) = \frac{yL}{cN} - \frac{F}{c}, \qquad p^*(N) = c\frac{\sigma(x^*(N), N)}{\sigma(x^*(N), N) - 1}, \tag{5.38}$$

and thus the equilibrium markup is

$$m^*(N) \equiv \frac{p^*(N) - c}{p^*(N)} = \frac{1}{\sigma(x^*(N), N)}, \tag{5.39}$$

which generalizes the expression (5.23) obtained under the CES. First of all, (5.39) suffices to show that, *in monopolistic competition working with a variable markup amounts to assuming a variable elasticity of substitution and non-isoelastic demands*. Furthermore, as in oligopoly theory, all variables depend on the mass of active firms. In particular, the equilibrium per variety consumption $x^*(N)$ always decreases with N, whereas the impact of N on $m^*(N)$ is a priori undetermined. To be precise, since $\sigma(x^*(N), N)$ may increase or decrease with the mass of firms, entry may generate pro- or anti-competitive effects. This in turn shows why comparative statics may give rise to diverging results in models where preferences are characterized by different functions $\sigma(x, N)$. In a nutshell, monopolistic competition is able to mimic oligopolistic competition. Finally, since $q^*(N)$ decreases with N, there is a business-stealing effect regardless of preferences.

Using (5.38) yields the operating profits earned by a firm:

$$\Pi^*(N) = \frac{c}{\sigma\left(x^*(N), N\right) - 1} Lx^*(N) - F, \tag{5.40}$$

Solving the ZPC $\Pi^*(N) = 0$ with respect to m yields a single equilibrium condition:

$$m^*(N) = \frac{NF}{Ly}. \tag{5.41}$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Setting $m \equiv FN/(Ly)$, (5.39) may be rewritten as a function of m only:

$$m\sigma\left(\frac{F}{cL}\frac{1-m}{m}, \frac{Ly}{F}m\right) = 1. \tag{5.42}$$

This expression shows that *a variable elasticity of substitution $\sigma(x, N)$ is sufficient to characterize the market outcome under general symmetric preferences and symmetric firms.* Note that (5.42) implies that σ *must* be a function, and not a constant, for the markup to be variable in our general framework. Since (5.42) involves the four structural parameters of the economy (L, y, c and F), how the market outcome varies with these parameters depends on how σ varies with x and N.

Although the above framework allows for very different patterns of substitution across varieties, it should be clear that they are not equally plausible. This is why most applications of monopolistic competition focus on different subclasses of utilities to cope with particular effects. Admittedly, making "realistic" assumptions on how the elasticity of substitution varies with x and N is not an easy task. That said, it is worth recalling with Stigler (1969) that "it is often impossible to determine whether assumption A is more or less realistic than assumption B, except by comparing the agreement between their implications and the observable course of events." This is what we will do below.

Spatial and discrete choice models of product differentiation suggest that varieties become closer substitutes when the number of competing varieties rises (Salop, 1979; Anderson et al., 1995). This leads Feenstra and Weinstein (2017) to use the translog expenditure function, where $\sigma(N) = 1 + \beta N$ increases with N, to capture the pro-competitive effects of entry. Therefore, $\mathcal{E}_N(\sigma) \geq 0$ seems to be a reasonable assumption. In contrast, how σ varies with x is a priori less clear. Nevertheless, this question can be answered by appealing to the literature on pass-through.

If a firm's demand is not too convex, the pass-through of a cost change triggered by a trade liberalization or productivity shock is smaller than 100 percent for a very large family of demand functions (Greenhut, Norman and Hung, 1987). More importantly, the empirical evidence strongly suggests that the pass-through is incomplete (De Loecker et al., 2016). Which assumption about σ leads to this result? The intuition is easy to grasp when preferences are additive, that is, $m(x) = r_u(x) = \sigma(x)$. Incomplete pass-through amounts to saying that p/c increases when c decreases, which means that firms have more market power or, equivalently, varieties are more differentiated. As firms facing a lower marginal cost produce more, the per capita consumption increases. Therefore, it must be that $\sigma(x)$ decreases with x. In the case of general symmetric preferences, Parenti et al. (2017) show that the pass-through is smaller than 100 percent *if and only if* $\mathcal{E}_x(\sigma) < 0$ holds. In addition, the pass-through must be equal to 100 percent when preferences are homothetic because $\mathcal{E}_x(\sigma) = 0$.

This discussion suggests the following conditions:

$$\mathcal{E}_x(\sigma) \leq 0 \leq \mathcal{E}_N(\sigma). \tag{5.43}$$

Even though these inequalities do not hold for some preferences, it is convenient to assume here that (5.43) holds. Applying Propositions 1 to 4 of Parenti et al. to (5.43) then implies:

Proposition 3 *Assume that* (**A**) *and* (5.43) *hold. Then,* (i) *there exists a free-entry equilibrium for all $c > 0$ and $F > 0$*; (ii) *this equilibrium is unique and symmetric;* (iii) *a larger market or*

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

a higher income leads to lower markups, bigger firms and a larger number of varieties; (iv) *the pass-through rate of a cost change is smaller than* 100 *percent.*

The pro-competitive effects associated with the extent of the market are intuitively plausible and supported by empirical evidence (Amiti, Itskhoki and Konings, 2016; De Loecker et al., 2016). Furthermore, a model in which the pass-through is smaller than 100 percent allows one to study firms' pricing policies, such as spatial price discrimination where firms charge different producer prices in response to differences in demand conditions, something that is hard to accomplish with the CES. Nevertheless, one should bear in mind that the industrial organization literature highlights the possibility of anti-competitive effects (Chen and Riordan, 2008). Moreover, result (iv) of Proposition 3 may be used to study how firms react to a shock which affects aggregate productivity, as in Bilbiie, Ghironi and Melitz (2012). To capture the versatility of the market outcome in the present setting, Parenti et al. (2017) provide a complete description of the comparative static effects through necessary and sufficient conditions, which may be used to pin down the restrictions on preferences for the equilibrium outcome to be consistent with the stylized facts.

Last, since we focus on monopolistic competition, the markup (5.39) stems directly from preferences through only the elasticity of substitution. This stands in stark contrast to oligopoly models where the markup emerges as the outcome of the interplay between preferences *and* strategic interactions. Nevertheless, by choosing appropriately the elasticity of substitution as a function of x and N, monopolistic competition is able to replicate the direction of comparative static effects generated in symmetric oligopoly models with free entry, as well as their magnitude. Therefore, as conjectured by Mas-Colell (1984), *monopolistic competition may be considered as the marriage between the negligibility hypothesis and oligopolistic competition.*

4.3.1 Additive preferences

Let $u(\cdot)$ be a strictly increasing and strictly concave function. Assume that the utility functional is as follows:

$$U(\mathbf{x}) = \int_0^{\mathbb{N}} u(x_i)\mathrm{d}i. \tag{5.44}$$

Since $D(x_i, \mathbf{x}) = u'(x_i)$, the marginal utility of variety i is independent of the other varieties' consumption. This property suggests that additive models retain, at least partially, the tractability of the CES. And indeed, since $\mathcal{E}_N(\sigma) = 0$, the equilibrium condition (5.42) becomes simpler:

$$m = r_u\left(\frac{F}{cL}\frac{1-m}{m}\right). \tag{5.45}$$

The equilibrium markup m^* is a fixed point of the function $r_u(\cdot)$, which maps the interval $[0, 1]$ into itself. If $r_u'(x) \geq 0$ for all $x \geq 0$, then the right-hand side of (5.45) is weakly decreasing, and thus m^* is always unique. When $r_u'(x) < 0$, showing uniqueness is less straightforward. However, if (**A**) holds, the right-hand side of (5.45) is a contraction mapping over $[0, 1]$, which implies that the equilibrium exists and is unique.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

To illustrate, consider the CARA utility $u(x) = 1 - \exp(-\alpha x)$ studied in Behrens and Murata (2007). Since the RLV is given by $r_u(x) = \alpha x$, (5.45) is the following quadratic equation:

$$m^2 + \frac{\alpha F}{cL} m - \frac{\alpha F}{cL} = 0,$$

the solution of which is as follows:

$$m^* = \frac{\alpha F}{2cL}\left(\sqrt{1 + 4\frac{cL}{\alpha F}} - 1\right). \tag{5.46}$$

Equation (5.46) gives a clue to understanding the asymptotic behavior of the market outcome: when the market is arbitrarily large, the equilibrium markup is arbitrarily close to zero. Thus, the economy features a competitive limit, which echoes what we saw in Section 2.1. Note that this is not so under the CES where $m^* = 1/\sigma > 0$ for all L.

Using (5.46) and recalling that $m = NF/(Ly)$ yields the equilibrium number of firms:

$$N^* = \frac{\alpha y}{2c}\left(\sqrt{1 + 4\frac{cL}{\alpha F}} - 1\right). \tag{5.47}$$

Plugging (5.47) into (5.38) pins down the equilibrium values of the remaining variables:

$$q^* = \frac{F}{2c}\left(\sqrt{1 + 4\frac{cL}{\alpha F}} - 1\right), \qquad p^* = c + \frac{\alpha F}{2L}\left(\sqrt{1 + 4\frac{cL}{\alpha F}} + 1\right). \tag{5.48}$$

Expressions (5.46)–(5.48) provide a complete solution of the CARA model. Furthermore, they imply unambiguous comparative statics with respect to L: *an increase in population leads to a drop in markup, price, and per variety consumption, an increase in firm size, and a less than proportional increase in the number of firms*.

Are these findings robust against the choice of alternative specifications for u? Zhelobodko et al. (2012) show the following result: *if r_u is strictly increasing in x, then a larger market leads to a lower markup, bigger firms and a larger number of varieties, whereas the opposite holds when r_u is strictly decreasing in x.* Evidently, when r_u is constant, whence preferences are CES, L has no impact on the market outcome.

The above discussion also shows that the individual income y has no impact on the market solution. This led Bertoletti and Etro (2017) to work with indirectly additive preferences:

$$\mathcal{V}(\mathbf{p}; y) \equiv \int_0^{\mathbb{N}} v\left(\frac{y}{p_i}\right) \mathrm{d}i, \tag{5.49}$$

where v is strictly increasing, strictly concave, and homogeneous of degree zero. Such preferences mean that $\mathcal{E}_x(\sigma) = \mathcal{E}_N(\sigma)$. Applying the concept of RLV to v, Bertoletti and Etro show that the equilibrium price depends on y but not on L. Since $\mathcal{E}_N(\sigma) = 0$ for non-CES

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

additive preferences and $\mathcal{E}_x(\sigma) = 0$ for non-CES homothetic preferences, indirectly additive preferences are disjoint from these two classes of preferences apart from the CES.

4.3.2 Homothetic preferences

There are several reasons for paying attention to homothetic preferences. First, such preferences retain much of the CES tractability. In particular, the marginal utility $D(x_i, \mathbf{x})$ of variety i is positive homogeneous of degree zero: $D(tx_i, t\mathbf{x}) = D(x_i, \mathbf{x})$ for all $t > 0$. By implication, an increase in income y leads to a proportional change in the consumption pattern $\mathbf{x}$ and leaves the relative consumptions x_i/x_j unchanged. Second, an appealing feature of homothetic preferences is that they can be easily nested into a multi-sectoral setting, for the aggregate price index is always well defined.

Homothetic preferences were used by Bilbiie et al. (2012) to study real business cycles to capture the fact that both markups and the number of firms are highly procyclical variables, while Feenstra and Weinstein (2016) use *translog* preferences for studying international trade. It is well known that there is no closed-form expression of the translog utility functional, which is instead defined by the expenditure functional:

$$\ln E(\mathbf{p}, U) = \ln U + \frac{1}{\mathbb{N}} \int_0^{\mathbb{N}} \ln p_i \mathrm{d}i - \frac{\beta}{2\mathbb{N}} \left[\int_0^{\mathbb{N}} (\ln p_i)^2 \mathrm{d}i - \frac{1}{\mathbb{N}} \left(\int_0^{\mathbb{N}} \ln p_i \mathrm{d}i \right)^2 \right]. \tag{5.50}$$

Using (5.37), we find that under homothetic preferences the equilibrium condition (5.42) reduces to

$$m = \mathcal{M}\left(\frac{Ly}{F} m\right). \tag{5.51}$$

Under (5.43), $\mathcal{M}(N)$ is a decreasing function of N, and thus there exists a unique equilibrium markup.

Contrasting the properties of (5.51) with those of (5.45) provides an insightful comparison of the market outcomes generated by, respectively, homothetic and additive preferences. The most striking difference is that (5.45) does not involve y as a parameter. In other words, assuming additive preferences implies that per capita income shocks are irrelevant for understanding changes in markups, prices and firm sizes. This property of additive preferences justifies the choice of population size as a measure of the market size. In contrast, (5.51) involves both L and y through the product Ly, i.e., the total GDP. Another interesting feature of homothetic preferences is that, unlike (5.45), (5.51) does not involve the marginal cost c. This yields an important comparative statics result: *under monopolistic competition with non-CES homothetic preferences, the markup is variable but the pass-through is always* 1.

As in the case of additive preferences, we proceed by studying an analytically solvable non-CES example. We choose to work with translog preferences (5.50). In this case, $\mathcal{M}(N) = 1/(1 + \beta N)$, while (5.51) is given by the following quadratic equation:

$$m^2 + \frac{F}{\beta Ly} m - \frac{F}{\beta Ly} = 0,$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

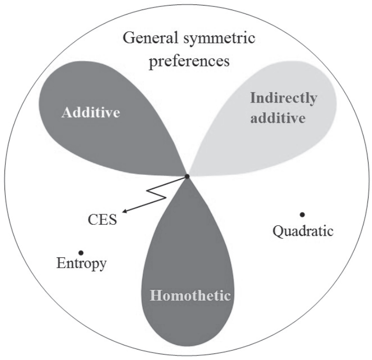


Figure 5.1 The space of preferences

whose solution is

$$m^* = \frac{F}{2\beta Ly}\left(\sqrt{1+4\frac{\beta Ly}{F}}-1\right). \tag{5.52}$$

Despite the differences between additive and homothetic preferences, the equilibrium markup (5.52) bears a remarkable resemblance to that obtained under the CARA utility, given by (5.46). In particular, (5.52) implies that $m^* \to 0$ in a large economy, i.e., when $Ly \to \infty$.

The equilibrium mass of firms can be determined by combining $m = FN/(Ly)$ with (5.52):

$$N^* = \frac{1}{2\beta}\left(\sqrt{1+4\frac{\beta Ly}{F}}-1\right). \tag{5.53}$$

Plugging (5.53) into (5.38) and rearranging terms yields:

$$q^* = \frac{F}{2c}\left(\sqrt{1+4\frac{\beta Ly}{F}}-1\right), \qquad p^* = c + \frac{cF}{2\beta Ly}\left(\sqrt{1+4\frac{\beta Ly}{F}}+1\right). \tag{5.54}$$

Expressions (5.52)–(5.54) yield a complete analytical solution of the translog model and entail unambiguous comparative statics results: *an increase in GDP triggers a decrease in prices and markups, increases firm size, and invites more firms to enter the market*. The same holds for any symmetric homothetic preference satisfying (5.43). What is more, (5.52)–(5.54) are strikingly similar to (5.46)–(5.48). To be precise, the CARA and translog models yield the same market outcome up to replacing the population L by the total GDP Ly.

Figure 5.1 shows the three subclasses of preferences used in the literature. The CES is the only one that belongs to all of them, which highlights how peculiar these preferences are.

5 HETEROGENEOUS FIRMS

In this section, we follow Melitz (2003) and assume that firms face different marginal costs. In this context, the key question is how the market selects the operating firms. We consider

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

the one-period framework used by Jean (2002) and Melitz and Ottaviano (2008). Prior to entry, risk-neutral firms face uncertainty about their marginal cost while entry requires a sunk cost F_e. Once this cost is paid, firms observe their marginal cost drawn randomly from the continuous probability distribution $\Gamma(c)$ defined over $\mathbb{R}_+$. After observing its type c, each entrant decides whether to produce or not, given that an active firm must incur a fixed production cost F. Under such circumstances, the mass of entrants, N_e, is larger than the mass of operating firms, N.

Even though varieties are differentiated from the consumer's point of view, firms sharing the same marginal cost c behave in the same way and earn the same profit at the equilibrium. As a consequence, we may refer to any variety/firm by its c-type only. Furthermore, the envelope theorem implies that equilibrium profits always decrease with c. Hence, there is perfect sorting of firms by increasing order of marginal cost. In other words, there exists a value $\bar{c}$ such that all operating firms have a marginal cost smaller than or equal to $\bar{c}$, while firms having a marginal cost exceeding $\bar{c}$ choose not to produce. A consumer program may then be written as follows:

$$\max_{x_c(.)} \mathcal{U} \equiv N_e \int_0^{\bar{c}} u(x_c)\mathrm{d}\Gamma(c) \qquad \text{s.t. } N_e \int_0^{\bar{c}} p_c x_c \mathrm{d}\Gamma(c) = y,$$

where $x_c \geq 0$ is the individual consumption of a c-variety. The mass of operating firms is then given by $N = N_e\Gamma(\bar{c})$. Since the distribution Γ is given, the equilibrium consumption profile is entirely determined by $\bar{c}$ and N_e. For homogeneous firms, the variable N is sufficient to describe the set of active firms.

A *free-entry equilibrium* $(c^*, N_e^*, q^*_{c\leq c^*}, x^*_{c\leq c^*}, \lambda^*)$ must satisfy the following equilibrium conditions:

1. the profit-maximization condition for c-type firms:

$$\max_{x_c} \Pi_c(x_c, \mathbf{x}) \equiv \left[\frac{D(x_c, \mathbf{x})}{\lambda} - c\right] Lx_c - F;$$

2. the ZPC for the cutoff firm:

$$(p_{c^*} - c^*)q_{c^*} = F,$$

   where c^* is the cutoff cost. At the equilibrium, firms are sorted out by decreasing order of productivity, which implies that the mass of active firms is equal to $N \equiv N_e\Gamma(c^*)$;
3. the product market clearing condition:

$$q_c = Lx_c$$

   for all $c \in [0, c^*]$;
4. the labor market clearing condition:

$$N_e F_e + \int_0^{c^*} (F + cq_c)\mathrm{d}\Gamma(c) = yL;$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

5. firms enter the market until their expected profits net of the entry cost F_e are zero:

$$\int_0^{c^*} \Pi_c(x_c, \mathbf{x}) \mathrm{d}\Gamma(c) = F_e.$$

Although some entrants earn positive profits whereas others lose money, the last condition implies that total profits are zero. Hence, yL is the total income.

5.1 Additive Preferences

Melitz (2003) and successors assume that consumers have CES preferences. This vastly simplifies the analysis because the equilibrium price of a c-type firm, $p^*(c) = c\sigma/(\sigma - 1)$, does not depend on the cost distribution Γ, although the price index does. Given that many properties derived under CES preferences are not robust, we focus below on additive preferences.

The inverse demand function (5.34) becomes $p_c(x_c) = u'(x_c)/\lambda$, which implies that the demand structure retains the simplicity of the homogeneous firm case. The profits made by a c-type firm are given by

$$\Pi(x_c; \lambda) = \left[\frac{u'(x_c)}{\lambda} - c\right] Lx_c - F.$$

Rewriting (5.39) for each type c implies that the equilibrium markup of a c-type firm is given by

$$m_c^* = r_u(x^*(c)) = 1/\sigma(x^*(c)), \tag{5.55}$$

which extends (5.23) to markets where firms are heterogeneous. It follows immediately from (5.55) that the elasticity of substitution is now c-specific in that it is the same *within* each type, whereas it varies *between* types. Furthermore, as firms of different types charge different prices, the individual consumption $x^*(c)$ varies with the firm's type, so that the equilibrium markup also varies with c. Evidently, a more efficient firm sells at a lower price than a less efficient firm. Therefore, consumers buy more from the former than from the latter, so that *a firm's markup increases* (decreases) *with its degree of efficiency when the RLV is increasing* (decreasing).

Given the second-order condition for profit-maximization ($r_{u'}(\cdot) < 2$), for each type c the expression

$$\bar{\pi}(c, \lambda; L) \equiv \max_{q \geq 0} \left[\frac{u'(q/L)}{\lambda} q - cq\right]$$

is a well-defined and continuous function. Since $\bar{\pi}(c, \lambda, L)$ is strictly decreasing in c, the solution $\bar{c}(\lambda; L)$ to the equation $\bar{\pi}(c; \lambda, L) - F = 0$ is unique. Clearly, the free-entry condition may be rewritten as follows:

$$\int_0^{\bar{c}(\lambda;L)} [\bar{\pi}(c, \lambda; L) - F] \,\mathrm{d}\Gamma(c) - F_e = 0. \tag{5.56}$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Using the envelope theorem and the ZPC at $\bar{c}(\lambda; L)$, we find that $\bar{\pi}(c, \lambda; L)$ and $\bar{c}(\lambda; L)$ are both decreasing functions of λ, which implies that the left-hand side of (5.56) is also decreasing in λ. As a consequence, the above equation has a unique solution $\bar{\lambda}(L)$. Plugging this expression into $\bar{c}(\lambda; L)$ yields the equilibrium cutoff $c^*(L)$. In other words, the free-entry equilibrium, if it exists, is unique. The expression (5.56) also shows that $c^*(L)$ exists when the fixed production cost F and entry cost F_e are not too large.

We are now equipped to study the impact of market size on the cutoff cost. The ZPC at $\bar{c}$ implies that

$$\frac{\partial \bar{\pi}}{\partial L} + \frac{\partial \bar{\pi}}{\partial c}\frac{d\,\bar{c}}{dL} + \frac{\partial \bar{\pi}}{\partial \lambda}\frac{d\bar{\lambda}}{dL} = 0.$$

Rewriting this expression in terms of elasticity and applying the envelope theorem to each term, it can be shown that the elasticity of c^* with respect to L is, up to a positive factor, equal to

$$\int_0^{\bar{\theta}} \left[r_u(x^*(c^*)) - r_u(x^*(c)) \right] R^*(c) \mathrm{d}\Gamma(c),$$

where $R^*(c)$ is the equilibrium revenue of a c-type firm. As a consequence, the elasticity of c^* is negative (positive) if r_u is increasing (decreasing). Therefore, we have:

Proposition 4 *Regardless of the cost distribution, the cutoff cost decreases with market size if and only if the RLV is increasing. Furthermore, the cutoff cost is independent of the market size if and only if preferences are CES.*

Thus, the number of firms selected when the market gets bigger depends only upon the behavior of the RLV. Intuitively, we expect a larger market to render competition tougher (the RLV increases), which in turn triggers the exit of the least productive firms. However, if a larger market happens to soften competition (the RLV decreases), then less productive firms are able to stay in business.

Zhelobodko et al. (2012) show that both the equilibrium mass of entrants and the mass of operating firms increase with L when the RLV increases. The same authors also establish that the equilibrium consumption $x^*(c)$ decreases with L for all $c < c^*$. Therefore, when the RLV is increasing, (5.55) implies that the equilibrium price $p^*(c)$ decreases for all the c-type firms that remain in business. Hence, when preferences are additive, prices move in the same direction in response to a market size shock whether firms are homogeneous or heterogeneous.

Note, finally, that even when preferences generate pro-competitive effects ($\mathcal{E}_x(\sigma) < 0$), the selection of firms associated with a bigger market may lead to a drop in aggregate productivity because the more productive firms need not gain more demand than the less productive firms. To be precise, Bertoletti and Epifani (2014) show that a hike in L may have a negative impact on the aggregate productivity if the elasticity of the marginal revenue is decreasing in q while the elasticity of substitution is decreasing in x.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

5.2 Linear-quadratic Preferences

Melitz and Ottaviano (2008) propose an alternative approach based on utility (5.27). Because such preferences generate linear demands that feature a finite choke-price α (see Section 3.2), there is no need to assume that firms face a positive fixed production cost ($F = 0$). In this case, total profits are generally different from zero. However, how profits/loses are shared does not matter because the upper-tier utility is linear.

Firm i operates if the demand for its variety is positive, that is,

$$p_i \leq p_{\max} \equiv \frac{\beta\alpha + \gamma N\bar{p}}{\beta + \gamma N}, \tag{5.57}$$

holds, where $\bar{p}$ is the average market price given by

$$\bar{p} \equiv \frac{1}{N}\int_0^N p_i \mathrm{d}i.$$

The market demand for a variety i is such that

$$q_i = \frac{L}{\beta}\left(p_{\max} - p_i\right). \tag{5.58}$$

Unlike the CES, the price elasticity of the demand for variety i is variable and equal to

$$\mathcal{E}_{p_i}(x_i) = \frac{p_i}{p_{\max} - p_i}. \tag{5.59}$$

The expressions (5.57) and (5.59) imply that the demand elasticity increases with its own price p_i, decreases with the average market price $\bar{p}$ (because varieties are substitutes), and increases when more firms are active.

Using (5.58), it is readily verified that the profit-maximizing price $p^*(c)$ set by a c-type firm must satisfy

$$p^*(c) = \frac{p_{\max} + c}{2}, \tag{5.60}$$

which boils down to (5.31) when firms are homogeneous.

Assume that the support of cost distribution Γ is a compact interval $[0, c_M]$, where c_M is large. The cut-off cost $\bar{c} \in [0, c_M]$ satisfies $p^*(c) = c$, that is, the least productive operating firm earns zero profits and prices at the marginal cost. Combining the cut-off condition $p^*(c) = c$ with (5.60) yields

$$\bar{c} = p_{\max},$$

so that the equilibrium price $p^*(c)$, output $q^*(c)$ and profits $\pi^*(c)$ of a c-type firm are given by:

$$p^*(c) = \frac{\bar{c} + c}{2}, \qquad q^*(c) = \frac{L}{\beta}\frac{\bar{c} - c}{2}, \tag{5.61}$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

$$\pi^*(c) \equiv \left[p^*(c) - c\right] q^*(c) = \frac{L}{4\beta}(\bar{c} - c)^2. \tag{5.62}$$

By implication, *firms with a higher productivity have more monopoly power and higher profits*.

It is well known that linear demands allow for a simple relationship between the variances of prices and marginal costs:

$$\mathbf{V}(p) = \frac{1}{4}\mathbf{V}(c),$$

which means *prices are less dispersed than marginal costs*. This result complements the discussion on incomplete pass-through in Section 4.2.

It remains to pin down $\bar{c}$, which is given using the ZPC:

$$\int_0^{\bar{c}} \bar{\pi}(c)\mathrm{d}\Gamma(c) = F_e,$$

where $F_e > 0$ is the sunk entry cost. Using (5.62), we restate this condition as follows:

$$\frac{L}{4\beta}\int_0^{\bar{c}} (\bar{c} - c)^2 \mathrm{d}\Gamma(c) = F_e. \tag{5.63}$$

The left-hand side of this expression is an increasing function of $\bar{c}$, which implies that (5.63) has a unique solution c^*. This solution is interior ($0 < c^* < c_M$) if and only if

$$\mathbf{E}\left[(c_M - c)^2\right] > \frac{4\beta}{L}F_e$$

holds. Therefore, when the population L is very small, all firms choose to produce. Otherwise, as implied by (5.63), a hike in L drives c^* downwards, which confirms the idea that firms do not pass onto consumers the entire fall in cost (see Proposition 3). In other words, *a larger market skews the distribution of sales toward the varieties that are more efficiently produced*.

Finally, (5.57) and (5.61) can be used to pin down the mass of active firms:

$$N^* = \frac{2\beta}{\gamma}\frac{\alpha - c^*}{c^* - \mathbf{E}(c|c \le c^*)},$$

which indicates a decreasing relationship between c^* and N^*. In particular, an increase in L invites more firms to enter, even though a larger market pushes the least productive firms out of business.[7]

[7] Behrens et al. (2014) undertake a similar exercise within a full-fledged general equilibrium model with CARA preferences and income effects.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

5.3 VES Preferences

Working with general symmetric preferences and heterogeneous firms is tricky. Assume that the cutoff cost $\bar{c}$ and the number of entrants are given, so that the mass of active firms is determined. Unlike the CES, the equilibrium consumption of a given variety depends on the consumption levels of the other varieties. Hence, markets are independent across varieties. Unlike additive preferences, competition among firms is no longer described by an aggregative game. Unlike LQ preferences, a closed-form solution is not available. All of this implies that the way firms choose their output is through a non-atomic game with asymmetric players, which cannot be solved point-wise. Such an equilibrium $\bar{\mathbf{x}}(\bar{c}, N_e)$, which need not be unique, can be shown to exist if, when a non-zero measure set of firms raise their prices, it is profit-maximizing for the other firms to increase their prices, as in oligopoly games where prices are strategic complements (Parenti et al., 2017). The corresponding free-entry equilibrium is thus defined by a pair $\left(\bar{c}^*, N_e^*\right)$ that satisfies the zero-expected-profit condition for each firm:

$$\int_0^{\bar{c}} \left[\bar{\pi}_c(\bar{c}, N_e) - F\right] \mathrm{d}\Gamma(c) = F_e, \tag{5.64}$$

and the cutoff condition:

$$\bar{\pi}_{\bar{c}}(\bar{c}, N_e) = F. \tag{5.65}$$

Thus, regardless of the nature of preferences and the distribution of marginal costs, the heterogeneity of firms amounts to replacing N by $\bar{c}$ and N_e because $N = \Gamma(\bar{c})N_e$. As a consequence, *the complexity of the problem increases from one to two dimensions*.

Dividing (5.64) by (5.65) yields the following new equilibrium condition:

$$\int_0^{\bar{c}} \left[\frac{\bar{\pi}_c(\bar{c}, N_e)}{\bar{\pi}_{\bar{c}}(\bar{c}, N_e)} - 1\right] \mathrm{d}\Gamma(c) = \frac{F_e}{F}. \tag{5.66}$$

When firms are symmetric, we have seen that the sign of $\mathcal{E}_N(\sigma)$ plays a critical role in comparative statics. Since firms of a given type are symmetric, the same holds here. The difference is that the mass of operating firms is determined by the two endogenous variables $\bar{c}$ and N_e. As a consequence, understanding how the mass of active firms responds to a population hike requires studying the way the left-hand side of (5.66) varies with $\bar{c}$ and N_e. Let $\sigma_c(\bar{c}, N_e)$ be the equilibrium value of the elasticity of substitution between any two varieties supplied by c-type firms:

$$\sigma_c(\bar{c}, N_e) \equiv \bar{\sigma}\left[\bar{x}_c(\bar{c}, N_e), \bar{\mathbf{x}}(\bar{c}, N_e)\right].$$

In this case, we may rewrite $\bar{\pi}_c(\bar{c}, N_e)$ as follows:

$$\bar{\pi}_c(\bar{c}, N_e) = \frac{c}{\sigma_c(\bar{c}, N_e) - 1} L\bar{x}_c(\bar{c}, N_e),$$

which is the counter-part of (5.40), while the markup of a c-type firm is given by

$$m_c^*(\hat{c}, N_e) = \frac{1}{\sigma_c(\hat{c}, N_e)}.$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Hence, the elasticity of substitution can be used for studying heterogeneous firms at the cost of one additional dimension, i.e., the firm's type c. Following this approach, Parenti et al. (2017) prove the following result:

Proposition 5 *Assume that $\bar{\pi}_c(\bar{c}, N_e)$ decreases with $\bar{c}$ and N_e for all c. Then, the equilibrium mass of entrants increases with L. Furthermore, the equilibrium cutoff decreases with L when $\sigma_c(\bar{c}, N_e)$ increases with $\bar{c}$ and N_e, whereas it increases with L when $\sigma_c(\bar{c}, N_e)$ increases with $\bar{c}$ but decreases with N_e.*

Given $\bar{c}$, the number of operating firms is proportional to the number of entrants. Therefore, assuming that $\sigma_c(\bar{c}, N_e)$ increases with $\bar{c}$ and N_e may be considered as the counterpart of the condition $\mathcal{E}_N(\sigma) > 0$ discussed in subsection 5.3. In response to an increase in L, the two effects combine to induce the exit of the least efficient active firms. However, Proposition 5 also shows that predicting the direction of firms' selection is generally problematic.

6 EQUILIBRIUM VERSUS OPTIMUM PRODUCT DIVERSITY: A VARIETY OF RESULTS

Conventional wisdom holds that entry is desirable because it often triggers more competition and enhances social efficiency. However, when the entry of new firms involves additional fixed costs, the case for entry is less clear-cut. What is more, when goods are differentiated the extent of diversity comes into play. In this context, the following question arises: does the market provide too many or too few varieties?

Spence (1976) casts doubt on the possibility of coming up with a clear-cut answer to this question because two opposite forces are at work. First, the entrant disregards the negative impact its decision has on the incumbents by taking away from them some of their customers (the "business-stealing" effect). This effect pushes toward excessive diversity. Second, the entrant is unable to capture the entire social benefit it creates by increasing diversity because it does not price discriminate across consumers (the "incomplete appropriability" effect). This pushes toward insufficient diversity. As a consequence, the comparison between the market and optimal outcomes is likely to depend on the particular framework we believe to be a good representation of differentiated markets. Conventional wisdom holds that *spatial models foster excessive diversity, whereas the market and optimal outcomes do not differ much in the case of symmetric preferences*. The reason for this difference is that a firm has few neighboring rivals in spatial models, which facilitates entry. On the contrary, when competition is global, the entrant must compete with many rivals, which dampens entry.

6.1 Additive Preferences

The social planner aims to find the mass of firms and the consumption profile that maximize the common utility level and meet the labor balance constraint:

$$\max_{(\mathbf{x}, N)} U(\mathbf{x}) \qquad \text{s.t.} \quad cL \int_0^N x_i \,\mathrm{d}i + NF = L. \tag{5.67}$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Using additivity and symmetry, this program may rewritten as follows:

$$\max_{(q,N)} Nu(x) \qquad \text{s.t.} \quad N = \frac{L}{cLx + F},$$

which can be reduced to maximizing

$$\frac{Lu(x)}{cLx + F}$$

with respect to x. Applying the FOC yields

$$\mathcal{E}_x(u) = \frac{cLx}{cLx + F}. \tag{5.68}$$

Using the equilibrium condition (5.45), we obtain

$$1 - r_u(x) = \frac{cLx}{cLx + F}. \tag{5.69}$$

These two expressions show that firms care about consumers' marginal utility (see (5.69)), which determines the inverse demands, whereas the planner cares about consumers' utility (see (5.68)).

The equilibrium outcome is optimal for any L, c and F if and only if the utility $u(\cdot)$ satisfies both (5.68) and (5.69), that is, solves the following differential equation:

$$r_u(x) + \mathcal{E}_x(u) - 1 = 0. \tag{5.70}$$

Can this condition be given a simple economic interpretation? Let λ be the social value of labor, that is, the Lagrange multiplier of the social planner. Therefore, it must be that $u'(x) = \lambda cx$, so that

$$1 - \mathcal{E}_x(u) = \frac{u(x) - u'(x)x}{u(x)} = \frac{u(x) - \lambda cx}{u(x)},$$

and thus $1 - \mathcal{E}_x(u)$ may be interpreted as the "social markup" of a variety (Vives, 1999). Since $r_u(x)$ is a firm's markup, (5.70) means that *the market and social outcomes coincide if and only if the private and social markups are identical at the equilibrium consumption*.

It is readily verified that, up to an affine transformation, $u(x) = x^\rho$ is the only solution to (5.70). Furthermore, labor balance implies that each firm produces the optimal quantity. Accordingly, when firms are symmetric *the CES is the only model with additive preferences under which the market outcome is socially optimal*. Intuitively, under the CES everything works *as if* firms' marginal cost were $c\sigma/(\sigma - 1) > c$, while the market price equals $c\sigma/(\sigma - 1)$ in an otherwise perfectly competitive market. Under these circumstances, the amount $(p^* - c)q^* = cq^*/(\sigma - 1) > 0$ must be interpreted as a transfer from consumers to firms, which allows firms to exactly cover their fixed costs. Labor market clearing pins down the mass of firms, which is optimal because the surplus $(p^* - c)q^*$ generated by an

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

additional variety is equal to its launching cost F. As a consequence, the market equilibrium coincides with the socially optimal outcome. Dhingra and Morrow (2018) extend this result to heterogeneous firms. But how robust are these interesting optimality properties?

Comparing (5.68) with (5.69) implies that the market delivers excessive variety if and only if the private markup exceeds the social markup at the equilibrium consumption level:

$$\mathcal{E}_x(u)|_{x=x^*} > 1 - r_u(x^*). \tag{5.71}$$

For example, under the CARA utility, (5.71) may be written as follows:

$$\frac{\alpha x^*}{1 - \alpha x^*} > \exp(\alpha x^*) - 1.$$

Applying Taylor expansion to both sides of this expression yields

$$\sum_{k=1}^{\infty} (\alpha x)^k > \sum_{k=1}^{\infty} \frac{1}{k} (\alpha x)^k ,$$

which holds for any positive value of x. As a consequence, under the CARA, the market provides too many varieties while firms' output is too small. In addition, Behrens et al. (2016) show that, when firms are heterogeneous, the more productive firms under-produce, whereas the less productive firms over-produce, and thus the average productivity at the equilibrium is lower than at the social optimum.

Since

$$\frac{\mathrm{d}}{\mathrm{d}x} [1 - \mathcal{E}_x(u)] = \frac{\mathcal{E}_x(u)}{x} [r_u(x) - (1 - \mathcal{E}_x(u))] ,$$

there is always excessive diversity, hence firms' output is too small, if and only if $\mathcal{E}_x(u)$ is decreasing. Equivalently, there is always insufficient diversity, hence firms' outputs are too large, if and only if $\mathcal{E}_x(u)$ is increasing. For example, under the preferences given by $u(x) = (x + \alpha)^\rho$, there are too few (too many) varieties in equilibrium if $\alpha > 0$ ($\alpha < 0$).

Furthermore, in a multi-sector economy where firms are heterogeneous, the upper-tier utility is Cobb-Douglas, while each sub-utility is CES, the equilibrium and optimum coincide if and only if the elasticity of substitution is the same across sectors. Otherwise, too much labor is allocated to the more competitive sectors (Behrens et al., 2016). These results point to the lack of robustness of the CES welfare properties, which may lead to strong biases in policy assessment.

6.2 Homothetic Preferences

Since homothetic preferences are also widely used in applications, it is legitimate to ask how the above results change when preferences are homothetic. Without loss of generality, we assume that U is homogeneous of degree one in $\mathbf{x}$. In the case of symmetric consumption profiles $\mathbf{x} = xI_{[0,N]}$, we have

$$U\left(xI_{[0,N]}\right) \equiv \phi(N, x) = X\psi(N),$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

where $\psi(N) \equiv \phi(N,1)/N$ and $X \equiv xN$. The ratio of the first-order conditions is given by

$$X\frac{\psi'(N)}{\psi(N)} = \frac{F}{cL},$$

which is equivalent to

$$\mathcal{E}_{\psi}(N) \equiv N\frac{\psi'(N)}{\psi(N)} = \frac{F}{cLx}.$$

As for the market equilibrium condition (5.41) can be reformulated as follows:

$$\frac{\bar{m}(N)}{1-\bar{m}(N)} = \frac{F}{cLx}.$$

The social optimum and the market equilibrium are identical if and only if

$$\mathcal{E}_{\psi}(N) = \frac{\bar{m}(N)}{1-\bar{m}(N)}, \tag{5.72}$$

while there is excess (insufficient) variety if and only if the right-hand side term of (5.72) is larger (smaller) than the left-hand side term.

Given $\phi(X,N)$, it is reasonable to map this function into another homothetic preference $\mathbb{A}(N)\phi(X,N)$, where $\mathbb{A}(N)$ is a *shifter* that depends only on N. Observe that the utility $\mathbb{A}(N)U(\mathbf{x})$ is homothetic and generates the same equilibrium outcome as $U(\mathbf{x})$, for the elasticity of substitution $\sigma(N)$ is unaffected by introducing the shifter $\mathbb{A}(N)$. To determine the shifter $\mathbb{A}(N)$, (5.72) is to be rewritten as follows in the case of $\mathbb{A}(N)\phi(X,N)$:

$$\mathcal{E}_{\mathbb{A}}(N) + \mathcal{E}_{\psi}(N) = \frac{m(N)}{1-m(N)}. \tag{5.73}$$

For this expression to hold, $\mathbb{A}(N)$ must be the solution to the linear differential equation in N

$$\frac{\mathrm{d}A}{\mathrm{d}N} = \left[\frac{m(N)}{1-m(N)} - \frac{N}{\psi(N)}\frac{\mathrm{d}\psi}{\mathrm{d}N}\right]\frac{A(N)}{N},$$

which has a unique solution up to a positive constant. Therefore, *there always exists a shifter $A(N)$ such that* (5.73) *holds for all N if and only if $U(\mathbf{x})$ is replaced with $A(N)U(\mathbf{x})$*. The shifter aligns the optimum to the equilibrium, which remains the same. Furthermore, it is readily verified that there is excess (insufficient) variety if and only if the right-hand side term of (5.73) is larger (smaller) than the left-hand side term. Thus, even when one restricts oneself to homothetic and symmetric preferences, there is, a priori, no reason to expect a robust result to emerge.

In sum, care is needed, for the choice of (additive or homothetic) preferences is likely to affect the nature of the prescriptions based on quantitative models of monopolistic competition. In particular, *CES preferences, which occupy center stage in the growing flow of quantitative models, must be used with care*.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

7 CONCLUDING REMARKS

Accounting for oligopolistic competition in general equilibrium theory remains a worthy goal rather than an actual achievement. This is why many scholars have embraced the CES model of monopolistic competition. Although this model has great merits, it leads to knife-edge results or to findings that clash with fundamental principles of microeconomics and industrial organization. In addition, recent empirical evidence pointing out the shortcomings of the CES is growing fast. This does not mean, however, that we need a totally new framework; the emphasis on the elasticity of substitution is warranted when recognizing that it is variable, rather than constant. By mimicking the behavior of oligopolistic markets, the VES model of monopolistic competition offers an alternative solution to some of the difficulties uncovered by Gabszewicz and Vial (1972) in their pioneering work on imperfect competition in general equilibrium. In this respect, we have built a link between two bodies of the literature on imperfect markets that were perceived so far as totally separate.

Despite real progress, it should be clear that there is scope for more work. We provide here a short list of some major issues that should rank high on the research agenda. First, papers coping with several sectors typically assume a Cobb-Douglas upper-tier utility and CES lower-tier sub-utilities. Such a specification of preferences leaves much to desire as it does not allow for a genuine interaction across sectors because the income share spent on each product is given a priori. Behrens et al. (2016) is a worthy exception that should trigger new contributions.

Second, the demand side of existing models of monopolistic competition relies on the assumption of symmetric preferences, while heterogeneity is introduced on the supply side only. Yet, the recent empirical evidence gathered by Hottman, Redding and Weinstein (2016) find that 50 to 75 percent of the variance in US firm size can be attributed to differences in what these authors call "firms' appeal," that is, the demand side, and less than 20 percent to average marginal cost differences. As a consequence, one may safely conclude that it is time to pay more attention to the demand side in monopolistic competition theory. Besides the VES model, another step in this direction has been made by Di Comite, Thisse and Vandenbussche (2014) who embed taste heterogeneity into the LQ model. Absent a specific taste demand parameter, the model with heterogeneous costs and quality only explains 55 percent of the quantity variation in Belgian exports. Allowing for taste differences generates asymmetry in demand across countries and offers a rationale for the missing variability in sales.

Last, one may wonder what "heterogeneous firms" actually mean in a world where, despite a large number of producers, a handful of firms account for a very high share of total sales. There are at least two different modeling strategies that can be used to tackle this question. Ever since Melitz (2003), the first approach with firms operating under monopolistic competition but facing different marginal costs is dominant. However, as observed by Neary (2011), firms in this approach differ in types, not in kind, as all firms remain negligible to the market. A second line of research, developed by Shimomura and Thisse (2012), combines a continuum of negligible (non-atomic) players and a few large (atomic) players who are able to manipulate the market. Hence, firms now differ in kind. This leads to a hybrid market structure blending the features of oligopoly and monopolistic competition. Despite its empirical relevance, this approach has attracted little attention in the profession.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

REFERENCES

Allanson, P. and C. Montagna (2005) Multiproduct firms and market structure: An explorative application to the product life cycle. *International Journal of Industrial Organization* 23: 587–97.

Amity, M., O. Itskhoki and J. Konings (2016) International shocks and domestic prices: How large are strategic complementarities? *NBER Working Paper No. 22119*.

Anderson, S.P. and A. de Palma (2000) From local to global competition. *European Economic Review* 44: 423–48.

Anderson, S.P., A. de Palma and Y. Nesterov (1995) Oligopolistic competition and the optimal provision of products. *Econometrica* 63: 1281–302.

Anderson, S.P., A. de Palma and J.-F. Thisse (1992) *Discrete Choice Theory of Product Differentiation*. Cambridge, MA: MIT Press.

Anderson, S.P., N. Erkal and D. Piccinin (2015) Aggregative oligopoly games with entry. University of Virginia, memo.

Aumann, R.J. (1964) Markets with a continuum of traders. *Econometrica* 32: 39–50.

Beckmann, M.J. (1972) Spatial Cournot oligopoly. *Papers and Proceedings of the Regional Science Association* 28: 37–47.

Behrens, K. and Y. Murata (2007) General equilibrium models of monopolistic competition: A new approach. *Journal of Economic Theory* 136: 776–87.

Behrens, K., G. Mion, Y. Murata and J. Südekum (2014) Trade, wages, and productivity. *International Economic Review* 55: 1305–48.

Behrens, K., G. Mion, Y. Murata and J. Südekum (2016) Distorted monopolistic competition. Université du Québec à Montréal, mimeo.

Belleflamme, P., P. Picard and J.-F. Thisse (2000) An economic theory of regional clusters. *Journal of Urban Economics* 48: 158–84.

Bernard, A.B., S.J. Redding and P.K. Schott (2010) Multiple-product firms and product switching. *American Economic Review* 100: 70–97.

Bernard, A.B., S.J. Redding and P.K. Schott (2011) Multi-product firms and trade liberalization. *Quarterly Journal of Economics* 126: 1271–318.

Bertoletti, P. and P. Epifani (2014) Monopolistic competition: CES redux? *Journal of International Economics* 93: 227–38.

Bertoletti, P. and F. Etro (2017) Monopolistic competition when income matters. *Economic Journal* 127: 1217–43.

Bilbiie, F., F. Ghironi and M. Melitz (2012) Endogenous entry, product variety, and business cycles. *Journal of Political Economy* 120: 304–45.

Chamberlin, E. (1933) *The Theory of Monopolistic Competition*. Cambridge, MA: Harvard University Press.

Chen, Y. and M.H. Riordan (2007) Price and variety in the spokes model. *Economic Journal* 117: 897–921.

Chen, Y. and M.H. Riordan (2008) Price-increasing competition. *Rand Journal of Economics* 39, 1042–58.

d'Aspremont, C., R. Dos Santos Ferreira and L.-A. Gérard-Varet (1996) On the Dixit-Stiglitz model of monopolistic competition. *American Economic Review* 86: 623–29.

De Loecker, J. and P.K. Goldberg (2014) Firm performance in a global economy. *Annual Review of Economics* 6: 201–27.

De Loecker, J., P.K. Goldberg, A.K. Khandelwal and N. Pavcnik (2016) Prices, markups and trade reform. *Econometrica* 84: 445–510.

Deneckere, R. and M. Rotschild (1992) Monopolistic competition and preference diversity. *Review of Economic Studies* 59: 361–73.

Dhingra, S. (2013) Trading away wide brands for cheap brands. *American Economic Review* 103: 2554–84.

Dhingra, S. and J. Morrow (2018) Monopolistic competition and optimum product diversity under firm heterogeneity. *Journal of Political Economy*, forthcoming.

Di Comite, F., J.-F. Thisse and H. Vandenbussche (2014) Verti-zontal differentiation in export markets. *Journal of International Economics* 93: 50–66.

Dixit, A. and J.E. Stiglitz (1977) Monopolistic competition and optimum product diversity. *American Economic Review* 67: 297–308.

Eaton, B.C. and R.G. Lipsey (1977) The introduction of space into the neoclassical model of value theory. In: M. Artis and A. Nobay (eds.), *Studies in Modern Economics*. Oxford: Basil Blackwell, pp. 59–96.

Eaton, B.C. and N. Schmitt (1994) Flexible manufacturing and market structure. *American Economic Review* 84: 875–88.

Feenstra, R.C. (2003) A homothetic utility function for monopolistic competition models, without constant price elasticity. *Economics Letters* 78: 79–86.

Feenstra, R.C. and D. Weinstein (2017) Globalization, markups, and U.S. welfare. *Journal of Political Economy* 125(4): 1040–74.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Foster, L., J.C. Haltiwanger and C. Syverson (2008) Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review* 98: 394–425.

Gabaix, X., D. Laibson, D. Li, H. Li, S. Resnick and C.G. de Vries (2016) The impact of competition on prices with numerous firms. *Journal of Economic Theory* 165: 1–24.

Gabszewicz, J.J. and J.-P. Vial (1972) Oligopoly "a la Cournot" in a general equilibrium analysis. *Journal of Economic Theory* 4: 381–400.

Greenhut, M.L., G. Norman and C.-S. Hung (1987) *The Economics of Imperfect Competition. A Spatial Approach*. Cambridge, UK: Cambridge University Press.

Hart, O.D. (1985a) Imperfect competition in general equilibrium: An overview of recent work. In K.J. Arrow and S. Honkapohja (eds.), *Frontiers in Economics*. Oxford: Basil Blackwell, pp. 100–149.

Hart, O.D. (1985b) Monopolistic competition in the spirit of Chamberlin: A general model. *Review of Economic Studies* LII: 529–46.

Hart, O.D. (1985c) Monopolistic competition in the spirit of Chamberlin: Special results. *Economic Journal* 95: 889–908.

Hotelling, H. (1929) Stability in competition. *Economic Journal* 39: 41–57.

Hottman, C., S.J. Redding and D.E. Weinstein (2016) Quantifying the sources of firm heterogeneity. *Quarterly Journal of Economics* 131: 1291–364.

Jean, S. (2002) International trade and firms' heterogeneity under monopolistic competition. *Open Economies Review* 13: 291–311.

Kaldor, N. (1935) Market imperfection and excess capacity. *Economica* 2: 35–50.

Kugler, M. and E. Verhoogen (2012) Prices, plant size, and product quality. *Review of Economic Studies* 79: 307–39.

Lancaster, K. (1979) *Variety, Equity, and Efficiency*. New York: Columbia University Press.

MacLeod, W.B., G. Norman and J.-F. Thisse (1988) Price discrimination and equilibrium in monopolistic competition. *International Journal of Industrial Organization* 6: 429–46.

Mas-Colell, A. (1984) On the theory of perfect competition. The 1984 Nancy L. Schwartz Memorial Lecture, Northwestern University.

Matsuyama, K. (1995) Complementarities and cumulative process in models of monopolistic competition. *Journal of Economic Literature* 33: 701–29.

Mayer, T., M. Melitz and G.I.P. Ottaviano (2014) Market size, competition, and the product mix of exporters. *American Economic Review* 104: 495–536.

Mayer, T., M. Melitz and G.I.P. Ottaviano (2016) Product mix and firm productivity responses to trade competition. *CEPR Discussion Paper No. 11389*.

McFadden, D. (1974) Conditional logit analysis of qualitative choice behavior. In: P. Zarembka (ed.), *Frontiers in Econometrics*. New York: Academic Press, pp. 105–42.

Melitz, M.J. (2003) The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica* 71: 1695–725.

Melitz, M.J. and G.I.P. Ottaviano (2008) Market size, trade, and productivity. *Review of Economic Studies* 75: 295–316.

Neary, P. (2010) Two and a half theories of trade. *The World Economy* 33: 1–19.

Nocke, V. (2006) A gap for me: Entrepreneurs and entry. *Journal of the European Economic Association* 4: 929–55.

Norman, G. and J.-F. Thisse (1999) Technology choice and market structure: Strategic aspects of flexible manufacturing. *Journal of Industrial Economics* 47: 345–72.

Ottaviano, G.I.P. and J.-F. Thisse (2004) Agglomeration and economic geography. In J.V. Henderson and J.-F. Thisse (eds.), *Handbook of Regional and Urban Economics, Volume 4*. Amsterdam: North-Holland, pp. 2563–608.

Ottaviano, G.I.P., T. Tabuchi and J.-F. Thisse (2002) Agglomeration and trade revisited. *International Economic Review* 43: 409–36.

Parenti, M., P. Ushchev and J.-F. Thisse (2017) Toward a theory of monopolistic competition. *Journal of Economic Theory* 167: 86–115.

Perloff, J.M. and S.C. Salop (1985) Equilibrium with product differentiation. *Review of Economic Studies* 52: 107–20.

Redding, S.J. (2011) Theories of heterogeneous firms and trade. *Annual Review of Economics* 3: 77–105.

Salop, S.C. (1979) Monopolistic competition with outside goods. *Bell Journal of Economics* 10: 141–56.

Sattinger, M. (1984) Value of an additional firm in monopolistic competition. *Review of Economic Studies* 51: 321–32.

Shimomura, K.-I. and J.-F. Thisse (2012) Competition among the big and the small, *RAND Journal of Economics* 43: 329–47.

Simonovska, I. (2015) Income differences and prices of tradables: Insights from an online retailer. *Review of Economic Studies* 82: 1612–56.

Spence, M. (1976) Product selection, fixed costs, and monopolistic competition. *Review of Economic Studies* 43: 217–35.

Stigler, G. (1969) *The Organization of Industry*. Chicago, IL: The University of Chicago Press.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Triffin, R. (1940) *Monopolistic Competition and General Equilibrium Theory*. Cambridge, MA: Harvard University Press.
Ushchev, P. and Y. Zenou (2016) Price competitition in product variety networks. *CEPR Discussion Paper No. 10862*.
Vickrey, W.S. (1964) *Microstatics*, New York: Harcourt, Brace & World.
Vives, X. (1985) On the efficiency of Bertrand and Cournot equilibria with product differentiation. *Journal of Economic Theory* 36: 166–75.
Vives, X. (1999) *Oligopoly Pricing. Old Ideas and New Tools*. Cambridge, MA: MIT Press.
Vogel, J. (2011) Spatial price discrimination with heterogeneous firms. *Journal of Industrial Economics* LIX: 661–76.
Zhelobodko, E., S. Kokovin, M. Parenti and J.-F. Thisse (2012) Monopolistic competition: Beyond the constant elasticity of substitution. *Econometrica* 80: 2765–84.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

APPENDIX

That $\mathcal{E}_{p_i}(P) > 0$ is straightforward because $P(\mathbf{p})$ is increasing in any p_i. To show that $\mathcal{E}_{p_i}(P) < 1$ also holds, observe that $P(\mathbf{p})$ satisfies $P(t\,\mathbf{p}) < tP(\mathbf{p})$ for any $t > 1$. Indeed, (5.6) implies that for any given $\mathbf{p}$ and any given $t > 1$, $P_1 \equiv P(t\mathbf{p})$ and $P_2 \equiv tP(\mathbf{p})$ must solve, respectively, the following equations:

$$P_1 = \varphi'\left[\sum_{j=1}^{n} u\left(\xi\left(\frac{tp_j}{P_1}\right)\right)\right], \qquad P_2 = t\varphi'\left[\sum_{j=1}^{n} u\left(\xi\left(\frac{tp_j}{P_2}\right)\right)\right].$$

As $t > 1$, the right-hand side of the second equation is greater than the right-hand side of the first equation. Therefore, $P_1 < P_2$, that is, $P(t\,\mathbf{p}) < tP(\mathbf{p})$. This, in turn, implies

$$\mathcal{E}_t\left[P(t\,\mathbf{p})\right]\Big|_{t=1} < 1,$$

or, equivalently,

$$\sum_{i=1}^{n} \mathcal{E}_{p_i}(P) < 1. \tag{A.1}$$

Since $\mathcal{E}_{p_i}(P) > 0$, it must be that $\mathcal{E}_{p_i}(P) < 1$. Q.E.D.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

6. Oligopoly and product differentiation

Jean J. Gabszewicz and Ornella Tarola

1 INTRODUCTION

Several reasons can explain why the variants of many products are so numerous in developed economies. The main reason is certainly related to consumers' heterogeneity. Not only are their preferences and tastes different among them, but also their income. Imagine on the other hand, a completely homogeneous world in which all consumers were identical. There would be no reason to make different cars, with different colors; all houses would be built using the same style. There would be no luxuries, but only standard variants of the products since all consumers would have the same income. The world would be uniform, deprived of any diversity. Clearly, taste and income heterogeneity constitute the main factors explaining variants' proliferation.

As for the firms, one could think *a priori* that there exist counterincentives to multiply the number of variants of a product. In particular, it seems that producing two different variants of the same product is more expensive than producing a single one and duplicating it. Nevertheless, this is not always the case. Consider cars of different colors: a red car of some given type is not more expensive than a blue one of the same type while there certainly exist consumers who prefer a red car to a blue one. Accordingly, it can be more advantageous for the firm to produce both models, then all consumers can be supplied, while some of them would probably not be willing to buy the car if it does not correspond to their preferred color. In some cases, it could even be less costly to produce different variants than simply reproducing the original one. This would be the case, for instance, when a flexible technology is able to produce different variants, allowing the firm to reach a production level large enough to benefit from increasing returns to scale and satisfy the higher demand resulting from this very model diversity. Then the working stock is more rapidly deadened and the unit cost is accordingly reduced. In such cases, the firm wishes to increase the diversity of its supply to take full advantage of economies of scope.

Yet, there is a deeper reason why firms may wish to differentiate their products from those supplied by their rivals. When firms sell close substitutes, a small price differential is sufficient in order to attract all consumers to the least expensive variant. As a consequence, one should expect exacerbated competition among firms to capture demand, possibly leading to a price war, and devastating the profits realized by all firms in the industry. On the other hand, when proposing variants not "too" close to each other, firms can soften price competition. When the rival firm decreases its price, it succeeds in attracting only a small fringe of the competitor's customers. Each firm then benefits from a local monopoly position. Thus, all economic agents – consumers as well as firms – have strong incentives to favour product differentiation.

A competitive market is often idealized as a central place where many sellers and buyers meet in order to exchange units of a homogeneous product. From this interaction a unique price emerges, the competitive price, summarizing all the transactions and arbitrages among

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

them. A typical example of such a market is the stock exchange market. In the real world however, economic activities are generally performed at different points in space, even when they concern the same homogeneous product. The market operates through a complex network of scattered sellers and buyers located in different points of the geographical space. Examples abound of such dispersed activities: supermarkets located in different areas of the urban space, gasoline stations, shops like bakeries, butchers, groceries, to cite only a few. An immediate consequence of this spatial dispersion is to make the key assumption of pure competition inapplicable: a large number of buyers and sellers participating in the exchange of the product. When firms are dispersed over space, each of them finds at most a small number of other firms selling the same product in its immediate environment. Yet more such firms exist in a remoter area, but their influence is weaker due to the transportation cost that consumers would pay in order to buy from them. In the same way, consumers far enough from a particular seller have no incentives to buy from it. The market of a particular seller is thus limited to a specific area: the more restricted this area, the smaller the transportation cost between the consumers and the seller.

The difference between a competitive market and a geographically dispersed market finds its counterpart, in industrial economics, in the difference existing between a homogeneous and a differentiated industry. In a competitive market for a homogeneous good, all variants proposed for sale are all agglutinated on the same point in the space of *characteristics*: they are perfect substitutes.[1] Furthermore, nothing prevents sellers and buyers from being numerous. By contrast, in a differentiated industry, products are dispersed in the space of characteristics. Furthermore, the seller of each variant is almost a monopolist with respect to potential consumers since only those firms selling rather close substitutes could attract these consumers by setting a lower price. These substitute products are thus located in a small neighborhood of the space of characteristics.

The analogy between a geographically dispersed market and a differentiated industry spontaneously invites the analysis of spatial competition, and extends it later, by analogy, to competition among differentiated products. This is the proposal made by Hotelling (1929) in his celebrated article on spatial competition, which we shall analyze later in full. At this stage, we simply examine how the spatial metaphor allows a precise definition of the industry.

Figure 6.1 represents a two-dimensional geographical space. Points identify consumers' locations in this space, while crosses represent the position of the sellers of a perfectly homogeneous product. Consumers and sellers are dispersed over the space. Consumers pay a transportation cost, increasing with the distance they are separated from each seller whenever they want to buy a unit of this homogeneous good. Let the index i identify a specific consumer, and the index j a particular seller. We assume there are m consumers $i,\ i = 1, \ldots, m$ and n sellers $j, j = 1, \ldots, n$. We denote by $t(i, j)$ the transportation cost of consumer i when moving from his home to seller j. Assume that seller j produces the good at a constant unit cost $c(j)$ and that the reservation price of consumer i for buying a unit of the good is equal to $r(i)$. Consumers' choice operates among a finite number of alternatives, equal to $n + 1$ when there are n sellers: to buy one unit of the good from one of the n sellers at the exclusion of the others, or not to buy at all. In a purely competitive market, the price to be paid by each consumer to seller j would be equal to its marginal cost, $c(j)$.

[1] The model of characteristics was invented by Gorman (see Gorman, 1956 and 1980) and popularized by Lancaster (see Lancaster, 1966).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

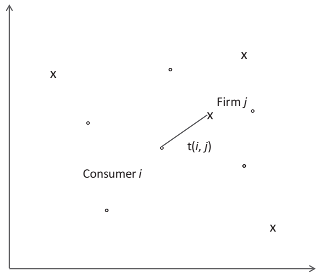


Figure 6.1 *A two-dimensional geographical space*

A necessary condition for firm j to capture consumer i when setting the price $c(j)$ is that the magnitude $c(j) + t(i,j)$ does not exceed the reservation price $r(i)$ corresponding to this consumer. On the contrary, this consumer would refrain from buying the good since his/her reservation price is lower than the effective price augmented with the transportation cost he/she has to incur to travel to seller j. Consequently, all consumers for whom this condition does not hold will never buy from seller j: even at a price leaving a zero profit margin to the seller, they are not willing to buy the product from him. This observation explains the following definitions. The *potential market* of seller j is the set of all consumers i for which the condition

$$c(j) + t(i,j) < r(i)$$

is satisfied. For all consumers in the potential market of seller j, their reservation price is smaller than their transportation cost to seller j plus the production cost $c(j)$: all these consumers are thus potential buyers of firm j if the latter sets a price equal to its marginal cost.

The notion of potential market suggests the intuitive idea of potential competition between two sellers. If there is no consumer in the intersection of their potential markets, it means that, even if these sellers set their price each at their unit cost – the competitive price – there is no consumer who would contemplate buying from either of them at these prices. Thus there is no reason for these sellers to compete with each other since they have no potential customer who is common to both of them. On the other hand, two sellers are *direct potential competitors* when there is at least one consumer who belongs to the intersection of their potential markets. When two firms are direct potential competitors, a price always exists for each of them such that, setting this price, they do not go bankrupt while attracting, however, some customers who would be willing to buy from them.

The notion of direct potential competition is not yet sufficient to totally capture the ingredients of spatial competition. In particular it does not take into account the indirect competition between sellers belonging to the same geographic space, who, while not being direct potential competitors with other firms, indirectly influence their price decisions. This indirect influence transits along a chain formed by a sequence of firms who are, two by two, direct potential competitors. More precisely, two firms are *indirect* potential competitors when (i) they are not direct potential competitors and (ii) there is a chain of sellers between them such that each pair of successive firms in the chain are direct potential competitors. The notion of indirect potential competition captures the idea that each seller in the chain of firms can influence the others' market shares by the reactions it provokes in their strategies

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

when changing its own price. Finally, an *industry* is defined as a group of sellers such that all of them are either direct or indirect competitors of some other firm(s) in the group, and no seller outside the group is a direct or indirect competitor of some firm in the group.

The above conceptual machinery seems at first sight to be a little abstract. It is necessary, however, to define unambiguously when and which firms compete with each other in the context of spatial competition. In other words, it is important to define precisely where an industry starts and where it finishes. The following examples illustrate how to apply the above concepts to more specific environments. As a first example, consider a geographical space made of a road connecting five villages located at points *A, B, C, E* and *F,* separated from each other by a distance equal to d, except between C and E where the distance is equal to $2d$ (see Figure 6.2). A bakery is located in each of these villages, and produces bread at a unit cost c. Each inhabitant in every village eats bread every day, obtaining a resulting utility level equal to s from this consumption. Furthermore, they pay a transportation cost equal to zero when buying the bread in their own village while they pay a cost equal to d when buying the bread in an adjacent village. Finally, we suppose that $c + d < s < c + 2d$. As a consequence, if a baker located in the group of villages (A, B, C) or (E, F) sells the bread at a price $p = c$, he/she can attract inhabitants of an adjacent village to buy bread in his/her shop since the price paid plus the transportation cost is smaller than their willingness to pay s. However, this baker cannot succeed in attracting inhabitants located in more remote villages. In particular, the baker located in village C has no hope of attracting the inhabitants of village E because, even when selling at price $p = c$, the transportation cost $2d$ plus the price c exceeds their willingness to pay (remember that we have assumed $s < c + 2d$). Similarly, the bakers in the group (E, F) can attract the customers of the next village's baker, but the baker in E cannot attract those located in village C.

In this example there are in fact two industries: the bakeries located in the group of villages (A, B, C) and those in the group (E, F). The baker in village A is a direct potential competitor of the baker located in *B,* and reciprocally: at price $p = c$, consumers in B (resp. A) could consider buying their bread from the baker in village A (resp. B) since their willingness to pay s exceeds $c + d$. Similarly, and for the same reason, bakeries located in B and C are direct potential competitors. Furthermore, even bakers located in A and C are not direct potential competitors (since $s < c + 2d$); they are, however, indirect ones via the bakery located in village B: the baker in A is a direct potential competitor of the baker in B who, in turn, is a direct potential competitor of the baker in C. Thus there is a chain of bakeries connecting bakeries A and C such that any link in the chain is made up of direct potential competitors. The group of bakeries (A, B, C) thus corresponds to the definition of an industry given above. As for the group (E, F), the two bakeries are mutually direct potential competitors. Since there is no other bakery that is a direct potential competitor of one of these firms, the group (E, F) also satisfies the conditions required to constitute an industry. In order to introduce a second example, consider the geographical space made up of two roads crossing each other

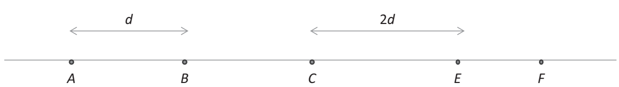


Figure 6.2 An example of spatial competition

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

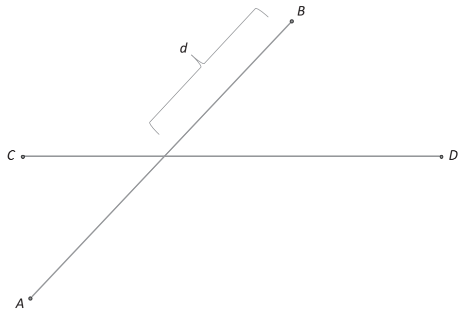


Figure 6.3 *A star network*

and whose extremities are occupied each by a village and a grocery selling the same product. This star network is represented in Figure 6.3.

Suppose the distance separating each village from the crossing point of the two roads is equal to d, also representing the transportation cost from each village to the crossing point. Again suppose that the willingness to pay for the product is equal to s and that the inequalities $c + d < s < c + 2d$ hold, with c representing the unit cost of producing the good. Under these conditions, each of the sellers is a pure monopolist in its own village since no consumer belonging to another village would consider buying from him at $p = c$ and, *a fortiori*, at any price strictly higher than c. We have to conclude that four industries exist in this example, each consisting of a single firm. Now imagine that a new grocery is installed at the crossing point of the two roads and that it also sells the product at a price p equal to c. Then the existing market conditions are completely upset by this new entry since this entry now makes each grocery a direct potential competitor of at least another one (the entrant is now a direct potential competitor of all the groceries existing before entry while each existing grocery is now a direct potential competitor of the entrant).

The concept of industry introduced above is based on the microeconomic characteristics of the agents, firms and consumers. More precisely, it takes as parameters (i) the geographical distribution of consumers, their number, as well as their willingness to pay and their transportation costs; (ii) the location of firms, their number, as well as their unit production costs. As revealed by the last example, changing the value of these parameters is likely to deeply affect the nature of competition among rival firms. Yet the entry of a new firm can entail a brutal change in the competitive environment. Similarly, a change in the distribution of firms in the space of characteristics is likely to modify the structure of potential competition among them.

Finally, it is worth noting that extending the analysis from the spatial metaphor to product differentiation is almost immediate. In the case of differentiated products, the potential market of firm j producing a variant whose cost is equal to $c(j)$ is made up of all consumers i for whom their willingness to pay for variant j, $s(i,j)$, exceeds $c(j)$. The definition of an industry immediately follows from the definition of potential competition. Notice that the spatial model represents only a particular case of the general definition above when the willingness to pay $s(i,j)$ is set equal to $r(i) - t(i,j)$. We shall frequently use the spatial analog in the following pages to illustrate competition among differentiated products, but it must be clearly understood that these illustrations are used only for simplicity, and not because the analysis would apply to the spatial context only. Thanks to the spatial analogy, the concepts introduced

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

above, like potential market or industry, do still hold in all situations in which firms sell substitute products represented in the space of characteristics. Yet, it remains true that, in most applications, the analysis of competition is limited to cases where products' variants are defined with respect to a single dimension only. In such cases variants are identified with only one index. Similarly a particular consumer is identified with a specific point along the line, and this point is viewed as the most preferred variant of this consumer, his/her ideal variant. The farther another point (variant) from this one, the lower the utility of this consumer for that variant. An example, due to Salop (1979), in which consumers are continuously distributed on the space of characteristics is provided below, in the form of a circular market. At each point of the circle is one consumer for whom this point constitutes his/her ideal variant because the transportation cost for this consumer at that point is equal to zero, while it is strictly increasing when moving away from this point in either direction.

It is easy to derive the consumer's demand function for the product's variants from his/her preferences on the space of characteristics. First let us show how to obtain this demand for consumer i, $i = 1, \ldots, m$, when the choice operates among the variants j, $j = 1, \ldots, n$, supplied by the n firms j, $j = 1, \ldots, n$, and according to the rule of *mutually exclusive purchases*: if the consumer decides to buy, he/she buys one unit only, and from only one firm. Let $r(i,j)$ denote the reservation price of consumer i for variant j and $p(j)$ the price of that variant. Consumer i, facing the vector of prices $(p(1), p(2), \ldots, p(j), \ldots p(n))$, chooses to buy variant j if, and only if, the two conditions: (i) $r(i,j) > p(j)$ and (ii) $r(i,j) - p(j) = max_{k=1,\ldots,n}\ (r(i,k) - p(k))$ are simultaneously satisfied. The first condition guarantees that consumer i prefers to consume variant j at price $p(j)$ than not to buy at all, and the second that consumer i selects the variant offering him the largest surplus in utility among all the variants supplied in the industry at the corresponding prices. This variant is called the *ideal variant*. The above procedure divides the population of consumers into two categories: the first gathers all consumers who buy a specific variant at this vector of prices, and the second one consists of all consumers who do not want to buy any variant at these prices. More precisely, the demand for variant j is defined as equal to the number of consumers i for which conditions (i) and (ii) defined above are simultaneously satisfied. Taking into account that each consumer buys at most one unit of a variant, this number correctly expresses the number of variants bought at this price vector. In order to get rid of the discontinuities due to the finite number of consumers, it is often assumed that the set of consumers is represented by an atomless continuum.

As an example, consider the circular market due to Salop (1979) (see Figure 6.4). Now we assume a market with consumers uniformly distributed along a circle. In this representation, there are four firms, *A, B, C,* and *D*, located equidistantly from each other around the circle. As for the consumers, they are uniformly dispersed around the circular boulevard, with one consumer located at each point. We shall denote a typical consumer by the letter t. This notation designates the consumer as well as the place where he/she is located: point t constitutes consumer t's ideal variant since the transportation cost to this point is equal to zero for him/her, and strictly positive to any other point on the circle. Assume that all consumers have the same willingness to pay for the product, s, and that the transportation cost of consumer t, to buy from firm $j, j = A, BC$ or D, is equal to the distance $d(j,t)$ separating consumer t from firm j, with $d(j,t)$ smaller than 1 for all j and all t. Assume that consumers have to move along the boulevard to buy from one of the sellers. If all sellers set the same price smaller than $s - 1$, condition (ii) is satisfied so that all consumers buy the product,

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

and buy it from the firm they are the closest to. Accordingly, the market share of each firm (its demand) consists of a quarter of the circumference, as on Figure 6.4, where the segment (a_1, a_2) constitutes the market share of firm A when all sellers set the same price. Consider indeed two adjacent firms k and j. For the consumer $t(p)$ located at the border between two market shares, he is necessarily indifferent between buying the product from the seller located on his/her left side and buying from the one located on the right. Accordingly the equality $d(k,t)+p = d(j,t)+p$ must necessarily hold: the sum of the transportation cost and the price must be the same for both sellers. All consumers located between $t(p)$ and firm j *a fortiori* buy from firm j since their transportation cost is lower than for $t(p)$. For the same reason, consumers located between $t(p)$ and firm k buy from firm k. Now assume that a firm, say firm A, decides to lower its price while the other firms still keep their price equal to p. This price drop pushes farther away the consumers indifferent between buying from firm A and firms B and D in the direction of these adjacent firms. The price drop makes firm A more attractive for the fringes of consumers located on both sides of its original market share who were buying from the adjacent firms before the price drop. The market share of firm A now coincides with the segment (a'_1, a'_2), as depicted on Figure 6.4.

We notice that firms B and D are direct potential competitors of firm A: the price drop directly affects the market share of the adjacent firms. On the other hand, the price drop does not affect the market share of firm C, which is only an indirect potential competitor of firm A via firms B and D.

An important distinction about differentiation often follows from the intrinsic characteristics of the products, changing from one variant to the other: shirts can be synthetic or linen, and TV sets can be black and white, or color. Some cars can be sports cars, while others are family cars. Similarly, shops can be close to some buyers and far from others, and vice versa. The first two examples correspond to cases of *vertical* product differentiation, while the latter to *horizontal* differentiation. Under vertical differentiation, *all* consumers buy the same variant when both are sold at the same price. This unanimity is not satisfied under horizontal differentiation: some consumers prefer to buy one variant and some the other, when they are sold at the same price. It turns out that market properties of product differentiation are not the same with the two types of differentiation. While vertical differentiation mainly evokes the differences in quality, all consumers ranking quality in the same way, horizontal product differentiation rests on the differences in tastes. A consumer close to a given shop prefers to buy from that, while another consumer closer to another shop would prefer to buy from the latter at equal price. As we shall see, the market properties differ from one type to the other.

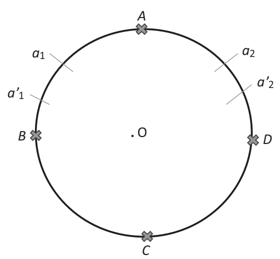


Figure 6.4 A circular market (Salop, 1979)

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

This chapter is subdivided into four main sections. Section 2 is devoted to horizontal differentiation while vertical differentiation is considered in Section 3. In Section 4 an essay on nesting the two approaches is provided. Section 5 is devoted to applications like environmental issues and credit card markets.

2 HORIZONTAL PRODUCT DIFFERENTIATION

The study of competition under horizontal product differentiation is developed in this section. Recall the definitions given in the introduction. Two variants of a product are horizontally differentiated when, sold at the same price, some consumers prefer to buy one variant over the other while the reverse is true for the other consumers. The archetypal model of horizontal product differentiation is the *Hotelling model* of spatial competition. The first subsections below study price competition in this spatial model, and its implication concerning the selection of the variants by the competing firms. Two opposite forces should *a priori* be considered in this respect. On the one hand, the firm selecting a variant remote from the space of characteristics of the variant selected by the rival firm, benefits from the advantage of protecting its market against the rival's price cuts: the firm enjoys an almost local monopoly position. But, conversely, this choice makes it more difficult for this firm to attract the rival's clients by setting a lower price, which should induce the firm to locate its own variant closer to the variant selected by its competitor. The above analysis is performed assuming that consumers have perfect information on the prices selected by the firms. This assumption is partially abandoned in the next section, where a less strict assumption is made concerning the information structure of the customers: here it is assumed that consumers know the price set by the firm closer to them, but are ignorant of the price set by the more remote firm.

The following subsection is devoted to the problem of entry in a horizontally differentiated market, as well as its consequences on the number of variants offered at equilibrium: how many variants are supplied? Are they too few, or too many? Are they the "good" ones? How does entry in a horizontally differentiated market differ from entry in a market for a homogeneous product? Also, some applications of the model are evoked at the end of this section, like competition between political parties, or competition in media markets. The preceding analysis has been made assuming that each firm supplies only a single variant. Firms, however, often sell several products and variants simultaneously. At the end of this chapter, different problems raised by multiproduct firms are studied, like brand proliferation to create barriers to entry, or the danger of cannibalization, the increase in sales of one variant produced by the firm decreasing the sales of another one also produced by the same firm. See, for example, Sections 3.4, 3.5 and 5.

2.1 Price Competition and Horizontal Differentiation

2.1.1 Price competition when the good is a homogeneous product

As already underlined in the introduction, a major motivation of firms to differentiate their products comes from the disastrous consequences of price competition on profits, when they sell a homogeneous product. When the sellers of a product are few, each of them should *a priori* benefit from some market power, resulting from its capacity to influence the selling price that will emerge as a consequence of aggregate supply. When the product is

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

homogeneous, and consumers have perfect information on prices, total demand is addressed to the firm quoting the lowest price. As a consequence, the other sellers must in turn decrease their own price in order to keep their customers. Accordingly, the price should progressively decrease down to the level where a further drop would cost more than would be beneficial. Under price competition, the price must finally decrease down to the competitive level, where it is equal to marginal cost. The whole market power resulting from competition among the few completely vanishes when the product is homogeneous. Probably this is why a market where few firms sell very similar variants of a product is rarely observed. Perhaps this is the case of gasoline stations located close to each other at the entrance of a highway or of some products hardly differentiable by modulating the level of their characteristics, like soaps or other housekeeping products. But, even in these extreme cases, sellers try to personalize their variant by choosing particular packaging, or by advertising specific virtues of their product that the rivals' variants do not share. In the extreme case where such a minimal differentiation appears impossible, the survival of firms necessarily requires a tacit collusive agreement on price, in order to keep it at a level exceeding the competitive price (see Vives, 2000 on this).

As soon as it is possible, one should expect firms to adopt any shelter available from price competition: it is harsher the closer the similitude in the space of characteristics between their products. Nonetheless, as stressed in the preceding section, there is a force operating in the reverse direction, which induces competing firms to locate in the space of characteristics not too far apart from their rivals. This force allows firms to choose their characteristics close enough to those selected by their rivals so as to capture their customers without having to consent to substantial price cuts. An endogenous determination of both prices and characteristics' levels emerges from the equilibrium conditions resulting from these two opposite forces.

2.1.2 Price competition in the spatial model

Hotelling (1929) proposed the following analysis of price competition in a spatial context. Two bakeries are located along a road of length L uniformly scattered with inhabitants. They constitute their potential consumers. Each individual consumes one unit of a homogeneous good, produced at zero cost, which he/she can buy either from seller 1, located at a distance a from the left extremity of the road, or from seller 2, located at a distance b from its right extremity (see Figure 6.5).

All customers are assumed to have a sufficiently high willingness to pay to be sure that all of them want to buy a unit of bread from either of the bakeries. Furthermore, each individual pays a transportation cost equal to the square of the distance he/she has to travel to join the chosen selling point. The problems raised by Hotelling are as follows. Which prices $p_i, i = 1, 2$, are the firms willing to set? How did they select their location, point a for seller 1, and $L - b$ for seller 2? As we have just seen, these two questions are intimately related. Here we concentrate on the first, assuming that the sellers have already selected their location.

Whichever consumer t in the interval $[0, L]$, he/she will choose the seller offering him/her the best opportunity. Since the bread is identical for both sellers, a consumer located between

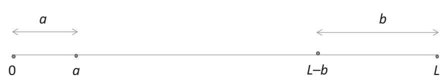


Figure 6.5 Price competition in the spatial model

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

the two sellers will select to buy bread at the seller for whom the transportation cost to travel to him plus the price is the lowest. If p_1 (resp. p_2) denotes the price at bakery 1 (resp. bakery 2), and x (resp. y) the distance between the consumer and bakery 1 (resp. 2), the consumer will arbitrate for bakery 1 (resp. 2) whenever $p_1 + cx^2 < p_2 + cy^2$. When these two numbers are equal then the corresponding consumer – call him/her $t(p_1, p_2)$ – is indifferent between the two alternatives and is just at the border between the two markets: those consumers located at the left of consumer $t(p_1, p_2)$ buy from bakery 1 while those at his left buy from bakery 2. Similarly, those consumers located at the left of bakery 1 buy from bakery 1 and those located at the right of bakery 2 buy from bakery 2. The market shares of the two bakeries at the pair of prices (p_1, p_2) are thus represented by the two intervals $\left[0, t(p_1, p_2)\right]$ and $\left[t(p_1, p_2), L\right]$, respectively. It is not difficult to determine the consumer $t(p_1, p_2)$ who is at the border of the two markets. First the equality $p_1 + cx^2 = p_2 + cy^2$ must necessarily be satisfied since this consumer is indifferent between buying his bread at either bakery. Furthermore, the equality $x + y = L - a - b$ must also hold: the distance between the two sellers is equal to the total distance (L) minus the distances from seller 1 to 0 (a) and from seller 2 to $L(b)$. Accordingly, the explicit value of $t(p_1, p_2)$ can be found solving the linear system formed by these two equations in x and y, i.e.

$$x = \frac{p_2 - p_1}{2c(L - a - b)} + \frac{L - a - b}{2};$$
$$y = \frac{L - a - b}{2} - \frac{p_2 - p_1}{2c(L - a - b)}.$$

It is now easy to determine the demand functions of the two sellers as a function of their prices. If we denote by $D_i(p_1, p_2)$ the quantity demanded from firm $i, i = 1, 2$ at the pair of prices (p_1, p_2), we get

$$D_1(p_1, p_2) = a + x = a + \frac{p_2 - p_1}{2c(L - a - b)} + \frac{L - a - b}{2};$$
$$D_2(p_1, p_2) = b + y = b - \frac{p_2 - p_1}{2c(L - a - b)} + \frac{L - a - b}{2}.$$

The profits of the sellers corresponding to the pair of prices (p_2, p_1) are thus equal to $p_1 D_1(p_1, p_2)$ for firm 1 and $p_2 D_2(p_1, p_2)$ for firm 2. Now all the ingredients with which to define a game are brought together: the players are represented by the two bakeries, the strategies correspond to the prices chosen by each firm and the payoffs to the profit functions we have just identified. Now we use the Nash equilibrium concept proposed by game theory to represent the choices of prices made by each firm. A *Nash equilibrium* is a pair of prices such that no player can increase his/her payoff by deviating unilaterally from the strategy he/she has selected at this pair. Applying this criterion to our situation leads us to identify, for each bakery, the price maximizing its profits, given the price selected by the rival. From the first-order conditions we easily derive the equilibrium values for the prices, namely,

$$p_1^* = \frac{c(L - a - b)}{L + (a - \frac{b}{3})};\ p_2^* = \frac{c(L - a - b)}{L - (a - \frac{b}{3})}.$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

A look at these equilibrium prices immediately shows that their values explicitly depend on the locations that sellers have selected, revealing that price competition is directly related to the positioning of products in the space of characteristics (here represented by the set of points on the line (a, b)). In particular, when the sellers locate exactly at the same place (in this case, $a = L - b$), the two variants are perceived by all consumers as perfectly homogeneous. We discover that both prices are then equal to zero, the competitive price.

The previous analysis has been conducted assuming that the sellers had already selected their location before choosing their price. Thanks to this assumption, we were able to answer the first question raised by Hotelling: which prices will the competitors select? We have just identified their values at the Nash equilibrium of the price competition game. It appears, however, that these values depend on the locations of a and $L - b$ selected by the sellers. Accordingly, we still need a method to identify how the firms select their location in order to answer the second question raised by Hotelling: how do firms choose their location, point a for seller 1, and $L - b$ for seller 2?

2.1.3 The choice of locations: Spatial competition

The previous analysis reveals that the value of prices (p_1^*, p_2^*) is known as soon as locations are given. If the sellers are able to anticipate these prices as a consequence of their location choice, they will use this information in order to evaluate what their profits are at each location. To this end, it is sufficient to evaluate their demands $D_i(p_1, p_2)$ at the pair of prices (p_1^*, p_2^*) and multiply the resulting magnitude by the corresponding equilibrium price $p_i^*, i = 1, 2$. For instance, the profits of seller 1 corresponding to the location $L - b$ of seller 2 are given by $\frac{c(L-a-b)}{18(a-b+3L)^2}$, while those of seller 2 corresponding to the location a of seller 1 are equal to $\frac{c(L-a-b)}{18(b-a+3L)^2}$. The first expression is a function of the location choice of seller 1 and, similarly, the second is a function of the location choice of seller 2. Assuming that the selection of locations takes place before the selection of prices, it is easy to determine the best location of each seller as a function of the location selected by the rival. For instance, imagine that seller 2 has chosen to locate at point $L - b$. Then, looking at the expression of profits of seller 1 as given above, it is easy to see that the profits of seller 1 are monotone decreasing with his location variable a. Another way to say the same thing is that the farther seller 1 locates from seller 2, the higher seller 1's profits. It follows that, whatever the location choice of his rival, the best decision of seller 1 is to locate at $a = 0$. Reasoning in the same manner from the viewpoint of seller 2, we observe that, whatever the location strategy a of seller 1, the optimal location choice of seller 2 obtains by choosing $b = 0$, entailing a location at the point $L - b$. Thus, we conclude that, if the sellers correctly anticipate the equilibrium prices as a function of the locations, and if the transportation costs are quadratic in distance, both firms will locate at the opposite extremities of the linear market.

It is easy to interpret this conclusion. We have seen above that, when the two sellers are located at the same place (in which case $a = L - b$), their variants are like perfect substitutes: equilibrium prices are then equal to zero. Also, a close scrutiny of equilibrium prices reveals that these prices tend to zero when the sellers' locations tend to each other. Accordingly, in spite of the agglomeration force that induces both sellers to get closer to each other in order to increase their market share, the loss in product differentiation implied by these moves reduces prices to such an extent that it ceases to be profitable. Between the agglomeration force and the strength of competition, the latter appears to be more significant: sellers differentiate their

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

variants as much as possible, by locating themselves at the two opposite extremities of the linear market. When the choice of location and price operates in a sequential manner without cooperation between the sellers, and with quadratic transportation costs for consumers, these sellers constitute "local monopolies" in order to protect themselves as much as possible from too harsh competition with their rival. A natural question immediately comes to mind: how robust is this conclusion?

2.1.4 Robustness of the analysis

The conclusion in the above paragraph seems to exclude the possibility of agglomeration as the major force operating in horizontal differentiation. On the contrary, as a result of spatial competition, firms seem to prefer to remain far apart from each other. But this result follows from the particular assumptions under which the analysis has been conducted. Hotelling (1929) was assuming in his original paper that the cost of distance is a *linear* function of the latter while it is assumed here that the cost varies as the square of the distance. It can be shown that, under Hotelling's specification of transportation costs, *there is no price equilibrium when the firms are located too close to each other* (d'Aspremont, Gabszewicz, and Thisse, 1979). This conclusion excludes the possibility of applying a similar analysis to the above in order to characterize the result of spatial competition when assuming linear transportation costs. Since no price equilibrium exists for some locations, one cannot theoretically conclude at which location sellers would finally stabilize their choice. Finally, we can say that no general conclusion can be drawn from Hotelling's analysis as whether firms have incentives to reduce competition by moving away from each other or on the other hand to agglomerate. Hotelling thought erroneously that a kind of *principle of minimum differentiation* should hold. But this conclusion was not robust to the choice of the transportation cost function.

In the real world, many signs seem to indicate a strong tendency of firms to locate in similar areas. Numerous examples, provided in particular by Hotelling himself, reinforce the desire to prove the existence of an agglomerating force dominating the dispersion of firms in order to relax competition. For instance, De Palma et al. (1985) show that, under sufficient heterogeneity of consumers and random preferences, the principle of minimum differentiation still holds. Similarly, Lin and Tu (2013) develop a variant of Hotelling's model involving an intermediate good's market in order to explore the location choices of firms. They consider interactions among three firms: a wholesale supplier of an essential input and two retail producers. One of these retailers is a vertically integrated firm. The other is an independent downstream firm. Then they discuss the role of strategic vertical outsourcing in determining optimal locations for firms, and the input pricing of wholesale suppliers. In general, the two firms were located more closely when the vertically integrated firm had a cost advantage without taking strategic outsourcing into consideration. However, the price of the input may increase when we take strategic vertical outsourcing into account and this may cause the two firms to move farther away, giving rise again to the principle of minimum differentiation. In his contribution to this stream of literature, Jehiel (1992) studies the effects of costs on differentiation. In the standard model, not including differentiation costs, firms choose to locate at the extremes of the linear city in order to dampen price competition. In his setup, the differentiation costs increase the attractiveness of the central locations, as the production cost increases with the distance to the center. He finds that low differentiation costs, relative to transportation costs, do not affect the equilibrium location choices of the

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

firms (maximum differentiation). However, sufficiently high differentiation costs induce firms to locate in the interior of the city (partial differentiation), increasingly closer to the center as the differentiation costs increase.

It is important to notice that the principle of minimum differentiation is restored when prices are not taken into account as strategies in the interaction between the rivals, as in political competition or in the case when firms compete only on the choice of characteristics of their variant. Applying the standard model described above to political competition between two candidates at an election, Downs (1956) considers two political candidates, the first representing the Left Party while the second represents the Right Party. The voters' population is uniformly spread on the political spectrum, represented by the interval $[0, 1]$: each point in this interval simultaneously identifies a voter and a political program, ranking from the extreme left (point 0) to the extreme right (point 1) and the voter for whom this political program constitute his/her "ideal" program. Each elector votes for the program lying closest to his/her own position. The objective of each candidate is to announce the program maximizing the number of voters for the selected program. Which programs will the two candidates select at equilibrium? In this representation, the main ingredients of the spatial competition model are inserted. The linear market is replaced here by the political spectrum, and the choice of a particular variant by the consumers becomes the selection by the voters of a particular program between those proposed by the candidates. Also, the transportation cost in the spatial model here becomes the cost for the voter of voting for a political program differing from his/her ideal one. But there ends the analogy. In particular, the other element conditioning the consumers' choice between the variants in the spatial model, namely their price, is not present here since the only element taken into account in the voters' choice is the inadequacy of the proposed programs to their own ideal. Accordingly, this interpretation is simpler since it does not deal with price selection but only with variant selection, where a variant is identified with a specific political program on the political spectrum. It gives rise to a game with the players being the political candidates, the strategies the programs to be chosen in the political spectrum $(0, 1)$ and the payoffs the number of votes obtained by each candidate at each pair of programs selected. The Nash equilibrium of this game can provide an answer to the question formulated above concerning which programs will be proposed to the vote of electors. It is easy to show that the only equilibrium of this game consists of the pair of programs $(1/2, 1/2)$, which obtains when each candidate selects the perfectly centrist program represented by point $\frac{1}{2}$ of the political spectrum. This property is called the *median voter theorem.*

To conclude, nothing can be said in the standard Hotelling model about the forces underlying competition in the case of horizontal product differentiation. Of course, introducing alternative assumptions in the model can restore the principle of minimum differentiation. However, in the pure model of horizontal differentiation, equilibrium prices crucially depend on the transportation costs, sometimes driving the equilibrium to maximal differentiation (quadratic transportation costs), or to non-existence of an equilibrium (linear transportation costs).

2.2 Entry and Horizontal Product Differentiation

Using the model of a circular market à la Salop, we have seen in the introduction how the chain structure of demands implies a localized competition between each seller and their two

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

neighbors in the space of characteristics. As a consequence, the entry of new competitors in a differentiated market requires taking into account the interdependence among sellers: even if they are very numerous in the market, they are in direct competition with only a few of them, their immediate neighbors. Our purpose now is to illustrate the entry of new sellers in horizontal differentiation, using the circular market as a paradigm to describe strategic aspects of entry on the circle. Let us recall the ingredients of this model. The market consists of a circular boulevard of length 1, with sellers equidistantly located around the circle. As for the consumers, they are uniformly dispersed along the boulevard so that, at each point, there corresponds a different consumer, denoted by the letter t: this letter denotes the consumer t, as well as the location of this consumer on the boulevard. The point t thus also constitutes the ideal variant of this consumer since the transportation cost to this point is equal to zero for him.

Let us suppose that for all consumers transportation varies linearly with the distance $d(j, t)$ to be traveled by consumer t to buy the good at shop $j, j = 1, \ldots, n$ (see Figure 6.6).

Furthermore, suppose that consumers are forced to travel along the boulevard in order to shop. To determine the demand addressed to firm j, fix all the prices p_k of the other sellers. The share of the market going to firm j is then constituted by the segment of the circle lying between the consumer indifferent between buying from seller $j-1$ or from firm j, and the consumer indifferent between buying from seller j and seller $j+1$. The position of the first of these consumers is determined by the condition

$$cx + p_{j-1} - 1 = c(1/n - x) + p_j,$$

where x denotes the consumer indifferent between buying from seller j or seller $j-1$ (recall that the n firms are located equidistantly on the circle, implying that the distance between two firms is equal to $1/n$) and cx represents the (linear) cost of traveling a distance x. Similarly, the position of the consumer y indifferent between buying from firm j and $j+1$ must satisfy the equality

$$cy + p_j + 1 = c(1/n - x) + p_{j+1}.$$

All consumers located between the consumer satisfying the first equation and firm j buy from firm j, because the distance between them and firm j is smaller than the distance between them and firm $j+1$. Similarly, all consumers located between the consumer satisfying the

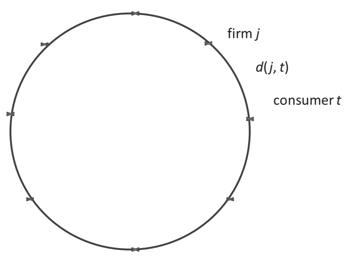


Figure 6.6 Entry and horizontal product differentiation

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

second equation and firm j buy from firm j as well, and for the same reason. Solving the two above equations in x and y, we obtain

$$x = \frac{\frac{1}{n}(p_{j-1} - 2p_j + p_{j+1})}{2c} + \frac{1}{2n}$$

and

$$y = \frac{\frac{1}{n}(p_{j+1} - 2p_j + p_{j-1})}{2c} + \frac{1}{2n}.$$

Thus, adding the two above expressions, we obtain the demand addressed to firm j, namely

$$D_j(p_{j-1}, p_j, p_{j+1}) = \frac{(p_{j-1} - 2p_j + p_{j+1})}{2c/n} + \frac{1}{n}.$$

From this expression, it clearly follows that the demand addressed to firm j depends on its own price and on the prices of its immediate neighbors, but not on the prices quoted by more remote firms. This is the essence of localized competition. Also, as expected, this demand decreases when firm j increases its price, but increases when neighboring firms $j-1$ or $j+1$ increase their prices: differentiated products are indeed imperfect substitutes. If the unit production cost b is constant and identical for all firms, the profits of firm j are written as:

$$(p_j - b)D_j(p_{j-1}, p_j, p_{j+1}).$$

Due to the complete symmetry of the model, one must expect that all firms sell their product at the same price p^* at equilibrium. First-order conditions for profit maximization reveal that all equilibrium prices are equal to $p^* = b + c/2n$.

First, notice from the preceding analysis the discrepancy between the equilibrium price p^* with n firms in the market, and the competitive price. In a competitive market all firms would have sold their product at a uniform price equal to the marginal cost, b. In this case, the product is the same in each firm: no differentiation exists among firms from the viewpoint of consumers. When the sellers are dispersed over the space, with the cost parameter c strictly positive, as in the situation described above, each seller benefits from a kind of local monopoly power allowing the firm to sell above the marginal cost. While two firms are sufficient to destroy the monopoly rent in a homogeneous market, horizontal differentiation maintains positive profit margins whatever the number n of sellers. The higher the profit margin, the higher the value of the cost parameter c representing the unit transportation cost. On the other hand, when the value of c is weak, the units of the product sold by different firms appear as more homogeneous since the transportation cost is weak: a small price differential between two firms may justify traveling from one to the other while a higher transportation cost would not have allowed it. In this case, competition is harsher since variants are less differentiated, and the equilibrium price p^* tends to the competitive price when c tends to zero.

A similar effect appears when increasing the number of firms, n. The equilibrium price p^* tends also to the competitive price b when the number n of firms tends to infinity: this reflects the decline in profits resulting from the entry of new firms, which reinforces competition by

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

diminishing the distance between adjacent firms (here we compare situations in which all firms relocate equidistantly on the circle when a new firm enters the market). Notice also that, without any fixed cost, nothing can stop the entry of new firms in a horizontally differentiated market. There is no upper bound on the number of firms. It is sufficient for the new entrant to locate its shop between two existing sellers and capture all the customers who are now located closer to the entrant than to the adjacent incumbents. By contrast, the existence of a fixed cost F will limit the proliferation: when there is free entry in the sector, the number of firms cannot increase beyond the level where the revenue of each firm would become smaller than the fixed cost, F. This property imposes an upper bound on the number of firms capable of surviving at equilibrium. In our example, the revenue per firm at equilibrium is equal to $\frac{c}{3n}$. Accordingly, the number of firms able to survive at equilibrium is obtained from the condition $\frac{c}{3n} \geq F$, or $n \leq 1/\sqrt{3}F$. The number of firms surviving at equilibrium results from a compromise between the transportation costs paid by the consumers and the fixed costs paid by the firms. When the fixed cost is small, the number of firms can be very high. As a consequence, the price p^* tends to the marginal cost, b, or the competitive price. When the number of firms becomes very large, each consumer finds one of them in an immediate neighborhood, and competition is very harsh. By contrast, when the transportation cost c is weak, only a small number of firms can survive at equilibrium. In spite of this, the equilibrium price tends to the marginal cost when c becomes very small. When products are horizontally differentiated, the equilibrium price can be close to the competitive price either because they are weakly differentiated (c is small), or because the fixed cost F is small, which allows the entry of a large number of firms.

In the preceding analysis, we assume that the incumbent firms do not anticipate the entry of new firms and relocate equidistantly at zero cost around the circle each time a new firm enters the market. Sequential entry when dropping this assumption has been analyzed by several authors (in particular Prescott and Visscher, 1977). This analysis introduces a new element. The incumbent firms are handicapped, being unable to relocate. Yet they are assumed to be able to anticipate the future entry of new competitors and accordingly develop strategies to choose the most profitable locations before entry of new competitors, taking into account *ex ante* these future entries. In the case of sequential entry, one can show the persistence of strictly positive profits at equilibrium, even under free entry (Eaton and Lipsey, 1980).

2.3 Social Optimum and Market Equilibrium in the Spatial Model

At the end of the preceding subsection we analyzed how the number of firms was influenced by the size of the fixed cost F and the transportation cost c. Now we try to compare the number of firms at a market equilibrium with the number corresponding to a Pareto-optimal allocation. To evaluate this number, we can write down the welfare W corresponding to n firms equidistantly located on the circular market as

$$W = 2n \int_0^{\frac{1}{2n}} (v - b - cx)dx - nF = (v - b - \frac{c}{4n}) - nF$$

(notation v expresses the willingness to pay for the product, assumed to be identical for all consumers, and sufficiently high for all consumers to buy a unit of it at equilibrium). The expression W can be explained as follows. When n equidistant firms operate in the industry,

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

the consumer at the border between two market segments travels a distance equal to $1/2n$ to the closest firm. Consequently a consumer located at a distance $x \leq 1/2n$ from this firm enjoys a surplus exceeding the marginal cost b, equal to $v-b-cx$. Maximizing the expression W with respect to n, we find that the number of firms maximizing total welfare is equal to $\frac{1}{2}\sqrt{c}/F$. A direct comparison from this value of n with the value of n resulting from the market mechanism ($n = \sqrt{3c}/F$) reveals that it is difficult to conclude whether competition in the spatial model generates too much, or too little, diversity. The result of the comparison depends crucially on the specific values of the cost parameter c and the size of the fixed cost F.

3 VERTICAL PRODUCT DIFFERENTIATION

This section studies competition in a vertically differentiated market, and its effects on product differentiation. Two variants of the same product are vertically differentiated when all consumers prefer to buy one of these variants over the other, when both variants are sold at equal price.

A major reason why variants are vertically differentiated follows from income disparities: there are consumers who have sufficient income to buy the luxurious variant while other consumers possibly have access to the standard one only.

3.1 The Various Dimensions of Quality

The "quality" of the different variants of a product have to be specified by reference to the set of characteristics defining this product. Assuming the utility of all consumers of this product is increasing in the amount of each characteristic, it is easy to compare the quality of the variants by comparing the amount that each of them embodies. In the particular case where one variant embodies a larger amount of each of them, compared with another one, it seems clear that the first variant is of higher quality than the other. It is rather exceptional for a product that the utility of a consumer would increase, whatever the characteristic defining this product. More generally, the utility is increasing with respect to some characteristics and decreasing with respect to others when climbing along the quality ladder. Consider, for instance, different types of French wines, like a bottle of Burgundy or Bordeaux. Everybody would agree that the wine is tantamount to higher quality, the higher the respect and devotion to the sanitary conditions of its production. Nevertheless, there is disagreement when appreciating the two bottles of wine since many consumers prefer the Burgundy wine to the Bordeaux, and vice versa for the remaining consumers (this example is borrowed from Coestler and Marette, 2004). Similarly the utility of a restaurant's customers probably increases with the freshness of the food, but some customers like a small restaurant while others prefer a larger one. This is also the case for tourist hotels in seaside resorts. These hotels are vertically differentiated by their comfort (number of "stars"), but also by their location. Some tourists would prefer to sacrifice a star to be in a hotel at Sharm-el-Sheikh rather in the Canary Islands, or vice versa. In all these cases, variants are vertically differentiated with respect to some characteristics, but horizontally with respect to others.

Situations in which a product can be defined with respect to a single characteristic only are simpler to analyze. It requires, however, the utility of all consumers to be monotone increasing in the amount of this unique characteristic. In this case, indeed, the quality of a variant is

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

simply measured by the amount of the single characteristic it embodies. A first characteristic frequently studied by industrial economists is the reliability of a variant. This reliability can be measured unequivocally by the probability of success of a variant, measured either by its functioning or by the expectation of the customer. The higher this probability, the higher the quality of the variant. This is the case, for instance, for a bulb endowed with a given probability of functioning when it is switched on. Similarly, concerning the life duration of a product: the longer the life, the higher the quality of the variant. This also can be the case with a bulb or a car or, dresses: their quality is often judged on their resistance to wear. Further, it is also possible to compare two variants according to their chemical components or to a particular index, like the number of carats of a jewel or the proportion of silver and tin content of a silver dish, or the number of watts of a bulb. Moreover, product quality can also be measured by the number of warranties accompanying the sale of the product, or the quality of the after-sales service. Finally notice that, for some goods, it is possible to substitute the quality and the quantity, so as to keep the utility obtained from consuming such goods constant. For instance, one can obtain the same quantity of light in a home either by using a single bulb of a given number of watts, or by combining several bulbs of different watts, the sum of which corresponds to the number of watts of this single bulb. Similarly, if the quality of a razor blade is identified as its life duration, and if there are two different types of blades, one of a life duration equal to half the duration of the other, it is possible to obtain the same service by purchasing two blades of the first type, or a single blade of the second!

3.2 Price Competition and Vertical Differentiation

3.2.1 Vertically differentiated duopoly

Consider the following simple model of two vertically differentiated variants of a given product, based on income disparities existing in the population of potential buyers. Suppose there are two firms, with one of them selling the "standard" product at price p_1 and the other the "luxurious" one at price p_2. To simplify, we shall temporarily assume that production costs of the two variants are equal and proportional to the quantity produced. One can then assume, without loss of generality, that production costs are equal to zero. All potential buyers agree that variant 2 is of a higher quality than variant 1. All consumers have identical preferences but differ by their income. We represent these differences by assuming that consumers are ranked by order of increasing income in the unit interval, from the poorest customer (the consumer represented by point 0 in the interval) owning an income R_1 to the richest one (represented by point 1 in the same interval) owning an income equal to $R_1 + R_2$. As for the consumers t located in the interior of the interval, they own an intermediate income equal to $R_1 + R_2 t$, t denoting a point in the interval. Figure 6.7 represents the different income levels assigned to the agents in the population of consumers.

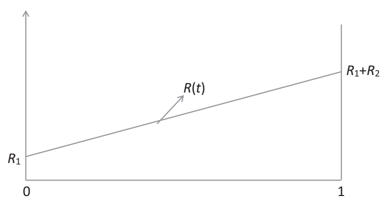


Figure 6.7 Vertically differentiated duopoly

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

This income distribution is uniform between the income R_1 and $R_1 + R_2$, respectively. Income R_1, corresponding to the income level of the poorer customer, is proportional to the population's average income: an increase in this level equally affects the income of every member of the population. Similarly, the parameter R_1 is proportional to the variance in the population's income. Suppose indeed that we keep the total mass of income constant while transferring a proportion of it from the richer to the poorer customers. This is equivalent to rotating the line joining R_1 to $R_1 + R_2$ while maintaining the surface constant below this line. This corresponds to a change in the dispersion of income distribution.

Let us now define consumers' preferences. As remarked above, we assume that all consumers are identical from this viewpoint. We also suppose that, if a consumer decides to buy one unit of some variant, he/she decides to buy it at the exclusion of the other. Furthermore, if he/she decides to buy a variant, he/she buys a single unit of it. These assumptions depict most purchasing decisions about manufactured goods, like pianos or television sets. If one consumer is considering buying a piano, he/she is satisfied with one piano only, and the choice is which one, among all brands available in the industry. This choice represents the classical arbitrage between price and quality when a consumer examines whether the increase in utility obtained when buying a higher-quality variant compensates for buying it at a higher price. In view of obtaining an explicit representation of consumers' preferences, we suppose finally that the consumer denoted by t, with income $R(t) = R_1 + R_2 t$, obtains a utility level given by $u_1(R_1 + R_2 t)$ when consuming variant 1 (low-quality variant), by $u_2(R_1 + R_2 t)$ when consuming variant 2 (high-quality variant) and $u_0(R_1 + R_2 t)$ when consuming neither of them. Since all consumers prefer owning a unit of the higher-quality variant to the lower one, and prefer to own a unit of the good, even if this unit is of lower quality, the parameters u_0, u_1 and u_2 must satisfy $u_2 > u_1 > u_0$. It is now easy to derive the reservation price of each consumer t according to his/her income $R(t)$ for each of the two variants. Taking into account the fact that the utility of the consumer buying a variant i must be equal, for each variant, to the utility obtained when he does not buy it, the equalities

$$u_i(R_1 + R_2 t - s_i(t)) = u_0(R_1 + R_2 t), i = 1, 2$$

have to be be satisfied, with $s_i(t)$ representing the reservation price of consumer t for variant $i, i = 1, 2$. Solving these equations with respect to $s_i(t)$,we obtain:

$$s_i(t) = \frac{u_i - u_0}{u_i}(R_1 + R_2 t), i = 1, 2.$$

Figure 6.8 represents the reservation price of each of the two variants. As expected, the reservation price of the luxurious variant is higher than the standard one, whichever the consumer t. This property reflects the unanimity of consumers about the ranking of variants on the quality ladder. Moreover, whatever the variant, the reservation price is increasing with consumers' income.

3.2.2 Monopoly pricing

We start by assuming that the only variant is the low-quality good, and we study what price would be set by the monopolist, and which consumers would be willing to buy at this price. Let p_1 be any price set by the monopolist. All consumers whose reservation price exceeds p_1

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

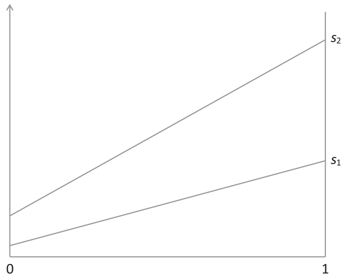


Figure 6.8 Vertical differentiation and reservation price

buy the product. Denote by $t(p_1)$ that consumer for whom the reservation price is equal to p_1. The demand at this price is equal to the length of the interval $\left[1, t(p1)\right]$. The explicit value of $t(p_1)$ is easily found from the condition $\frac{u_1 - u_0}{u_1}(R_1 + R_2 t) = p_1$. Solving this equation for t, we obtain

$$t(p_1) = \frac{u_1 p_1}{(u_1 - u_0)R_2} - \frac{R_1}{R_2}.$$

Demand at price p_1 is thus equal to $1 - t(p_1) = 1 - \frac{u_1 p_1}{(u_1 - u_0)R_2} + \frac{R_1}{R_2}$, with a revenue equal to $p_1(1 - \frac{u_1 p_1}{(u_1 - u_0)R_2} + \frac{R_1}{R_2})$. Maximizing this revenue with respect to price, and assuming that the solution is interior to the domain of prices $(u_1 R_1, u_1(R_1 + R_2))$ in which the demand does not cover the set of consumers, we find the optimal price from the necessary first-order condition, namely,

$$p_M = \frac{(R_1 + R_2)(u_1 - u_0)}{2u_1},$$

providing a receipt equal to $(R_1 + R_2)^2(u_1 - u_0)/(4u_1 R_2)$.

It is important, however, to recall the assumption we have made to derive this optimal price: we have assumed that it strictly belongs to the interval

$$[u_1 R_1, u_1(R_1 + R_2)]$$

of those prices at which the market is not covered. Thus, the optimal price should exceed the value $u_1 R_1$ corresponding to the reservation price of the poorest customer present in this market. In the opposite case, when the optimal price p_M does not satisfy this condition, it must take the value $p_M = u_1 R_1$: this is the highest price among all those whose corresponding demand is equal to 1. At this price, the revenue of monopoly is also equal to 1. When comparing this revenue with that corresponding to the interior solution, we find that the first exceeds the second if, and only if, $\frac{R_1}{R_2} > 1$.

We can draw the following conclusions from the preceding analysis. First, two market situations can emerge from profit maximization by the monopolist, according to the values characterizing the income distribution. According as $\frac{R_1}{R_2} \lesseqgtr 1$, the whole market is served or,

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

on the contrary, some potential customers prefer to refrain from buying because they find the price too high. These customers are of course those with the smaller income. This situation must arise when incomes are highly dispersed, with “very” rich and “very” poor customers. The first have higher reservation prices while the remaining ones rather low ones. Thus, the monopolist has an interest in setting a higher price in order to capture the surplus from richer customers, even if it leads to abandoning the share of the market consisting of the poorer ones. The reverse holds when the income distribution is “flatter”. The monopolist in this case has an interest in fixing the price equal to the reservation price of the poorest customer and serving the whole market. Concerning the choice of quality, we notice that both price and monopoly revenue are increasing functions of u_1. Thus, the higher the quality, the higher the monopoly revenue and price. Consequently, if the monopolist has control over the quality, he will always choose the highest possible one. This conclusion should, however, be tempered by the fact that we have assumed so far zero production costs. In fact, production costs are generally increasing with the quality of the variant. Consequently, when costs are introduced, the monopolist should arbitrate between the increase in quality and the resulting increase in cost, a property we have not considered so far in the analysis. This will be done when costs are introduced.

3.2.3 Equilibrium prices under duopoly

Let us now return to our initial situation consisting of a differentiated duopoly, with firm 1 selling a standard product and firm 2 a luxurious variant of it. This situation could be interpreted as the entry of a new producer in the market, selling a higher-quality variant than the quality sold by the incumbent. Which price equilibrium will then follow? With respect to the earlier situation with only one low-quality variant on the market, when the consumer had only to decide to buy the product at the running price or not, the consumer is now faced with a new problem: which variant should he/she decide to buy? Given the pair of prices p_1 and p_2, a consumer t buys one unit of variant 2 if and only if, the utility level $u_2(R_1+R_2t-p_2)$ obtained when buying one unit of variant 2 exceeds the utility level $u_1(R_1+R_2t-p_1)$ obtained when buying one unit of variant 1. In the case when the inequality is satisfied in the reverse order, the consumer buys the first variant. In all cases, a second requirement is that the reservation price $s_i(t)$ of variant i exceeds the price p_i. Thus, a purchase of variant 2 takes place if, and only if, the two following inequalities $u_2p_2 - u_1p_1 < u_2s_2(t) - u_1s_1(t)$ and $s_2(t) \geq p_2$ are satisfied.

Figure 6.9 describes how the market is partitioned for three different pairs of prices.

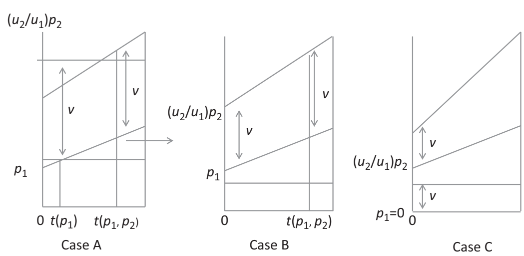


Figure 6.9 The market for three different pairs of prices

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

In the first case (case A), the set of consumers is partitioned into three parts. In the interval $[0, t(p_1)]$, consumers buy neither variant 1 nor variant 2: in such situations, their reservation price is smaller than the price corresponding to the associated variant. All consumers in this set have the weaker income levels. By contrast, all consumers lying in the interval $[t(p_1), t(p_1, p)]$ buy variant 1. This group consists, on the one hand, of consumers who, in any case, would not buy variant 2 at price p_2, and of consumers who would have bought variant 2 without the presence of variant 1, but prefer to refrain from buying this variant, given the difference in prices. This group of consumers represents the class of "average" income levels. Finally remains the set of consumers in the interval $[1, t(p_1, p_2)]$, namely, those who buy variant 2. These are the richer consumers. The sharing of the market between those who buy variant 1 and variant 2 takes place at the "frontier" consumer $t(p_1, p_2)$ for whom the equality

$$u_2 p_2 - u_1 p_1 = \frac{s_2(t)}{u_2 - u_1}$$

is satisfied. Denoting by V the difference $u_2 p_2 - u_1 p_1$, we find the consumer $t(p_1, p_2)$ as the consumer t for whom the distance between the two lines on Figure 6.9 is exactly equal to V.

Suppose now that firm 1 decides to lower its price p_1 to some level below $u_1 R_1$, the reservation price of the poorest customer t, $t = 0$. This fall in price influences the size of firm 1's market share in two different ways. On the one hand it increases its market share by "stealing" some customers from the original market share of its rival, those who are located just at the right of $t(p_1, p_2)$, who now prefer to buy from firm 1 due to the fall in price p_1. On the other hand, firm 1 also increases its market share by acquiring customers to the left of customer $t(p_1)$ who did not buy anything before the fall in price, but now start to be attracted by variant 1 due to the fall of its price. Further, this fall in price has the effect of covering the market when it is large enough to lower the price below the reservation price of the poorest customer, namely, below $u_1 R_1$: all consumers now buy either from firm 1, or from firm 2. Then we are in case B on Figure 6.9. All customers buy a single unit either from firm 1 (those consumers in the interval $[0, t(p_1, p_2)]$, or from firm 2 (the interval $[1, t(p_1, p_2)]$. Finally, in case C in Figure 6.9, firm 1 has no customers because all of them prefer to buy the luxurious product. Everything happens as if the firm entering the market has succeeded in eliminating the incumbent and is now a monopolist in place of the latter.

The preceding analysis reveals that the domain of a pair of prices can be partitioned into three subdomains. In the first one, corresponding to case A, both firms obtain a positive market share at running prices, but the market remains uncovered: some customers, the poorest ones, do not buy any of the variants. In the second, case B, both firms have a strictly positive market share, but the market is entirely served. Finally, in case C, firm 2 is the only one present in the market: firm 1 cannot find any customers, even setting a price equal to zero. It can be shown that, for any values of the parameters (u_1, u_2, R_1, R_2) characterizing the income distribution and the preferences of the customers, there is one, and only one, equilibrium pair of prices, depending on these values. One can show that this unique pair of prices lies either in domain A, or B, or C, according to the particular values of the parameters (u_1, u_2, R_1, R_2). When these parameters satisfy the inequalities:

$$0 \leq R_1/R_2 \leq \frac{u_2 - u_1}{3(u_2 - u_0)},$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

the pair of equilibrium prices lie in the subdomain A. In other words, these parameters lead to a market situation where both firms have a positive market share and the whole market is not served. On the other hand, when the parameters satisfy

$$\frac{u_2 - u_1}{3(u_2 - u_0)} < R_1/R_2 \leq 1,$$

the two sellers share the whole market, as in case B. Finally, when $1 \leq R_1/R_2$, the whole market is now served by firm 2 selling the high-quality variant, and firm 1 is excluded from the market: now we are in case C. Finally, an explicit solution for equilibrium prices can be found for each of the three cases (for more details, see Gabszewicz and Thisse, 1979).

3.2.4 Properties of equilibrium prices

It follows from the above discussion that, when $R_1/R_2 \leq 1$, the entry in the market of a firm selling a higher-quality variant than the incumbent is compatible with the presence of this incumbent in the market: only prices have to readjust at their new equilibrium values. It can be verified that these prices are lower than those set by each seller as a monopolist. By contrast, when $R_1/R_2 > 1$, the entry of a new variant can only be realized with a higher-quality variant, and is always accompanied by the exit of the standard product. Moreover, one can verify that, in this case, the equilibrum price is equal to $\frac{u_2-u_1}{3(u_2-u_1)}$, a price not only inferior to the pure monopoly price, but also the highest price compatible with the presence of the standard product on the market. Even setting a price equal to zero, there is no hope for the standard product to be sold to any customer. This surprising result is easily explained when the quality differential between the incumbent's variant and the variant penetrating the market is taken into account. The increase in quality justifies paying a strictly positive price for variant 2 even if the standard variant is made available at zero price! Notice finally that entry of the high-quality variant is always beneficial for the consumers. Not only is the set of possible purchases widened, but the price of the standard variable is lower when it can remain in the market!

3.3 Product Choice

Let us now consider the choice of quality by the entrant. We recognize here the same problem as in horizontal product differentiation. Is it more profitable to select a variant whose quality is close to the incumbent's quality or, on the contrary, to select a variant strongly differentiated so as to constitute a kind of "local monopoly"? When $R_1/R_2 > 1$, the answer is immediate. Since the entrant remains alone in the market after entry, with a demand equal to 1, its profits are equal to $\frac{u_2-u_1}{3(u_2-u_0)}$, a function monotone increasing in the quality u_2. Thus the monopoly chooses the highest possible quality. Now suppose that the reverse inequality holds, namely, $\frac{R_1}{R_2} \leq 1$. When the entrant selects a quality sufficiently close to the existing standard variant, the condition $0 \leq R_1/R_2 \leq \frac{u_2-u_1}{3(u_2-u_0)}$ is necessarily satisfied, so that the price equilibrium is located in case B in Figure 6.9: both sellers have a positive market share and the market is covered. One can easily show that, when u_2 tends to u_1, both equilibrium prices and profits simultaneously tend to zero. This is not surprising since, in the latter case, both variants tend to be homogeneous and we know since Bertrand that two homogeneous products have zero price at equilibrium. It is easy to show that the optimal choice for the entrant again consists of

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

choosing the highest possible quality, confirming the outcome already observed in the case of horizontal product differentiation. There is a strong incentive for competitors to differentiate their variants in view of relaxing price competition when products are too similar.

Finally, we point out two properties of price equilibria related to the parameters of the income distribution, R_1 and R_2. First notice that increasing the value R_1 from 0, without changing R_2, leads the equilibrium prices from domain A to B and, then, to C: the ratio R_1/R_2 increases when R_1 increases and R_2 remains constant. Considering the explicit values of the equilibrium prices, we conclude that increasing the value of R_1 from 0 first leads to a simultaneous increase in both equilibrium prices. Beyond some level, a further increase of R_1 drives the equilibrium prices in domain B, with the price p_1 of the low-quality product decreasing to zero and reaching this value exactly when the ratio R_1/R_2 is equal to 1. The decrease in p_1 follows because more and more consumers reject the standard variant, in proportion as the average income increases with R_1. At the limit, when R_1/R_2 is equal to 1, the price p_1 is equal to zero: even with a null price, the seller of the low-quality variant cannot retain any consumers! A higher value of R_1 then reverberates into an increase of p_2.

Changing the value of R_2 and, thus, the redistribution of income, shows more ambiguous effects. Starting with a high dispersion of incomes over the population, reducing the value of R_2 entails oscillatory behavior of equilibrium prices in domains A and B, until R_1/R_2 becomes equal to 1. Reducing further the parameter R_2 again leads to the exclusion of the standard product. It also leads to an increase in the price of the luxurious variant. Consequently a further reduction of R_2 can only decrease consumers' welfare. In spite of this surprising proposition, it shows that it is never optimal to promote a pure egalitarian income distribution. With weak heterogeneity in the income distribution, firms can more easily capture the consumers' surplus than they can when income is dispersed. When income is further dispersed, they must lower their price in view of capturing the surplus of the poorer consumers.

3.4 Vertical Differentiation and Entry: The Finiteness Property

In order to analyze how entry modifies the market, we shall assume that entry takes place by superior quality: the kth firm entering the industry sells a variant dominating all the existing ones in quality. This assumption is not unreasonable since it captures the fact that, being the latest to enter, it benefits from technological progress since the entry of firm $k-1$. The preceding analysis can then be extended to a differentiated oligopoly embodying n firms, ranked by order of increasing quality, variant k being of higher quality than variant $k-1$, namely, $u_k - 1 < u_k$ for all $k, k = 2, \ldots, n$.

Suppose now the entry of a $(n+1)$th variant of superior quality than u^n. Two situations can be observed at the new Nash equilibrium after entry. In the first, there are still some consumers who do not buy any of the variants available after entry: this situation corresponds to case A identified above in Figure 6.9. In spite of the entry of variant 2, the market is not yet fully covered: some consumers among the poorest prefer to refrain from consuming either of the two variants. In the second situation, the entry drives the market to be fully covered at the new equilibrium prices. Even the poorest customer buys a variant. We are then in the situations corresponding to case B or case C analyzed above. The first one corresponds to case B. The entry of firm $n+1$ does not cancel the share of any firm existing before entry: all of them keep a strictly positive market share. The second situation corresponds to case C. The entry of the $(n+1)$th firm entails a new price equilibrium excluding the lowest-quality seller. Even setting

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

a price equal to zero, the latter cannot avoid being excluded from the market by higher-quality variants. The seller $n+1$ selling the highest quality enters the market, but simultaneously excludes from it the seller producing the lowest quality. Everything happens as if the market would be too "narrow" to allow more than n firms to cohabit with a strictly positive market share. The entry of the highest quality must necessarily be accompanied with the exit of the lowest-quality firm. This result generalizes the the property already observed with $n = 2$ when the condition $R_1/R_2 > 1$ is satisfied: under this assumption, firm 2 drives out firm 1 from the market. It can be proved that, whatever the value of the ratio R_1/R_2, there *is always a maximal number of variants that can cohabit in the market with strictly positive market shares.* Moreover, this number is increasing with the degree of income dispersion. Finally, it follows that, under some conditions, the entry of a higher quality always leads to a decrease in the equilibrium prices of the variants already existing in the market. This property, which seems natural when the number of variants is increasing, is more surprising when the entry of a new variant is necessarily accompanied by the exit of the least-quality variant!

Thus we conclude that entry with vertically differentiated variants follows a very different mechanics from entry with identical or horizontally differentiated products. In the latter case, without barriers of entry and/or fixed costs, a large number of firms must be expected, and pure competition is naturally observed: each seller takes the price as given since he/she resembles a drop in the ocean. By contrast, entry in a vertically differentiated industry drives the market more or less rapidly to a "natural" oligopoly structure: according to the degree of dispersion, the market becomes overcrowded with a too large number of firms and a relatively small number among them can cohabit (finiteness property). A strategic behavior must be expected from the sellers. This context prevents perfect competition from being realized.

If tastes and incomes are little differentiated throughout the population, all individuals have almost the same income and agree to buy the high-quality variant at the reservation price of the poorest customer. By contrast, when income is more dispersed, richer consumers add more to the surplus when buying the high-quality product at a higher price than the reservation price of the poorest customer. Thus, in this case, the profit is higher at a higher price and some poorer customers remain unserved. Then a lower-quality firm can enter the market and sell its product to these poorer customers. Paradoxically, the weaker the income dispersion, the smaller the number of sellers at equilibrium, and the weaker their competition!

3.5 The Role of Production Costs

We have assumed so far that the average production cost of each variant with respect to quality is equal to zero, or to some constant equal for each variant. This is not realistic. Generally, the higher the quality of a variant, the higher its unit cost. Thus it seems natural to assume that there is a cost function $c(u)$ increasing with u, and representing the unit cost of the variant u. Moreover, let us assume that there is an interval $\left[u_*, u^*\right]$ of qualities where u_* (resp. u^*) represents the lowest (resp. highest) quality available. If the unit cost is increasing very quickly with quality, the gain obtained by increasing quality could be lost due to the increase in unit cost. In particular, it must be verified that the finiteness property still holds after unit costs are introduced. To this end, assume that each unit of variant u is sold at a price equal to its average cost $c(u)$. The utility level U of a consumer with income R choosing variant u is equal to $u(R - c(u))$. This equation can be rewritten as $uc(u) = uR - U$. Figure 6.10 represents the curve $uc(u)$ in the plane $(u, c(u))$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

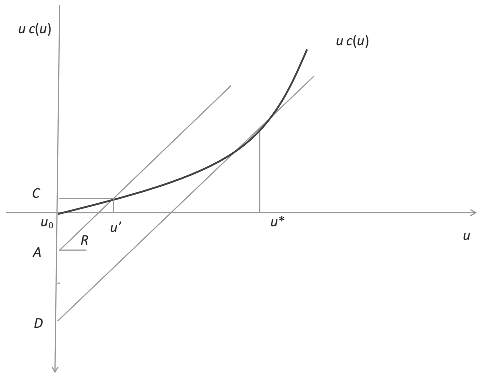


Figure 6.10 The role of production costs

Suppose that a consumer with income R decides to buy quality u', and consider the line passing through $(u', c(u'))$ with a slope equal to R. This line cuts the vertical axis at point A. By an elementary rule, the distance AC is equal to $u'R$. Since $C = u'c(u')$, we find that A is equal to $u'c(u') - u'R$. Consequently, the distance between point A and the origin u_0 measures the utility of a consumer with income R and buying one unit of variant u' at average cost $c(u')$. Suppose that the consumer chooses the variant leading to the highest utility. The optimal variant u^* maximizing utility is defined by the marginal condition $R = u^*c'(u^*) + c(u^*)$ where c' denotes the first derivative of c with respect to u. This optimal variant satisfying the first-order condition is represented in Figure 6.10: the utility obtained is equal to D.

The preceding analysis was performed for an arbitrary consumer with income R. In our analysis, incomes are spread over the interval [0, 1] representing the population of consumers. Let us examine now which variants are effectively consumed with the income distribution $R_1 + R_2t$,$t \in [0, 1]$. We first identify the variants consumed by the poorest and the richest consumers, with revenue R_1 and $R_1 + R_2$, respectively, if the price of each variant u would be given by $c(u)$. In Figure 6.11, the two lines corresponding to these levels of income are represented. The higher line corresponds to income level R_1, and the lower one to income level $R_1 + R_2$. The curve $uc(u)$ is also represented in Figure 6.11. To any income level $R_1 + R_2t$, $t \in [0, 1]$, there corresponds one line whose slope is in between the slopes of lines (a) and (b).

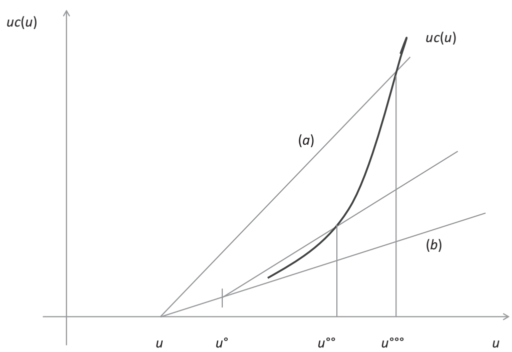


Figure 6.11 Production costs and income distribution

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Suppose first that the technologically feasible variants are between u and u°, as in Figure 6.11: only the branch of $uc(u)$ between these two values has to be taken into account. For all u in $[u, u^\circ]$, all consumers maximize their utility by selecting the highest quality since it is at u° that the distance from the origin is the largest for the line going through u, for all u in $[u, u^\circ]$. Furthermore, all consumers rank the variants in the feasible domain in the same increasing way. However, all consumers would have preferred to consume higher variants, but they are not available. Now suppose that the feasible domain is the interval $[u^\circ, u^{\circ\circ}]$: now only the section of $uc(u)$ between u° and $u^{\circ\circ}$ has to be taken into account. Then, at a price equal to marginal cost, the poorest consumer selects variant u° and the richest one variant $u^{\circ\circ}$. Furthermore, all consumers with an intermediate income choose a different variant. This variant is determined by the tangency between the line corresponding to income of this consumer and the curve $uc(u)$. Finally, suppose that the feasible variants correspond to the interval $[u^{\circ\circ}, u^{\circ\circ\circ}]$. Now all consumers rank the available variants in the same manner but in a decreasing way. Whatever the level of income, the farther away from $u^{\circ\circ\circ}$, the least it is desired if its price is equal to marginal cost. We are in the situation evoked above: if the unit cost increases too fast with quality, what can be gained in quality is more than proportionately lost by the increase in cost.

What can be said about the preceding analysis, due to Mussa and Rosen (1978)? The first case arises when the unit cost is constant (equal to zero) independently from the level of quality: at constant unit cost, all consumers rank the variants in the same increasing way. The second case is clearly different. At a price equal to average cost, all consumers prefer a different variant. This is similar to horizontal product differentiation: pricing at marginal cost allows the entry of an arbitrary number of firms and the finiteness property does not hold. Finally, in the third case, all variants are again unanimously ranked, but in a decreasing order: the increase in the unit cost, or the price, is stronger than the quality increase. Finally, only the first case leads to the finiteness property when costs are explicitly introduced and variants are priced at their marginal cost. To obtain this situation, it is necessary that the unit cost does not increase too rapidly with respect to quality. This condition is of course satisfied when the unit cost is assumed to be equal to zero whatever the quality.

The analysis has so far been conducted assuming that the variants are priced at their marginal cost. In fact, prices are set at the Nash equilibrium of the game, complicating the analysis a little more. In spite of these difficulties, Shaked and Sutton (1983) have found a necessary and sufficient condition for the finiteness property to hold at a Nash equilibrium in prices. This condition specifies that either the first-order condition $R = uc'(u) + c(u)$ is never satisfied by any consumer in the interval of qualities supplied, or, if there is such a consumer, he/she prefers to refrain from buying any variant and spends all of his/her income on other goods and services.

Let us summarize the main insights of our preceding paragraphs. They have revealed a new element to the case of vertical differentiation. Contrary to horizontal product differentiation, there might exist an upper bound on the number of variants who can cohabit in a vertically differentiated market. Generally, economists justify the presence of only a small number of firms by the existence of fixed costs or of barriers to entry. In a vertically differentiated industry, it is not necessary to evoke these arguments because the small number of firms can follow from the finiteness property: the entry of a new firm offering a higher quality product should be accompanied by the exit of the lowest-quality variant. This conclusion should, however, be tempered: in order to hold, the average cost should not increase too quickly with

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

quality. Otherwise, the finiteness property does not hold and an arbitrarily large number of firms can invade the market and realize the conditions required by perfect competition.

3.6 Vertical Differentiation and Spatial Competition

Hotelling proposed a spatial metaphor for analyzing horizontal differentiation. Similarly, we define a spatial version of the vertical differentiation model (Gabszewicz and Thisse, 1986). Let us suppose there is a linear market uniformly covered with consumers on the interval [0, 1] and willing to buy a unit of a homogeneous product. Nevertheless, differing from Hotelling, we suppose that the two sellers are located *outside* the market (see Figure 6.12, where s_i identifies the location of seller $i, i = 1, 2$). Also, all consumers have a transportation cost $c(x) = cx^2$ to move from their position on the line to the shops. Of course, all of them prefer to buy from the seller located at the closest end of the market if the sellers set the same price. Accordingly, the two variants are vertically differentiated since variant 1 located in s_1 is unanimously ranked above variant 2 located in s_2.

It is easy to find the price equilibrium in this model. Given a pair of prices (p_1, p_2), the consumer who is indifferent between buying variant 1 at price p_1 and variant 2 at price p_2 is located at the intersection of the two curves in Figure 6.12. The consumers located at the left of the latter buy from seller 2 and those located at the right buy from the more remote seller since they benefit from their proximity in terms of transportation costs. Formally, the marginal consumer is defined by the condition

$$p_1 + c(s_1 - x)^2 = p_2 + c(s_2 - x)^2.$$

Accordingly, profit functions are defined as

$$R_1(p_1, p_2) = p_1(1 - c(s_1 + s_2)/2 + (p_2 - p_1))$$
$$R_2(p_1, p_2) = p_2(c(s_1 + s_2)/2 + (p_1 - p_2)).$$

From the first-order conditions it is easy to derive equilibrium prices, namely,

$$p_1^* = (s_2 - s_1)(c(s_2 + s_1 + 2)/3)$$
$$p_2^* = (s_2 - s_1)(c(4 - s_2 - s_1)/3).$$

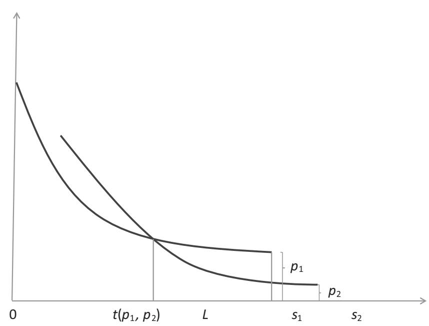


Figure 6.12 Vertical differentiation and spatial competition

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Notice, however, that if condition $4 < s_2 - s_1$ is not satisfied, the price of the most remote firm becomes negative. Then, its price is given by zero: even by setting a zero price, the most remote firm is unable to attract any consumer to buy the product at his/her shop.

How do the sellers choose their location? Substituting p_1^* and p_2^* in the profit functions given above, we obtain the payoffs of the sellers expressed as a function of their location at equilibrium prices corresponding to these locations. The resulting Nash equilibrium obtains when firm 1 selects to locate as close as possible to the market, i.e., $s_1^* = 1$, while firm 2 chooses its location by maximizing its profit under the constraint that $s_1^* = 1$. The next section is devoted to the problem of nesting the two forms of differentiation: horizontal and vertical.

4 NESTING HORIZONTAL AND VERTICAL DIFFERENTIATION

Under horizontal differentiation, when tastes are heterogenous, it is possible to segment the population of consumers according to their preferred variant. We may then associate to each variant the group of consumers who prefer that variant over the other and define this group as its natural market. In the extreme case of vertical product differentiation, the natural market of one firm consists of the whole market while the other has a zero market share at equal price. By contrast, horizontal differentiation accommodates a very large class of the natural markets' configurations. For instance, in the case of spatial competition à la Hotelling (Hotelling, 1929), when firm 1 is located at the left extremity of the linear market while firm 2 stands at the other extremity, the market does not view one firm as more "desirable" on average than the other since both natural markets are exactly of equal size.

At the other extreme, firms can be located in the linear market in such a manner that almost all consumers would prefer to buy from one of the two firms, in spite of the fact that these firms set the same price. This would be the case, again in the classical Hotelling location model, when firm 2 is located at the right extremity of the linear market while firm 1 now stands very close to it. Due to transportation costs, almost all consumers buy from firm 1 when it quotes the same price as firm 2. Thus, this situation corresponds very closely to the definition of vertical differentiation, even if, *sensu stricto*, it should fall into the alternative category. In all such hybrid cases, and although, differently from vertical differentiation, no variant holds a definite advantage over the other when horizontally differentiated, one may argue that the firm counting a larger number of consumers in its natural market should somehow benefit from a larger market power: on average, the market views this variant as more desirable. A first natural question therefore comes to mind: to what extent do differences in natural market sizes translate into different equilibrium market valuations for the product? Another natural question is whether, when the size of the natural market of a particular firm tends to the size of the whole market, the corresponding equilibrium prices tend monotonically to the prices prevailing at equilibrium of the corresponding vertical product differentiation market. Such a conclusion would then allow nesting vertical and horizontal product differentiation models in a natural way. We develop hereafter a duopoly model that addresses these questions in a precise way. To this end, we adapt the canonical Hotelling model to allow for natural markets of different sizes. In the symmetric linear model with firms located at the extremities of the unit interval, natural markets are defined by the $[0, 1/2]$ and $[1/2, 1]$ intervals, respectively. In order to allow for natural markets with different sizes, we then assume that the density differs from one interval to the other. Notice that, in this model, a vertical configuration appears as a

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

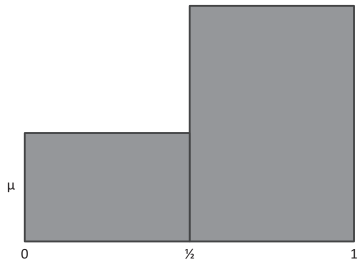


Figure 6.13 Horizontal and vertical differentiation

limiting case where the density of one of the intervals tends to zero while the density of the other tends to 1. In this setup, we show that equilibrium prices display two key properties: first, the level of prices at equilibrium decreases with the disparities in the natural market size. Second, the equilibrium price differential increases with the disparities. In other words, the more unequal in size the natural markets, the fiercer the price competition.

Let [0, 1] be the set of types of consumers and consider two firms, the first one being located at point 0 and the other at point 1. The density over the types in $T_1 = \left[0, \frac{1}{2}\right]$ is equal to μ and to $1-\mu$ over the types in $T_2 = \left[\frac{1}{2}, 1\right]$. Figure 6.13 illustrates this situation.

The preferences of consumer of type x in [0, 1], are defined by

$$U(x) = S - tx - p_1$$

when the consumers buy from firm 1 (firm located at point 0), and by

$$U(x) = S - (1-t)x - p_2$$

when the consumers buy from firm 2 (firm located at point 1), with S denoting the absolute reservation price and with t denoting the unit transportation cost and p_i the price set by firm $i, i = 1, 2$. Notice that, at equal prices, all consumers in T_1 prefer to buy from firm 1 while all consumers in T_2 prefer to buy from firm 2 so that, when $\mu < 1/2$, there is a majority of consumers who prefer buying from firm 2 than from firm 1 at equal prices, and vice versa when $\mu > 1/2$. In particular, when $\mu = 0$, all consumers prefer buying from 2, which corresponds to the (extreme) case of vertical differentiation and, when $\mu = 1/2$, we obtain the (opposite extreme) case of symmetric horizontal differentiation. For different values of μ, we get hybrid cases of horizontal product differentiation, with a majority preferring to buy from firm 2 (resp. firm 1) than firm 1 (resp. firm 2) according as $\mu < 1/2$ (resp. $\mu > 1/2$). In the following, we normalize the transportation cost t by putting $t = 1$. We shall also assume without loss of generality that $\mu < 1/2$, so that there is a majority of consumers who prefer to buy from firm 2 than from firm 1 at equal prices. Finally, we assume that the constant S is large enough to guarantee that the market is covered.

Let $x(p_1, p_2)$ be the solution to

$$S - x - p_1 = S - (1-x) - p_2,$$

namely, $x(p_1, p_2) = \frac{1}{2}(p_2 - p_1 + 1)$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Notice that, if $p_1 < p_2$, we have $x(p_1,p_2) > 1/2$, so that the interval $T_1 = [0,1/2]$ is included in the set of consumers who buy from firm 1 at prices (p_1,p_2). To this set, one must add the interval of types $[1/2, x(p_1,p_2)]$ corresponding to consumers who prefer to buy from firm 2 than from firm 1 at equal price, but who prefer to buy from firm 1 at prices (p_1,p_2). Consequently, the demand $D_1(p_1,p_2)$ to firm 1 at prices (p_1,p_2) with $p_1 > p_2$ obtains as

$$D_1(p_1,p_2) = \frac{\mu}{2} + (1-\mu)(p_1 - p_2).$$

Now, if $p_1 > p_2$, the point $x(p_1,p_2)$ is located at the left of $1/2$ and

$$D_1(p_1,p_2) = \frac{\mu}{2}(p_2 - p_1 + 1).$$

Notice that, since $\mu < 1/2$, the demand function of firm 1 is a linear convex with a kink at $p_1 = p_2$. Therefore, the revenue function might not be concave in own price. The demand $D_2(p_1,p_2)$ and revenue functions for firm 2 are easily derived as

$$D_2(p_1,p_2) = \frac{1-\mu}{2} + \frac{\mu}{2}(p_1 - p_2)$$

when $p_1 < p_2$, and

$$D_2(p_1,p_2) = \frac{1-\mu}{2}(p_1 - p_2 + 1)$$

when $p_2 > p_1$. The corresponding revenue function is concave in own price. We first identify the unique candidate pure-strategy equilibrium. Given the firms' best replies, it is first clear that there is no symmetric pure-strategy equilibrium. Additional computations directly show that there is no asymmetric equilibrium in which firm 2 quotes the lowest price. The only remaining price equilibrium candidate is such that $p_1 < p_2$. Combining the corresponding best reply functions, one obtains

$$p_1^* = \frac{\mu+1}{3(1-\mu)};\ p_2^* = \frac{2-\mu}{3(1-\mu)}.$$

It is then easy to check that $p_1^* < p_2^*$ holds if and only if $\mu < 1/2$. In other words, the equilibrium candidate indeed yields the desired price hierarchy. Finally, one can show that the price candidate p_2^* is located at the right of the point where the best-reply function of seller 2 is discontinuous, which guarantees the existence of (p_1^*, p_2^*) as a price equilibrium.

We may consider now the equilibrium prices corresponding to the "extreme" situations in which either all agents prefer to buy from firm 2 than from firm 1 at equal prices ($\mu = 0$: vertical product differentiation), or half of them prefer to buy from firm 1 and half of them from firm 2 under the same condition ($\mu < 1/2$: symmetric horizontal product differentiation). In the first case, $p_1^* = 1/3; p_2^* = 2/3$. In the second case, we notice that when $\mu < 1/2$, equilibrium prices are equal to each other and equal to one. We also notice that, when μ tends to zero, the model gets closer and closer to a situation of vertical differentiation, in which a larger and larger majority prefers variant 2 to variant 1 ($1-\mu$ tends to 1), and the corresponding equilibrium prices converge to the equilibrium prices in the limit model.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

It is easy to check that the equilibrium analysis covering the case when $1/2 > \mu > 1$ is, *mutatis mutandis*, identical to the preceding one: firm 1 now plays the role of firm 2 in the definition of demands and profits, firm 2 now selling the variant preferred by the majority. Additional properties of the equilibrium are worth mentioning, namely, the equilibrium price differential decreases with μ, while absolute price levels both increase with μ. In other words, a larger symmetry in the population's tastes means that the natural market of firm 2 gets bigger, which implies that it is less attractive to challenge the other's natural market in relative terms. Market valuations of the product tend to reflect the distribution of tastes among variants in the population. Equilibrium market valuations reflect the disparities in natural market sizes.

5 APPLICATIONS

In the following, we provide two applications of the above approach to product differentiation via models based on characteristics. The first deals with externalities, both simple and cross-network externalities. Simple networks arise when the utility of consumers for a good depends on the number of its users, like in the case of a communication network. Cross-network externalities are observed in several markets when the utility of the consumers in one market depends on the number of consumers in another, and vice versa. This the case, for instance, in the competition between bank card companies when the utility of merchants for accepting a credit card depends on the number of cardholders and the utility of subscribing a credit card to a bank card company depends on the numbers of merchants who accept it.

The second application is concerned with *social preferences*: people are conscious of their peers and the social image they convey to them. Two noticeable approaches to this social attitude are the "other-regarding preferences" (Andreoni, 1990, Andreoni and Miller, 2002) on the one hand, and "identity utility" (Akerlof, 1997) on the other. In the former approach, the individual choice is driven by the others' payoff. In the latter, the individual choice is shaped simultaneously by the desire to satisfy material needs *and* to comply with some normative ideal. Of course, when these drivers are in conflict, the actual choice depends both on the disutility of deviating from one's normative ideal and the cost of sacrificing some material needs to comply with the norm. Rather than being mutually exclusive, we combine these approaches of other-regarding preferences and identity utility by assuming that people do not pursue an absolute egoistic goal, but a *social position* among peers. This position is obtained through *goods* that provide their buyers with some social/psychological benefits beyond the material needs that products traditionally satisfy. In particular, these benefits can enable the buyers to differentiate themselves from other consumers, thereby satisfying their *vanity*. Otherwise, consumers can decide to purchase some brands to satisfy a *conformity* desire. Whatever the specific driver of this social behavior, consumers' choice is guided by a social and/or psychological incentive in addition to the *individual rationality-based motive.*

5.1 Network Effects

5.1.1 Simple network effects and product differentiation

In numerous situations, the demand for a product is correlated with the number of its users: for example, the utility of a cell is directly related to the size of the communication network in which it operates. The existence of such network externalities is a privileged

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

field of application of product differentiation theory. We shall limit our analysis to the case of horizontal product differentiation combined with the existence of positive consumption externalities. The analysis easily extends to vertical product differentiation and negative externalities (like congestion effects where the utility of consuming a product decreases with the length of the queue).

As in the Hotelling model, consider two firms selling a homogeneous product at zero cost to customers uniformly distributed along the unit interval. Each customer can buy a unit of the product either from seller 1 located in x_1, $x_1 \in [0, 1]$, with x_1 measuring the distance between the selling point and the left extremity of the interval, or from seller 2, located at a distance equal to x_2 from the same extremity. We suppose that customers share the same reservation price for one unit of the good, K. Moreover, each customer incurs a transportation cost proportional to the square of the distance to the seller he has selected for buying the good from.

This model differs, however, from the traditional model because it is assumed that consumers' utility now not only depends on the intrinsic value of the good (stand-alone value), but also on the number of buyers n_i of seller $i, i = 1, 2$ (network value).

More precisely, it is supposed that the utility of customer x of buying from seller i located at x_i is equal to $K - p_i - t(x - x_i)^2 + an_i, i = 1, 2$, where n_i represents the number of customers buying from seller i at price p_i. If the parameter a is positive (resp. negative), the network externality is positive (resp. negative). We suppose $a > 0$ (positive externalities). Also notice that the larger the value of a, the stronger the network effect since the externality increases with a. Given a pair of prices (p_1, p_2), two situations can appear. Either there is a consumer $t(p_1, p_2)$ in the interval $[0, 1]$ who is indifferent between buying at this pair of prices from either seller, in which case all consumers to the left of this consumer buy from seller 1 and those to his/her right from seller 2. Or there is no such consumer, in which case all consumers buy from a single seller. The first situation takes place when there is a value x in the interval $[0, 1]$, solving the equation

$$K - p_1 - t(x - x_1)^2 + an_1 = K - p_2 - t(x - x_2)^2 + an_2,$$

with n_1 and n_2 representing the expectations of buyers concerning the size of demand addressed to seller 1 and seller 2, respectively. We assume that these expectations are self-fulfilling: the sizes of demand expected by the buyers at the pair of prices (p_1, p_2) are corresponding to their effective sizes. Then, we may replace in the above equation the values n_1 and n_2 with the sizes x and $1 - x$, respectively, and we get the following equation to be solved for x, namely,

$$x = x(p_1, p_2) = p_1 - p_1 - a/2(t(2 - x_1 - x_2) - a).$$

This value belongs to the interior of the interval [0, 1] if, and only if, the two conditions

$$p_1 - p_2 > a - t(x_2 - x_1)(2 - x_1 - x_2)$$
$$p_1 - p_2 < t(x_2 - x_1)t(x_2 + x_1) - a$$

are simultaneously satisfied. The domain of a pair of prices for which these two conditions are simultaneously satisfied differs from the empty set if and only if the inequality $t(x_2 - x_1) > a$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

is satisfied. This condition states that the cost of the distance between the two sellers must exceed the intensity of the network effect. Now suppose that these conditions are not satisfied. Then, there is no solution in x to the preceding equation. Either the first equation is violated and the value of $x(p_1,p_2)$ is equal to 1 (and $1 - x(p_1,p_2) = 0$); or the second inequality is not satisfied and $x(p_1,p_2) = 0$ (and $1 - x(p_1,p_2) = 1$). Accordingly, given a pair of prices (p_1,p_2), three situations can be realized concerning the sharing of the market between the two sellers. In the first one, $x(p_1,p_2)$ belongs to the interior of the interval $[0,1]$ and the two sellers each have a strictly positive market share. In the two alternative cases, one of the sellers eliminates its competitor from the market.

5.1.2 Simple network effects and price equilibrium

It remains to determine the equilibrium prices and, in particular, under which conditions both sellers have a strictly market share or, on the contrary, one of them is eliminated from the market. When $x(p_1,p_2)$ belongs to the interior of the interval $[0,1]$, the demand from seller 1 is then equal to $x(p_1,p_2) = p_2 - p_1 - a/2(t(2 - x_1 - x_2) - a)$ and the demand of seller 2 $x(p_1,p_2) = 0$, with the corresponding profits obtained by multiplying these market shares by the corresponding prices. If the price equilibrium is interior to the admissible domain $[0,1]$, it is easy to determine the equilibrium values by using, as usual, the first-order necessary and sufficient conditions, leading to the equilibrium values:

$$p_1^* = \frac{t}{3}((x_2 - x_1)(2 + x_1 + x_2)) - a$$

$$p_2^* = \frac{t}{3}((x_2 - x_1)(4 - x_1 - x_2)) - a.$$

Substituting these values in the value of $x(p_1,p_2)$ given above and assuming $x_1 + x_2 > 1$, we notice that $x(p_1,p_2)$ effectively belongs to the interior of $[0,1]$ if, and only if, the condition $a \leq \frac{t}{3}(x_2 - x_1)(4 - x_1 - x_2)$ is satisfied. In the reverse case, only one seller covers the market with its sales. Suppose it is seller 1. Then, the equation $p_1 - p_2 > a - t(x_2 - x_1)(2 - x_1 - x_2)$ is violated, so that the reverse of this inequality must hold at equilibrium. Then, the price equilibrium p_1^* obtains as the highest price allowing seller 1 to serve the market ($p_1^* = a - \frac{t}{3}((x_2 - x_1)(2 - x_1 - x_2))$) guaranteeing that seller 2, even setting the lowest price $p_2^* = 0$, cannot serve any share of the market. At the equilibrium prices, both sellers find it profitable to serve the market if and only if the network effect is sufficiently low, namely $a \leq \frac{t}{3}((x_2 - x_1)(4 - x_1 - x_2))$.

When both sellers enjoy a positive market share at equilibrium we notice that equilibrium prices are smaller than they would be without network effects. Also we notice that equilibrium prices decrease with the size of the network effects as measured by a. The existence of network effects increases the elasticity of demand with respect to price, increasing the incentive of sellers to lower their price to conquer the rival's customers. When the above inequality is not satisfied, only one seller serves the whole market. Accordingly, network effects can engender a situation where, even though frequently observed in vertical differentiation, cannot arise under horizontal product differentiation without such effects, one of the sellers bars entry to its competitor at equilibrium. This can be explained as follows. With network effects, the good is desired not only for its intrinsic characteristics, but also for the number of its consumers. Accordingly, the seller with a larger market share can attract a larger number of consumers,

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

even if the product quality is the same. Thus equilibrium determines in an endogenous way the "quality" of the product by the size of the market of each seller at equilibrium.

5.1.3 Cross-network effects and product differentiation

In the preceding analysis, we assume that the externalities are induced by the size of demand in the industry considered. It happens, however, that externalities develop from one industry to another, when the utility for a product depends on the size of demand for a product developed in another. Thus, we deal with *cross*-network externalities. There are many examples leading to this situation. For instance, consider shopping malls. The larger the number of different shops in a shopping mall, the more attractive it is to visit. But, conversely, the larger the number of visitors, the more attractive for a merchant to rent a shop in the shopping mall! Another example is provided by credit cards. The utility of owning a credit card depends on the number of merchants who are willing to accept it as a means of payment. Conversely, the utility for a merchant to pay a subscription to the bank card company to be able to accept the credit card depends on the number of cardholders who are willing to pay with a credit card. Such a situation is called a *two-sided market.* If the two-sides market are men on the one hand and women on the other, the interest in subscribing to a dating website depends for men on the number of women and, for women, on the number of men who have subscribed to the website. Finally, a last example is provided by the market for media and the market for advertising, both linked by cross-network externalities. For instance, the readers of a newspaper are sensitive to the ads placed in it. In particular, ads are interesting to those who try to buy a second-hand car or rent an apartment. As for the advertisers, they are interested in the number of readers of the newspaper since the more readers, the higher the probability that the ad will be noticed.

All these situations give birth to new opportunities to all firms working at the interface of these markets (we call them *platforms* in the following). One intuitively understands that platforms play the major role of facilitating transactions between the buyers of some products and their sellers, transactions that could not have been realized without the platform's intervention. "Embarking" buyers and sellers on the platform makes possible transactions that would not be possible without their existence.

To model the preceding situations, we use a model combining two distinct models of vertical product differentiation, one for each side of the market. The first one constitutes bank card companies producing credit cards and selling them to customers, while they make available to merchants the right to accept them as a means of payment in exchange for paying a subscription. Then the second group constitutes consumers, paying their credit card bill to bank card companies and using it to pay for their purchases to the merchants. Finally, the third group comprises the merchants. They acquire from the banks the machine allowing them to accept the credit card as a means of payment and accept the payment for purchases by consumers using their credit card. In this representation, bank card companies are the platforms of the two-sided market: consumers the demand on one side and merchants the demand on the other side. The role of platforms consists in facilitating the transactions between the agents on the two sides. To simplify we shall assume there are only two platforms, for instance Visa and Mastercard.

We assume that the set of consumers is represented by the interval $[0, 1]$. The utility of a consumer t in this interval is measured by the expression $tx_i - p_i$, where x_i represents the number of merchants accepting the credit card of type i and p_i the price to be paid at the bank card company to hold a credit card of type $i, i = 1, 2$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

It is important to know whether a consumer has one or two credit cards and if merchants accept one or two of them. In the first case, one speaks of *single-homing* agents and *multi-homing* in the second. In cross-network externalities multi-homing is often realized. A large number of agents own several credit cards and merchants also accept several of them; also many consumers subscribe to several dating sites. In the credit card case, and supposing the number of the merchants accepting credit card 2 exceeds the number of those accepting credit card 1 ($x_2 > x_1$), it is reasonable to define the utility of consumer t if he owns both cards as equal to $tx_3 - p_1 - p_3$, with x_3 a number belonging to the interval $[x_2, x + x_3]$. When $x_3 = x_2$, it means that all merchants accepting card 2 also accept card 1 so that this latter is perfectly useless. At the other extreme, when $x_3 = x_2 + x_1$, it means that no merchant accepts the two cards simultaneously.

Finally we suppose that the set of merchants is also represented by the interval $[0, 1]$, the merchant v in this interval having a utility measured by the expression $\alpha_i v - s_i$, with α_i representing the number of consumers who own the credit card of type i and s_i the price imposed by the platform to accept payments of consumers by means of card $i, i = 1, 2$. In the case of merchants' multi-homing, we shall assume that the utility of a merchant v of accepting two cards is equal to $\alpha_3 v - s_1 - s_2$, where α_3 belongs to the interval $[\alpha_2, \alpha_3 + \alpha_2]$ when $\alpha_2 > \alpha_3$.

The intuition underlying the above model is as follows. From the viewpoint of the customers, the platforms appear as two sellers on a differentiated market, the seller having constituted the larger network on one side of the market looking like selling a product of higher "quality" to the customers of the other side. The hierarchical value on the quality ladder reflects the size of their respective networks: the larger the network on one side, the higher the quality of this "product" for the other side. Accordingly, the cross-network effects in this model appear as generating two parallel differentiated markets according to the price they choose on each side. Furthermore, the quality of the variants sold on each side of the market is endogenous and determined by the price each platform chooses at the price equilibrium. Notice that the platform can, by lowering its price on one side of the market, make the product more attractive to the buyers operating on the other side, thereby allowing it to set a higher price on the latter! Thus one can expect the existence of asymmetric equilibria such that, while the price of the product on one side is very low, the price set on the other side is, on the other hand, very high.

5.2 Product Differentiation and Environmental Economics

A natural entry point for combining social preferences with product differentiation is environmental economics. The theoretical framework of product differentiation has often been applied by the literature in environmental economics to the analysis of equilibrium configurations, environmental policy and firms' behavior in markets with polluting goods (Moraga-Gonzales and Padron-Fumero, 2002). The starting point of this literature is that *consumers are environmentally concerned.* This concern can be driven by *traditional* and *self-regarding preferences*. Then, when caring about environment, people care about their own health and safety, thereby being guided by purely private motives. Otherwise, this concern for environmentally friendly goods can be justified on different grounds, based on the judgment that *human actions are driven both by individual rationality and social concern.* For example, people can purchase green products to avoid a bad conscience. In this

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

case, consumers have some degree of idealism and take into account their possible negative contribution to the environment when purchasing brown goods. Also, their willingness to pay a price premium can be driven by a positional concern: consumers do not pursue an absolute end when going green, but a social status or a relative position among peers. In this perspective, the traditional *self-regarding preferences* no longer suffice to formalize consumption behavior: this opens the door to new ingredients in the utility function. Hereafter, we briefly present two models nesting this environmental concern in the utility function.[2] The first model is borrowed from the literature on horizontal product differentiation and assumes that, although consumers' individual decisions are based on utility-maximizing behavior, there may be a trade-off between the utility immediately derived from the preferred characteristics of a product and a moral constraint to buy environmentally friendly goods. Whenever a consumer purchases a brown product, she might incur a bad conscience and thus feel guilty. This moral guilt can be fostered by friends, parents, partners, or by the media. Since this guilt is internalized by consumers, the traditional utility function is enriched by some moral/social externality indirectly affecting the optimal characteristics of products.

Then, we consider a model of vertical product differentiation. In this work, the typical utility function à la Mussa-Rosen is modified with the aim of representing consumers that seek a relative position among peers when purchasing products. This position is obtained through *conspicuous goods* providing their buyers with some social benefits beyond the material needs that products traditionally satisfy. The higher the quality of these goods, namely their ranking along *the quality ladder*, the higher their social value and the corresponding position they confer to the buyer along the *social ladder*. Accordingly, utility function depends on the characteristics of products a consumer buys compared with those bought by his/her peers.

5.2.1 Horizontal product differentiation and social concern

The model we present in this section is inspired by Conrad (2005) and assumes that consumers are characterized by some moral concern, taking the form of a social externality. When making their decision about consumption, consumers value both the intrinsic characteristics and the green nature of the good. When the former satisfy the consumers, but the latter is missing, they can suffer social disapproval if they decide to buy the (brown) product. In this circumstance, the consumption behavior is driven by the solution to the dilemma between privileging the intrinsic characteristics of the product, thereby incurring social stigma, or sacrificing them and refraining from buying the polluting good. The duopoly theory developed hereafter aims at providing a solution to this dilemma.

The market is populated by two firms selling a heterogeneous product with characteristics $q_i \in [0, 1]$, $i = 1, 2$. On the one hand, goods are assumed to be horizontally differentiated. Horizontal product differentiation is formalized through a spatial duopoly model. On the other hand, they are labeled in increasing order of environmental friendliness and $q_1 \leq q_2$ so firm 1 is less concerned about not producing environmentally friendly goods compared with firm 2. There is a continuum of consumers uniformly distributed over the interval $[0, 1]$ and each

[2] The most frequently used utility function in these works is inspired by Mussa and Rosen (1978) (see also Ben Elhadj, Hili and Lahmandi-Ayed, 2013 and Coestier and Marette, 2004) so that the *self-regarding preferences* formalized in the traditional framework of vertical product differentiation are not questioned.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

consumer buys one unit of the product. The utility of consumer $\theta \in [0,1]$ for a unit of the good q_i is given by

$$U(q_i, \theta) = r - t(q_i - \theta)^2 - d(1 - q_i) - p_i$$

where r denotes the gross utility a consumer derives from consuming one unit of the product. The term $t(q_i - \theta)^2$ represents the costs a consumer, located at $\theta \in [0,1]$, faces when buying the good q_i, with p_i being the corresponding price. Finally, the new ingredient $d(1 - q_i)$ captures the social concern of the consumer. The consumer knows he should buy the most environmentally friendly product at the end of the characteristics' $[0,1]$ line. Still, his/her preferences induce him/her to buy a different good, namely q_i. Because of this, the consumer suffers a social stigma that increases with the distance between the ideal green good and the one actually bought.

In a two-stage game, either firm chooses the product's characteristic q_i at the first stage and then, at the second stage, the corresponding price. In producing the two characteristics, costs increase in q.

The aim of the analysis is to identify the effect, if any, of environmental awareness on the equilibrium configuration. In the traditional duopoly models of horizontal product differentiation, the two firms choose different locations (and thus differentiated goods) and both gain positive profits at the unique equilibrium. The key question here is whether under social externalities, the above finding still holds, namely whether (i) the existence of social concern among consumers induces firms to change location and thus price at equilibrium, with immediate consequence on the optimal profits and (ii) the equilibrium is unique. Conrad shows that when the environmental concern is strong compared with cost, firm 2 finds it optimal to produce $q_2^* = 1$ while firm 1 chooses $q_1^* > 0$. Under weak environmental concern, firm 1 optimally selects $q_1^* = 0$ while $q_2^* < 1$. Finally, in the case when the environmental concern is not very relevant so that the cost component can (weakly) prevail, then $q_1^* = 0$ while $q_2^* = 1$. The optimal characteristics of goods is unambiguously depending on the interplay between social concern and costs, the former ingredient favoring the production of a green good in terms of a higher market share and higher profits, the latter discouraging it. This statement immediately justifies the equilibria observed with the environmental concern is either extremely strong or weak, compared with costs. In the third scenario with a balanced environmental/cost situation, both firms choose maximal product differentiation since the incentive to improve the environmental characteristics of goods is countervailed by the costs incurred for this improvement.

5.2.2 Vertical product differentiation and relative preferences

In the duopoly theory presented hereafter, consumers' satisfaction is described by relative preferences. Formally, variants in the market are differentiated along two dimensions: hedonic quality and environmental quality. The hedonic dimension refers to the pure (intrinsic) performance of the good, whereas the latter dimension has a positional content: buying green goods satisfies the consumer's desire to be portrayed as a socially worthy citizen and thus provides some social rewards. In the case of brown consumption, people feel socially disapproved of, thereby incurring social *stigma*. As a by-product of this, whenever green purchases are used to obtain social approval or to escape from social *stigma*, the social benefit (resp. social punishment) from a pro-environmental (resp. brown) behavior is larger, the larger

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

the quality gap between the green and the brown variants. In line with this approach, the utility function of the consumer incorporates a new term depending, for each variant, on the relative position it can confer to the buyers, namely the environmental quality gap between "adjacent" variants.

It is worth noting that in some sectors a trade-off between hedonic and environmental quality is unavoidable as certain polluting goods meet consumers' requirements better than the green alternatives (Gupta and Ogden, 2009). For example, conventional internal combustion engine vehicles, although dominated by green alternatives in terms of polluting emissions, are still superior in most cases to electric or hybrid vehicles based on pure performance. Paper produced from trees instead of recycled paper is often preferred because it is softer to the touch. New-generation washing machines have energy-saving cycles labeled "green" or "eco"; they are, however, more time consuming in comparison with ordinary cycles. Nevertheless, in other sectors, such as cosmetics, household goods and sometimes food, a high hedonic quality can be obtained without sacrificing the corresponding environmental quality. In this circumstance, a variant can be unambiguously viewed of higher quality with respect to the alternative good since it meets the high hedonic standards valued by a consumer while conferring to her a high-ranked social position along the quality ladder.

The following analysis relies on Mantovani, Tarola, and Vergari (2015) and considers the former scenario where a trade-off between hedonic and environmental quality arises, so that a high hedonic quality variant is more polluting than a low hedonic quality alternative. Firms are assumed to compete in prices, the quality of variants being exogenously given.

Consider a vertically differentiated duopoly with two variants of the same good. The performance of the variant i, with $i = L, H$, determines its *intrinsic* or *hedonic quality* q_i. Accordingly, $q_H > q_L$: variant H has a higher intrinsic quality than variant L so that q_H is ranked higher than q_L along the *hedonic quality ladder*. Nevertheless, variant H generates more polluting emissions e_H per unit of production than variant L. Accordingly, variant q_H represents the brown good. Conversely, variant L is considered green because its emissions e_L are lower than those deriving from variant q_H, namely $e_L < e_H$. The *environmental quality* of L is then higher than that of H. The ranking between q_H and q_L is reversed when considered in terms of this latter dimension of quality: variant q_L is ranked higher than variant q_H along the *environmental quality ladder*. There is a continuum of consumers indexed by θ and uniformly distributed in the interval $[0, b]$ with density $1/b$. Formally, the indirect utility of consumer type θ is written as:

$$U(\theta) = \begin{cases} \theta q_H - p_H - \gamma\,(e_H - e_L)\,, \text{ if she buys the high-quality good,} \\ \theta q_L - p_L + \gamma\,(e_H - e_L)\,, \text{ if she buys the low-quality good,} \\ 0, \text{ if she refrains from buying.} \end{cases}$$

A further ingredient, namely $\gamma\,(e_H - e_L)$ with $i \neq j$, is added to the traditional component of the indirect utility function $(\theta q_i - p_i)$ such that the satisfaction of buying a product can be either magnified or dampened by the environmental characteristics of variant i, as compared with j. In line with the idea that green consumption has a positional driver, it is the *relative* pollution emitted by either good – captured by the term $\gamma\,(e_H - e_L)$ – that affects consumers' utility. Parameter $\gamma \geq 0$ measures the intensity of the relative dimension of consumption; the higher the value of γ, the stronger the relative (or social) preferences with respect to

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

the hedonic ones. From the above formulation of the utility function, the consumer who is indifferent between buying the low-quality good and not buying at all is given by θ_L, namely,

$$\theta_L = \gamma + \frac{p_L - \gamma q_H}{q_L} = \frac{p_L - \gamma(q_H - q_L)}{q_L}, \tag{6.1}$$

with $\theta_L > 0 \iff p_L > \gamma(q_H - q_L)$. The consumer indifferent between buying the low-quality good and the high-quality good is:

$$\theta_H = 2\gamma + \frac{p_H - p_L}{q_H - q_L}. \tag{6.2}$$

The introduction of other-regarding preferences is such that the traditional mechanism of competition between firms observed in a traditional vertically differentiated market holds *sensu stricto* as long as all consumers agree that the hedonic quality of a product is more valuable than its green nature (or the reverse). However, if a product is better than the alternative based on one characteristic but worse based on another one, then the defining property of vertical differentiation may cease to hold. This occurs whenever some consumers give more value to the hedonic dimension of a variant while others privilege its environmental quality so that, at the same price, both variants face a positive demand. Although this feature resembles the defining property of horizontal differentiation, the asymmetry between firms stemming from the two dimensions of product differentiation is such that, when the defining property of vertical differentiation stops holding, competition falls into a hybrid category where, at equal prices, both variants have a positive demand (horizontal differentiation) but at equilibrium their prices do not coincide (vertical differentiation).[3]

This hybrid category of product differentiation opens the door to several results that are not observed in a traditional vertically differentiated market. In particular, depending on b and γ driving consumers to the hedonic and environmental dimensions, respectively, one can characterize two relevant parametric regions. For *relatively low values* of b, both firms are active at interior equilibrium when γ takes intermediate values, while only the green (resp. brown) firm is active in the market when γ is sufficiently high (resp. low). Further, in the case when the market is monopolized by the green firm, it finds it profitable to cover the whole market if γ is extremely relevant. Conversely, for *relatively high values* of b, the brown producer never monopolizes the market. Moreover, under duopoly, the green firm can find it optimal to cover the market. This happens when γ is relatively high.

It is worth noting that, when both firms are active at equilibrium, there are circumstances in which the price of the green good (the low hedonic quality product) ceases to be lower than the price of the competing variant (*duopoly with price switch)*. This finding can never be observed under self-regarding preferences. Along the same rationale, in a traditional model of vertical product differentiation, a monopoly configuration can hold only under a restrictive assumption on consumer heterogeneity (natural monopoly). Further, at this configuration, the firm is typically induced to cover one-half of the entire population of consumers. In Mantovani et al. (2015), both the green monopolization of the market and its coverage are observed at

[3] There is a strand of literature considering different sources of product differentiation. For example, Gabszewicz and Thisse (1986), Neven and Thisse (1990), and Gabszewicz and Wauthy (2012) combine horizontal and vertical differentiation.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

equilibrium whenever the relative preferences are sufficiently strong that the social component of consumption prevails over the hedonic driver. A final remark deserves some attention. When the social component of consumption γ is high with respect to the willingness to pay for the hedonic quality b, then the green firm can monopolize the market. Still, the monopolization by the brown competitor does not arise when b is high compared with γ. So, one can easily conclude that the social component of consumption can benefit the green firm more than the willingness to pay for the hedonic quality can benefit the brown producer, thereby creating asymmetric effects on the equilibrium configurations.

It is worth noting that, if the two dimensions of product differentiation were not in contrast with each other, then the equilibrium analysis would change. If the green products were also of high hedonic quality, then one would never observe the price switch or the market coverage scenario. Indeed, the initial disadvantageous position of the low-quality firm would be enhanced by the social stigma associated with the brown nature of the good. Interestingly, this model can be reconciled with a setting where the trade-off is absent: for this to be evident, it suffices to focus on the range of parameters such that one dimension of quality, either the hedonic or the environmental one, unambiguously prevails over the other. As an example, consider the case in which the environmental quality dominates the hedonic quality. Then, the analysis can immediately be developed along the lines of traditional literature on vertical differentiation with the environmental quality being the driver of consumption. Thus, in a framework without trade-off, at equilibrium either the market would be served by two producers with the price of the green variant being higher than the brown one, or only the green firm would be active.[4]

6 DISCUSSION OF THE LITERATURE

Clearly the above survey does not give credit to many scientific contributions in the field of product differentiation and it would be impossible to quote all the literature related to this field. Accordingly, we have chosen to cite only some significant contributions. Many of them are related to *international trade*. Thus, the papers by Helpman (1981) and Krugman (1980) cover basic trade theory. One should also cite in the same field the papers by Gabszewicz et al. (1981) as well as Shaked and Sutton (1984). One could also add the papers by Lyons (1984), Flam and Helpman (1987), Hallak (2006), Herguera, Kujal, and Petrakis (2002), Hummels and Klenow (2005), Motta, Thisse, and Cabrales (1995) and Baltzer (2011), to cite just a few.

Another significant topic in product differentiation is the question of *quality standards*. This question has given rise to important contributions dealing with quality standards and competition (Crampes and Hollander, 1995, Boom, 1995, Scarpa, 1998, Garella, 2006, Valletti, 2000), quality standards and collusion (Ecchia and Lambertini, 1997, Lambertini, Poddar and Sasaki, 2002), quality standards and trade (Petropoulou, 2013, Baltzer, 2011), and quality standards and environment policy (Motta and Thisse, 1993).

Also, numerous contributions are devoted to *innovation* in the context of product differentiation. Dutta, Lach, and Rustichini (1995) and Lehmann-Grube (1997) study the relationship between technology adoption and vertical product differentiation. Gabszewicz and Tarola (2007) use a model of vertical differentiation to examine the question of firms' ownership in

4 See on this Ben Elhadj and Tarola (2015) and Ben Elhadj, Gabszewicz, and Tarola (2015).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

the context of product innovation. Hoppe and Lehmann-Grube (2001) examine the second-mover advantages in the context of dynamic quality competition, while Boone (2001) considers the incentive to innovate as a consequence of the intensity of competition. Finally, under the impetus of Lambertini, several contributions are devoted to the approach to product differentiation via dynamic games (Cellini and Lambertini, 2007, 2011, Lambertini, 2012, and Colombo and Lambertini, 2003).

7 CONCLUSION

In this survey we have concentrated on describing the basics of product differentiation theory and its place in the theories of imperfect competition. It is clear that the approaches proposed here are somewhat partial and suffer from the absence of robustness in their conclusions. What is really lacking to develop a satisfactory theory of product differentiation is a notion of *proximity* among different variants of a commodity, allowing us to define on a solid basis the idea that two variants are "close" or, on the other hand, "far" from each other. The notion of price substitutability is clearly insufficient to fill this gap, since it depends on prices of the *existing* variants. Thus, this notion could explain how the existing variants of a good mutually depend on their price variations, but cannot serve to explain *why these variants have been chosen by the firms.* It is suggested in this survey that the approach via characteristics can be a satisfactory way to introduce the notion of proximity among variants. Nevertheless, a better-performing method should be found to measure the distance between different variants of a good, allowing in particular the consideration of several variants, rather than a single variation as in this study. Until such a measure is developed, economists will have to rely on partial approaches to tackle the difficult problem raised by firms' product choice.

REFERENCES

Akerlof G. (1997). Social Distance and Social Decisions. *Econometrica,* 65, 1005–1027.

Andreoni J. (1990). Impure Altruism and Donations to Public Goods: A Theory of Warm-glow Giving, *The Economic Journal* 100, 464–477.

Andreoni J. and J. Miller (2002). Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism. *Econometrica*, 70, 737–753.

Baltzer K. (2011). Minimum Quality Standards and International Trade, *Review of International Economics* 19 (5), 936–949.

Ben Elhadj N. and O. Tarola (2015). Relative Quality-related (Dis)utility in Vertically Differentiated Oligopoly with an Environmental Externality, *Environment and Development Economics,* 20 (3), 354–379.

Ben Elhadj N., J.J. Gabszewicz and O. Tarola (2015). Social Awareness and Price Competition, *International Journal of Economic Theory*, 11 (1), 75–88.

Ben Elhadj N., A. Hili and R. Lahmandi-Ayed (2013). Essays on Consumption Behaviors in a Vertical Differentiation Model, in G. Leismuller and E.J. Schimpf (eds), *Economics of Competition*, New York: Nova Science Publishers.

Boom A. (1995). Asymmetric International Minimum Quality Standards and Vertical Differentiation, *Journal of Industrial Economics*, 43 (1), 101–119.

Boone J. (2001). Intensity of Competition and the Incentive to Innovate, *International Journal of Industrial Organization*, 19 (5), 705–726.

Cellini R. and L. Lambertini (2007). A Differential Oligopoly Game with Differentiated Goods and Sticky Prices, *European Journal of Operational Research*, 176, 1131–1144.

Cellini R. and L. Lambertini (2011). R&D Incentives under Bertrand Competition: A Differential Game, *Japanese Economic Review*, 62 (3), 387–400.

Coestier B. and S. Marette (2004). *Economie de la qualité*, Paris: La Découverte.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Colombo L. and L. Lambertini (2003). Dynamic Advertising Under Vertical Product Differentiation, *Journal of Optimization Theory and Applications*, 119 (2), 261–280.

Conrad K. (2005). Price Competition and Product Differentiation When Consumers Care for the Environment. *Environmental and Resource Economics*, 31 (1), 1–19.

Crampes C. and A. Hollander (1995). Duopoly and Quality Standards, *European Economic Review*, 39 (1), 71–82.

d'Aspremont C., J.J. Gabszewicz and J.-F. Thisse (1979). On Hotelling's "Stability in Competition", *Econometrica*, 47, 1145–1150.

De Palma A., V. Ginsburgh, Y.Y. Papageorgiou and J.F. Thisse (1985). The Principle of Minimum Differentiation Holds Under Sufficient Heterogeneity, *Econometrica*, 53, 767–781.

Downs A. (1956). *An Economic Theory of Democracy*, New York: Harper & Row.

Dutta P.K., S. Lach and A. Rustichini (1995). Better Late than Early: Vertical Differentiation and the Adoption of a New Technology, *Journal of Economics and Management Strategy*, 4 (4), 463–489.

Eaton C. and R. Lipsey (1980). Freedom and the Existence of Pure Profit, *Economic Journal*, 88 (35), 455–469.

Ecchia G. and L. Lambertini (1997). Minimum Quality Standards and Collusion, *The Journal of Industrial Economics*, 45 (1), 101–113.

Flam H. and E. Helpman (1987). Vertical Product Differentiation and North–South Trade, *American Economic Review*, 77 (5), 810–822.

Gabszewicz J.J. and O. Tarola (2012). Product Innovation and Firms' Ownership, *Economics of Innovation and New Technology*, 21 (4), 323–343.

Gabszewicz J.J. and J.F. Thisse (1979). Price Competition, Quality and Income Disparities, *Journal of Economic Theory*, 20 (3), 340–359.

Gabszewicz, J.J. and J.F. Thisse (1986). On the Nature of Competition with Differentiated Products, *Economic Journal*, 96, 160–172.

Gabszewicz J.J. and X.Y. Wauthy (2012). Nesting Horizontal and Vertical Differentiation, *Regional Science and Urban Economics*, 42 (6), 998–1002.

Gabszewicz J.J., A. Shaked, J. Sutton and J. Thisse (1981). International Trade in Differentiated Products, *International Economic Review*, 22 (3), 527–534.

Garella P.G. (2006). "Innocuous" Minimum Quality Standards, *Economics Letters*, 92 (3), 368–374.

Gorman W.M. (1956). The Demand for Related Goods, *Journal Paper J-3129*, Iowa Agricultural Station.

Gorman W.M. (1980). A Possible Procedure for Analysing Quality Differentials in the Egg Market, *The Review of Economic Studies*, 47 (5), 843–856.

Gupta S. and G.T. Ogden (2009). To Buy or Not to Buy? A Social Dilemma Perspective on Green Buying. *Journal of Consumer Marketing,* 26 (6), 376–391.

Hallak J.C. (2006). Product Quality and the Direction of Trade, *Journal of International Economics*, 68 (1), 238–265.

Helpman E. (1981). International Trade in the Presence of Product Differentiation, Economies of Scale and Monopolistic Competition: A Chamberlin-Heckscher-Ohlin Approach, *Journal of International Economics*, 11, 305–340.

Herguera I., P. Kujal and E. Petrakis (2002). Tariffs, Quality Reversals, and Exit in Vertically Differentiated Industries, *Journal of International Economics*, 58 (2), 467–492.

Hoppe H.C. and U. Lehmann-Grube (2001). Second-mover Advantages in Dynamic Quality Competition, *Journal of Economics and Management Strategy*, 10 (3), 419–433.

Hotelling H. (1929). Stability in Competition, *The Economic Journal*, 39 (153), 41–57.

Hummels D. and P.J. Klenow (2005). The Variety and Quality of a Nation's Exports, *American Economic Review*, 95 (3), 704–722.

Jehiel P. (1992). Product Differentiation and Price Collusion, *International Journal of Industrial Organization*, 10 (4), 633–641.

Krugman P. (1980). Scale Economies, Product Differentiation, and the Pattern of Trade, *American Economic Review*, 70 (5), 950–959.

Lambertini L. (2012). Dynamic Hotelling Duopoly with Linear Transportation Costs, *Optimal Control Applications and Methods*, 33 (1), 114–126.

Lambertini L., S. Poddar and D. Sasaki (2002). Research Joint Ventures, Product Differentiation, and Price Collusion, *International Journal of Industrial Organization*, 20 (6), 829–854.

Lancaster K.J. (1966). A New Approach to Consumer Theory, *Journal of Political Economy*, 74 (2), 132–157.

Lehmann-Grube U. (1997). Strategic Choice of Quality When Quality is Costly: The Persistence of the High-quality Advantage, *The RAND Journal of Economics* 28 (2), 372–384.

Lin Y.-J. and K.I. Tu (2013). Vertical Outsourcing and Location Choice, *Procedia Economics and Finance*, 5, 494–501.

Lyons B.R. (1984). The Pattern of International Trade in Differentiated Products: An Incentive for the Existence of Multinational Firms, in H. Kierzkowski (ed.), *Monopolistic Competition and International Trade*, Oxford: Clarendon Press.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Mantovani A., O. Tarola and C. Vergari (2015). Hedonic Quality and Social Norms: A Hybrid Model of Product Differentiation, *WP DSE 1029*, Bologna.

Moraga-Gonzalez J.L. and N. Padro-Fumero (2002). Environmental Policy in a Green Market, *Environmental & Resource Economics*, 22, 419–447.

Motta M. and J.F. Thisse (1993), Minimum Quality Standards as an Environmental Policy: Domestic and International Effects, mimeo, Foundation Mattei, Milan.

Motta M., J.F. Thisse and A. Cabrales (1997). On the Persistence of Leadership or Leapfrogging in International Trade, *International Economic Review*, 38 (4), 809–824.

Mussa M. and S. Rosen (1978). Monopoly and Product Quality. *Journal of Economic Theory,* 18, 301–317.

Neven D. and J. Thisse. (1990). On Quality and Variety Competition, in J. Gabszewicz, J.F. Richard, and L. Wolsey (eds), *Economic Decision Making: Games, Econometrics and Optimization; Contributions in Honour of Jacques Drèze,* Amsterdam: North-Holland. 175–199.

Petropoulou D. (2013). Vertical Product Differentiation, Minimum Quality Standards, and International Trade, *Oxford Economic Papers* 65 (2), 372–393.

Prescott E. and M. Visscher (1977). Sequential Location Among Firms with Foresight, *The Bell Journal of Economics*, 8 (2), 378–393.

Salop S. (1979). Monopolistic Competition with Outside Goods, *The Bell Journal of Economics*, 10 (1), 141–156.

Scarpa C. (1998). Minimum Quality Standards with More than Two Firms, *International Journal of Industrial Organization* 16 (5), 665–676.

Shaked A. and J. Sutton (1983). Natural Oligopolies, *Econometrica*, 51 (5), 1469–1484.

Shaked A. and J. Sutton (1984). Natural Oligopolies and International Trade, in H. Kierzkowski (ed.), *Monopolistic Competition and International Trade*, Oxford: Clarendon Press.

Valetti T.M. (2000). Minimum Quality Standards Under Cournot Competition, *Journal of Regulatory Economics*, 18 (3), 235–245.

Vives X. (2000). *Oligopoly Pricing*, Cambridge, MA: MIT Press.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

7. Oligopolistic competition and welfare

Robert A. Ritz *

1 INTRODUCTION

1.1 Objectives

Market performance under imperfect competition has been a classic question for economists since the time of Adam Smith. It remains a central concern of the theoretical and empirical industrial organization (IO) literature and antitrust policymakers dealing with competition issues in practice.

The objective of this chapter is to survey recent developments in the IO theory literature that speak to oligopoly and welfare. The coverage here is explicitly selective, concentrating on areas where the literature has substantially progressed over the last five to ten years. Related issues have been covered extensively by several authors in the past. Valuable resources remain: the early survey chapter by Shapiro (1989) as well as the oligopoly-theory books by Tirole (1988) and Vives (1999), which also contain significant material on welfare.

The uniqueness of this chapter lies in the following. First, the focus is specifically on welfare; most other treatments deal with this only as a by-product. Second, it covers recent developments that have not yet found their way into textbook treatments – but hopefully will do so in the near future. Third, it discusses separate strands of the recent literature in a way that highlights their common themes.

1.2 Scope

The scope of this chapter is limited to relatively simple static oligopoly models under partial equilibrium analysis.[1] It concentrates on theory – albeit in a way that it is informed by the empirical literature and speaks to industrial applications. Extensions to more complex settings are dealt with by other chapters contained in this volume.

Market power lies solely with firms, while buyers are atomistic; there is no price discrimination. The focus is on markets with varying degrees of competitive conduct – rather than tacit collusion or price fixing. Firms are assumed to be risk-neutral profit maximizers and are equally well informed about the market. There are no other market failures (such as environmental externalities) and no explicit role for regulation (such as price caps) or other policy interventions.

The definition of "welfare" (W) is mostly taken to be social surplus, that is, the unweighted sum of aggregate consumer surplus (CS) and aggregate producer surplus (Π): $W = CS + \Pi$.

* I am grateful to Anette Boom, Simon Cowan, Federico Etro, Pär Holmberg, Nathan Miller, Michael Pollitt, Andrew Rhodes, John Vickers and Glen Weyl for helpful comments and suggestions. Any remaining errors are mine.

[1] This excludes any general-equilibrium effects, which, for example, could arise due to interactions between supplier market power and imperfections in input markets (such as the labour market).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

A consumer welfare standard is highlighted in some places given that recent antitrust policy in jurisdictions such as the USA and EU is said to be geared more heavily towards consumers.[2]

The results discussed cover a range of models with homogeneous products as well as different forms of horizontal product differentiation. Some of the homogeneous products results apply equally to settings with vertical differentiation in which there are (known) differences in product quality across firms. Many of the models are "aggregative games" in which a firm's competitive environment can be captured using a single summary statistic of rivals' actions.

These models have useful applications across a wide array of industries. In the energy sector, similar homogeneous product models are widely employed in the analysis of electricity, natural gas and crude oil markets – as well as energy-intensive industry such as cement and steel. The differentiated price models covered form the basis for competition policy in sectors with branded products.

1.3 Plan for the Chapter

Section 2 presents the recent literature on the rate of cost pass-through as an economic tool with which to understand the market performance and the division of surplus between buyers and sellers. Section 3 discusses recent papers that quantify market performance in various Cournot-style models using welfare losses, that is, the comparison between equilibrium welfare and first-best. Section 4 covers recent developments in the theory of oligopoly with endogenous entry of firms, with a focus on the quantification of welfare losses and the impact of firm heterogeneity. Section 5 provides concluding remarks and suggestions for future research.

2 COST PASS-THROUGH AND THE DIVISION OF SURPLUS

Consider the treatment of monopoly in a textbook on microeconomics or industrial organization. With linear demand and costs, the monopolist captures 50 per cent of the (potential) gains from trade, with 25 per cent as consumer surplus – and the remainder as deadweight loss. So there is a ratio of 1:2 between consumer surplus and producer surplus.

Elsewhere, the textbook may turn to the question of cost pass-through: how much of a unit tax is passed onto the market price? For a linear monopoly, the rate of pass-through ($\Delta P / \Delta MC$) equals 50 per cent. So there is a ratio of 1:2 between the price change and the cost increase.

What textbooks do *not* say is that this is no coincidence. The ratio of consumer to producer surplus, in equilibrium, is equal to the rate of cost pass-through in that market. Weyl and Fabinger (2013) develop this insight more broadly, including for various representations of oligopoly, and argue that pass-through is a versatile tool with which to think about market performance.

---

[2] See Farrell and Katz (2006) for a discussion of welfare standards in antitrust. Armstrong and Vickers (2010) study a model in which a consumer welfare standard can, for strategic reasons, be optimal even if the regulator cares about total welfare (because the standard affects the set of mergers that is proposed by firms).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Much earlier, Bulow and Pfleiderer (1983) noted how monopoly cost pass-through varies with the shape of the demand curve, i.e., its curvature. Kimmel (1992) exploits this link to frame the profit impact of a unit tax in a Cournot oligopoly in terms of pass-through. Anderson and Renault (2003) study the relationship between demand curvature and the division of surplus under Cournot competition but do not explicitly cover pass-through. Weyl and Fabinger (2013) tie together these various antecedents.

2.1 Monopoly Case

Consider a monopolist that produces a single good with marginal cost $c + t$, where $t \geq 0$ is a parameter. The monopolist faces inverse demand $p(Q)$; let $D(p)$ be the corresponding direct demand. At the optimum, marginal revenue equals marginal cost, $MR(Q) = c + t$.

What is the impact of a small increase in t? Let $\kappa \equiv dp/dt$ denote the rate of cost pass-through, which measures how price responds to a \$1 increase in marginal cost.[3] Denote consumer surplus $CS = \int_p^\infty D(x)dx$, and observe that $dCS/dt = -\kappa Q$, at the optimum. Similarly, by the envelope theorem, the profit impact $d\Pi/dt = -Q$, as the indirect impact of the tax is zero since the monopolist is optimizing. Hence the burden of an infinitesimal tax, starting at zero, is split according to

$$\left.\frac{dCS/dt}{d\Pi/dt}\right|_{t=0} = \kappa(0),$$

where $\kappa(0)$ is pass-through at the price corresponding to initial zero tax rate.

Consider now a discrete increase in the tax from t_0 to $t_1 > t_0$. Write $Q(t)$ for the optimal quantity as a function of the tax. The changes in consumer surplus and monopoly profits satisfy $\Delta CS_{t_0}^{t_1} = -\int_{t_0}^{t_1} \kappa(t)Q(t)dt$ and $\Delta\Pi_{t_0}^{t_1} = -\int_{t_0}^{t_1} Q(t)dt$. Define the quantity-weighted pass-through over the interval $[t_0, t_1]$ as $\overline{\kappa}_{t_0}^{t_1} \equiv [\int_{t_0}^{t_1} \kappa(t)Q(t)dt]/\int_{t_0}^{t_1} Q(t)dt$. Define $\overline{t}$ as the hypothetical tax rate at which the market is eliminated, that is, $Q(\overline{t}) = 0$, and call the average quantity-weighted pass-through rate $\overline{\kappa} \equiv \overline{\kappa}_0^{\overline{t}}$.[4] Hence the surplus generated from the market's "birth" (at $\overline{t}$) to the equilibrium status quo (at $t = 0$) satisfies

$$\frac{\Delta CS_0^{\overline{t}}}{\Delta\Pi_0^{\overline{t}}} = \frac{\int_0^{\overline{t}} \kappa(t)Q(t)dt}{\int_0^{\overline{t}} Q(t)dt} = \overline{\kappa} = \frac{CS}{\Pi}.$$

Consumer surplus is generated by the market at a rate of monopoly profits times the pass-through rate, weighted over the inframarginal market quantities traded over the interval $[0, \overline{t}\,]$. This takes into account that the pass-through rate may not be a constant.

Intuitively, high pass-through means that price closely tracks marginal cost, so that (i) the monopolist's degree of market power is "low", and, conversely, (ii) realized social surplus is "high" and largely goes to consumers. With low pass-through, price follows more closely consumers' willingness-to-pay (WTP) so the monopolist captures the bulk of the gains from trade.

[3] Another formulation, more frequently used in the international trade literature, instead concerns the pass-through elasticity $(dp/p)/(dt/t) \leq \kappa$, which also incorporates the profit margin.

[4] Some demand curves have $\overline{t} = \infty$, though a finite choke price can be assumed.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Bulow and Pfleiderer (1983) showed that monopoly pass-through satisfies:

$$\kappa(t) = \frac{1}{[2-\xi(t)]} = \frac{\text{slope of inverse demand } p(Q)}{\text{slope of marginal revenue } MR(Q)}\bigg|_{Q=Q(t)},$$

where $\xi(t) \equiv -\left[p''(Q)/Qp'(Q)\right]_{Q=Q(t)}$ is the elasticity of the slope of inverse demand, which is a measure of demand curvature. The common theory assumption (Bagnoli and Bergstrom, 2005) that direct demand $D(p)$ is log-concave (i.e., $\log D(p)$ is concave in p), corresponds to $\xi \leq 1$, and hence to pass-through (weakly) less than 100 per cent. Loosely put, the monopolist then captures a greater share of the gains from trade than consumers. For very concave demand, $\xi \ll 0$, the "triangle" left as consumer surplus is very small; correspondingly the ratio CS/Π and pass-through κ are both small – as is the remaining deadweight loss.

For many familiar demand curves, the ratio $p'(Q)/MR'(Q)$ is constant, so pass-through is a constant with $\kappa(t) = \kappa$ for all $t \in [0, \bar{t}]$ – and so the "local" properties of demand are also "global". With linear demand, marginal revenue is everywhere twice as steep as demand, so pass-through $\kappa = \frac{1}{2}$. Other examples are constant elasticity demand, for which $\xi = 1 + 1/\eta > 1$ (violating log-concavity) where $\eta \equiv -p(Q)/Qp'(Q) > 0$ is the price elasticity, and exponential demand $D(p) = \exp((\alpha - p)/\beta)$, for which $\xi = 1$ as it is log-linear. In such cases, the marginal impact of a tax is equal to its average impact, $(dCS/dt)/(d\Pi/dt) = \Delta CS_{t_0}^{t_1}/\Delta\Pi_{t_0}^{t_1} = CS/\Pi = \kappa$.

The literature has found different ways of representing "constant" higher-order properties of demand. First, using the concept of ρ-concavity: demand $D(p)$ is ρ-concave if and only if demand curvature $\xi(Q) \leq (1-\rho)$ (Anderson and Renault, 2003). A ρ-linear demand curve thus has constant curvature $\xi = 1 - \rho$, and constant pass-through over its entire domain. Second, the demand curve $D(p)$ can be interpreted as arising from the values v of a distribution $F(v)$ of consumers with unit demand, so $1 - F(p)$ is the quantity sold at price p. The inverse hazard rate is $h(v) \equiv [1 - F(v)]/f(v)$ where $f(v)$ is the density. The monopolist's first-order condition $(p - c) = h(p)$, so pass-through is constant whenever the inverse hazard takes the linear form $h(v) = \lambda_0 + \lambda_1 v$. Third, Rostek, Weretka and Pycia (2009) show that a distribution has a linear inverse hazard rate if and only if it belongs to the generalized Pareto distribution, $F(v) = 1 - \left[1 + \frac{\omega}{\sigma}(v - \mu)\right]^{-1/\omega}$, where (μ, σ, ω) respectively describe its location, scale and shape (with $\lambda_0 = (\sigma - \omega\mu)$ and $\lambda_1 = \omega$).

2.2 Oligopoly Models

The preceding insights generalize to certain n-firm oligopoly models. Consider a general reduced-form model of competition in which firm i's profits $\pi_i = (p_i - c)q_i$ and the Lerner index (price–cost margin) with symmetric firms is determined as:

$$\varepsilon_D \frac{(p - c - t)}{p} = \theta,$$

where θ is a "conduct parameter" that measures the intensity of competition, and $\varepsilon_D \equiv -p(Q)/Qp'(Q)$ is the *market-level* price elasticity of demand.[5] The previous monopoly

[5] A large empirical literature reviewed by Reiss and Wolak (2007) has developed structural econometric techniques for estimating the intensity of competition.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

analysis corresponds to joint profit maximization with $\theta = 1$. This setup nests various widely used models of symmetric oligopoly, including the two following models:

Homogeneous product oligopoly Consider a Cournot model augmented with "conjectural variations": when firm i chooses its output it conjectures that each other firm j will adjust its quantity by $dq_j = [R/(n-1)]dq_i$. So the aggregate responses by all its rivals is given by $d(\sum_{j\neq i} q_j)/dq_i = R$. Cournot-Nash competition corresponds to $R = 0$ while Bertrand competition in effect has $R = -1$ (so the price stays fixed). Conjectural variations can be seen as a reduced-form way of incorporating (unmodelled) dynamic features of the game that firms play (Cabral, 1995).

The first-order condition for firm i has $MR_i = p(Q) + q_i p'(Q)(1+R) = c + t$, where $Q \equiv \sum_{i=1}^{n} q_i$ is industry output. This can be rearranged to give the symmetric equilibrium (with $q_i = Q/n$):

$$\varepsilon_D \frac{(p-c-t)}{p} = \frac{(1+R)}{n} = \theta.$$

Thus a constant conjectural variation R corresponds to a constant conduct parameter θ.

Differentiated products price competition Consider a model of price-setting competition with symmetrically differentiated products. Firm i's demand $q_i(p_i, \mathbf{p}_{-i})$ depends on its own price and those of its $n-1$ rivals. In symmetric equilibrium (with $q_i = q = Q/n$), the corresponding price can be written as $p(q)$, which captures how each price changes in response to a simultaneous change in *all* firms' outputs.

The first-order condition, at symmetric equilibrium, for firm i is given by the inverse-elasticity rule, $(p-c-t)/p = -(q/p)/(\partial q_i/\partial p_i)$. The elasticity of market demand is $\varepsilon_D = -(p/q)\sum_{j=1}^{n}(\partial q_i/\partial p_j)$, and so:

$$\varepsilon_D \frac{(p-c-t)}{p} = \frac{\sum_{j=1}^{n}(\partial q_i/\partial p_j)}{\partial q_i/\partial p_i} = 1 + \frac{\sum_{j\neq i}(\partial q_j/\partial p_i)}{\partial q_i/\partial p_i} = 1 - A = \theta,$$

where A is the "diversion ratio" from firm i to the rest of the industry as it raises its price (Shapiro, 1996).[6] With a linear demand system, for example, A is constant – and hence the conduct parameter is also constant.

As in the monopoly case, the envelope theorem together with the symmetric demand structure imply that the marginal impact of an increase in the tax rate on consumer surplus is given by $dCS/dt = -\kappa(t)Q(t)$. Weyl and Fabinger (2013) show that the marginal impact on producers is given by $d\Pi/dt = -[1-\kappa(t)(1-\theta(t))]Q(t)$, where industry profits $\Pi \equiv \sum_{i=1}^{n}\pi_i$. So the burden of an infinitesimal tax, starting at zero, is split according to:

$$\left.\frac{dCS/dt}{d\Pi/dt}\right|_{t=0} = \frac{\kappa(0)}{[1-\kappa(0)(1-\theta(0))]}.$$

[6] With the symmetric demand structure, $\frac{\sum_{j\neq i}(\partial q_i/\partial p_j)}{\partial q_i/\partial p_i} = \frac{\sum_{j\neq i}(\partial q_j/\partial p_i)}{\partial q_i/\partial p_i}$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

This is a clean generalization of the monopoly case, with some intuitive properties. For given pass-through $\kappa(0)$, less competitive conduct (higher θ) skews the division of surplus from consumers to producers. For given conduct θ, higher pass-through favours consumers.

The pass-through rate is here given endogenously by:

$$\kappa(t) = \left. \frac{1}{1 + \theta(t)\left[\varepsilon_{MCS} + \varepsilon_{\theta}\right]} \right|_{Q=Q(t)},$$

where $\varepsilon_{\theta} \equiv d \log \theta(q)/d \log q$ is the elasticity of the conduct parameters to changes in output, and $\varepsilon_{MCS} \equiv d \log CS'(Q)/d \log Q$ measures how responsive the *marginal* consumer surplus $CS'(Q) = -p'(Q)Q = [p(Q) - MR(Q)]$ (Bulow and Klemperer, 2012) is to changes in aggregate output.[7] The pass-through rate, in general, must capture how both of these metrics may vary as the tax affects equilibrium quantities. For example, if the tax reduces per-firm output ($dq/dt < 0$) and this makes the industry more competitive ($d\theta/dq > 0$), then this will tend to enhance pass-through. Note also that pass-through depends indirectly on the number of firms, since this will generally enter into $\theta(t)$.[8]

As in the monopoly case, it is possible to go from this local impact to the global division of surplus by appropriately weighting how pass-through rates change along the demand curve.

With "constant conduct" and "constant curvature", the global division of surplus again follows immediately from its local properties. As noted above, many oligopoly models feature $\theta(t) = \theta$ so that $\varepsilon_{\theta} = 0$. It is also instructive to write out $\varepsilon_{MCS} = [1 - \xi(t)]$ in terms of demand curvature. (Log-concave demand $\xi < 1$ corresponds to $\varepsilon_{MCS} > 0 \Leftrightarrow CS''(Q) > 0$.) With these modifications, pass-through becomes $\kappa = 1/\left[1 + \theta(1 - \xi)\right]_{Q=Q(t)}$, which nests the well-known Cournot-Nash oligopoly result (Kimmel, 1992) when $\theta = 1/n$.

2.3 Discussion

The insight that the division of surplus is pinned down by the rate of pass-through has a number of appealing features. First, it allows pass-through to be seen as a "sufficient statistic" for welfare analysis. Second, pass-through estimates already exist in the literature for many markets – based on studies of taxation, exchange rates, and other cost shifts. Third, it makes it easier to form intuitions about market performance since pass-through rates are often easier to think about than higher-order properties of demand.

Information on pass-through can also be used in the reverse direction. For price competition with differentiated products, Miller, Remer and Sheu (2013) instead emphasize how, *assuming* second-order demand properties (i.e., demand curvatures), the matrix of pass-through rates across products can be used to estimate a matrix of "first-order" cross-price elasticities. The attraction of this is that it sidesteps the problem of full-scale estimation of the demand system – which can be time-consuming or even infeasible.

While it is relatively easy to obtain empirical estimates of pass-through, it is more difficult to ascertain how pass-through itself varies along a demand curve. Yet, strictly speaking, the

[7] Note that $d \log CS'(Q)/d \log Q = d \log CS'(Q)/d \log q$ given the symmetric setup.

[8] For Bertrand competition (with $\theta \equiv 0$), note that $CS/\Pi = \kappa/(1-\kappa)$ but also $\kappa = 1$, so that $CS/\Pi \to 0$ (since firms make zero profits).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

theory requires the *quantity-weighted* pass-through $\overline{\kappa} \equiv \overline{\kappa}_0^{\bar{t}}$. MacKay et al. (2014) show how reduced-form regressions of price on cost may not yield reliable estimates of the rate of cost pass-through. Loosely put, such a regression can only yield consistent estimates in situations where the underlying environment is such that cost pass-through is constant over the range of prices in the data. Empirical implementation of the theory may have to resort to assuming $\kappa(t) = \kappa$ for all (or large parts of) $t \in [0, \bar{t}]$.

The above results are based on strong symmetry assumptions such as identical marginal costs and symmetrically differentiated products. These greatly simplify the welfare analysis but are likely to be violated in *any* oligopoly. Weyl and Fabinger (2013) also develop results from a general model that allows certain types of asymmetries. Other factors, such as the details of market structure, then come into play. Again, it is possible to adjust the definition of pass-through to incorporate these but this means that estimating this "adjusted" pass-through rate becomes increasingly difficult – and begins to merge into estimation of a full-scale market model. The power of pass-through is strongest for monopoly.

Another assumption is that the number of the firms in the market is fixed, and hence invariant to changes in costs. Ritz (2014b) shows that, with log-convex demand, a higher unit tax can induce additional entry into a market, and thus ultimately lead to a lower market price. Negative pass-through, also known as "Edgeworth's paradox of taxation", is ruled out in the models covered here. Conversely, a low pass-through rate can induce exit of weaker firms, which in turn causes price to jump back up.[9]

3 QUANTIFYING WELFARE LOSSES IN COURNOT-STYLE MODELS

Consider a textbook Cournot oligopoly with symmetric firms. How large are welfare losses due to market power? With three firms, they equal 6.66 per cent; in other words, a highly concentrated Cournot triopoly delivers over 93 per cent of the maximum possible welfare.[10] For a duopoly, the loss is 11 per cent – certainly not trivial, but not large either.

A recent literature quantifies market performance directly in terms of realized welfare (Corchón, 2008; Ritz, 2014a). It shows that welfare losses in familiar oligopoly models are often perhaps surprisingly small, and also shows what market factors can generate more substantial losses.

The approach is based on calculating equilibrium welfare losses relative to the first-best benchmark. It turns out that this ratio can naturally be determined in terms of observable metrics, notably firms' market shares. In this way, this literature is potentially useful also for policy purposes as a simple initial screening tool for market performance.[11]

---

9 Further afield, in the context of the commercial banking industry, Ritz and Walther (2015) show how risk aversion and informational frictions tend to dampen the pass-through of changes in interest rates across loan and deposit markets.

10 For a duopoly in which one firm is a Stackelberg leader, the welfare loss also equals 6.66 per cent – so the social value of leadership is equal to one additional entrant.

11 An older empirical literature going back to Harberger (1954) estimates welfare losses normalized relative to sales revenue. A disadvantage is that magnitudes are hard to intepret; for example, the ratio of equilbrium welfare to revenue can vary widely for reasons that have nothing to do with market power.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

3.1 Cournot-Nash Oligopoly

Consider a Cournot-Nash oligopoly with $n \geq 2$ active firms. Firm i has marginal cost c_i and chooses its output q_i to maximize its profits $\pi_i = (p - c_i)q_i$, where the price $p(Q)$ with industry output $Q \equiv \sum_{i=1}^n q_i$. Without loss of generality, firms are ordered such that $c_1 \leq c_2 \leq \ldots \leq c_n$. Inverse demand $p(Q) = \alpha - \beta Q^{1-\xi}$ is $(1-\xi)$-linear with constant curvature ξ, where $\xi < 2$ gives downward-sloping *industry* marginal revenue. This also ensures the uniqueness and stability of the Cournot equilibrium as well as a well-behaved consumer surplus function.

The first-best outcome, which maximizes social welfare $W \equiv CS + \Pi$, where $\Pi \equiv \sum_{i=1}^n \pi_i$, has price equal to the lowest marginal cost $p^{fb} = c_1$ with output $Q^{fb} = p^{-1}(c_1) = \left[(\alpha - c_1)/\beta\right]^{1/(1-\xi)}$. Denote the corresponding welfare level as W^{fb}.

The first-order condition for firm i is $MR_i = c_i$, and the sum of first-order conditions $\sum_{i=1}^n MR_i \equiv \left[np(Q) + Qp'(Q)\right] = \sum_{i=1}^n c_i$ pins down the equilibrium industry output Q^*. Hence the equilibrium price is given by:

$$\left[np^* - (1-\xi)(\alpha - p^*)\right] = \sum\nolimits_{i=1}^n c_i \Longrightarrow p^* = \frac{(1-\xi)\alpha + n\overline{c}}{(n+1-\xi)}.$$

This equilibrium pricing function $p^*(\overline{c})$ is affine in the unweighted average unit cost $\overline{c} \equiv \frac{1}{n}\sum_{i=1}^n$, so the pass-through of a cost change that affects all firms equally $\kappa \equiv dp^*/d\overline{c} = n/(n+1-\xi)$ is constant (i.e., $d^2p^*/d\overline{c}^2 = 0$).

Denote equilibrium welfare and consumer surplus under Cournot competition as W^* and CS^*, and define welfare losses relative to first-best as:

$$L \equiv \left(1 - \frac{W^*}{W^{fb}}\right),$$

which is a unit-free measure of welfare that lies on the unit interval, $L \in [0, 1]$.

3.1.1 Symmetric firms

To build intuition, it is useful to begin with the benchmark case in which firms have identical marginal costs, $c_i = \overline{c}$ for all i; Anderson and Renault (2003) showed that:

$$L(n, \xi) = 1 - \frac{n^{1/(1-\xi)}(n+2-\xi)}{(n+1-\xi)^{(2-\xi)/(1-\xi)}}.$$

Equilibrium welfare losses depend only on the number of (symmetric) firms and the curvature of demand. As expected, they decline with the number of firms and tend to zero at the limit as the competitors grows large. This reflects the classic result on convergence to perfect competition in large markets.

Welfare losses also tend to zero if the curvature of demand is extreme, either as $\xi \to 2$ or as $\xi \to -\infty$. The case with $\xi \to 2$ corresponds to very convex demand in which the total revenue to firms (and hence the total expenditure by consumers) become constant – and thus invariant to the number of firms competing; since production costs are symmetric, there is no other source of welfare losses. The case with $\xi \to -\infty$ corresponds to demand that becomes rectangular (infinitely concave) so all consumers have identical WTP of α for the

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

good; then firms extract all the gains from trade with a uniform price $p^* = \alpha$ while serving all consumers efficiently. Consistent with the previous monopoly discussion, this is the limit of zero pass-through, $\kappa \to 0$.

More generally, Corchón (2008) shows that welfare losses with symmetric firms tend to be quite "small". For example, with linear demand ($\xi = 0$) the above simplifies to $L(n) = 1/(n+1)^2$. So welfare losses are of order $1/n^2$ (as the price and output inefficiencies are both of order $1/n$) and in a quantitative sense decline quickly as the number of competitors rises, e.g., $L(n) \leq 4$ per cent if $n \geq 4$. For non-linear demands, Corchón (2008) derives the *maximal* welfare loss for a given number of firms, that is, $\widehat{L}(n) \equiv \max_\xi L(n, \xi)$. As long as there are at least four firms in the market, overall welfare losses are never greater than around 5.8 per cent. In fact, the textbook case with linear demand generally yields fairly *high* welfare losses.

3.1.2 Asymmetric firms

The symmetric case shows that welfare losses due to market power do not tend to be "large"– say well above 5 per cent – in Cournot-Nash models, except in some duopoly cases. However, the symmetry assumption switches off any role for welfare losses due to productive inefficiency. Indeed, it is well known that Cournot equilibria are not cost efficient since the lowest-cost firm does produce all output; high-cost firms serve too much of the market (Lahiri and Ono, 1988; Farrell and Shapiro, 1990; Aiginger and Pfaffermayr, 1997).[12]

More realistic results revert back to the case where firms' marginal costs may be asymmetric. The challenge is that costs are typically difficult to observe (or even reliably estimate), while there is an advantage in having a welfare measure that depends on observables as far as possible. The trick to resolve this is to use the first-order conditions to "substitute out" costs for market shares that are readily available for many markets.

In particular, let firm i's equilibrium market share $s_i^* \equiv \left(q_i^*/Q^*\right)$, and recall the first-order condition $MR_i = p(Q) + q_i p'(Q) = p - (1-\xi)\beta Q^{2-\xi} s_i = c_i$. Some rearranging shows that its equilibrium market share satisfies:

$$(1-\xi)(\alpha - p^*)s_i^* = (p^* - c_i),$$

which provides a direct mapping between (observable) market share and marginal cost, for a given p^* as determined above. Note that firm 1's market share s_1^* is the highest since it has the lowest marginal cost, and $s_1^* \geq s_2^* \geq \ldots \geq s_n^*$.

Based on this, Corchón (2008) shows that welfare losses with asymmetric firms are given by:

$$L(s_1^*, H^*, \xi) = 1 - \frac{\left[1 + (2-\xi)H^*\right]}{\left[1 + (1-\xi)s_1^*\right]^{(2-\xi)/(1-\xi)}},$$

where $H^* \equiv \left(\sum_{i=1}^n s_i^2\right)_{\{s_i^*\}_{i=1}^n}$ is the Herfindahl index of concentration, evaluated at the equilibrium market shares.[13] The expression for welfare losses remains simple: they now

[12] More generally, marginal costs are not equalized across firms (as would occur in a cost-minimizing allocation of any given industry production level).

[13] This expression simplifies to the symmetric case where $H^* = s_1^* = n^{-1}$ for all i.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

depend on s_1^* and the Herfindahl index H^*, both of which previously boiled down to the number of firms in the symmetric case.[14] Intuitively, market performance under Cournot is described by the Herfindahl index, while the largest market share captures how close this performance is to first-best – for which it should equal 100 per cent.

Welfare losses increase with the market share of the largest firm s_1^* (holding fixed the value of the Herfindahl index). Intuitively, the largest firm must have an above-average market share; further increasing its size relative to the market pushes the equilibrium closer to monopoly.

Welfare losses decline in the Herfindahl index (holding fixed s_1^*). While perhaps initially counterintuitive, the reason for the results is that a higher industry concentration shifts market share toward the more efficient firms (which have lower costs). This mitigates the productive inefficiency of the Cournot equilibrium. The more general point is that the Herfindahl index is not a reliable guide to market performance.

Corchón (2008) shows that welfare losses in asymmetric Cournot models can be *very* large. Specifically, it is possible to find combinations of demand conditions (ξ) and market structure $(s_1^*, s_2^*, \ldots, s_n^*)$ that yield welfare losses that are arbitrarily close to unity, $L(s_1^*, H^*, \xi) = 1-\epsilon$ for a small constant $\epsilon \to 0$. At the same time, the Herfindahl index may be arbitrarily low. The worst case for welfare is when the non-largest firms are symmetric, $s_2^* = s_3^* = \ldots = s_n^*$; then $\lim_{\xi\to-\infty}\left[\lim_{n\to\infty} L(s_1^*, n, \xi)\right] = 1 - s_1^*$, and clearly $H^* \approx 0$ while $L \approx 1$ for s_1^* small.

Welfare losses can be substantially higher than in symmetric cases, even with non-extreme assumptions about demand curvature and realistic market structures. As a numerical example, let firms' market shares $s_1^* = 40$ per cent, $s_2^* = 30$ per cent, $s_3^* = 20$ per cent, and $s_4^* = 20$ per cent which implies a Herfindahl index $H^* = 0.3$. Assuming linear demand ($\xi = 0$), it follows that welfare losses $L(s_1^*, H^*, \xi) \approx 18$ per cent. This is approximately *three* times as high as the *maximal* loss with four symmetric firms.

Surprisingly, it is possible for market performance under Cournot to be *worse* than for a monopoly. Corchón (2008) shows that with log-convex demand ($\xi > 1$), monopoly indeed generates the highest welfare loss. However, with log-concavity ($\xi < 1$), the socially worst outcome involves a "high" market share (at least 50 per cent) for one firm combined with a "tail" of very small firms. The intuition is that the small firms add little to competition but substantially reduce productive efficiency.

Finally, with asymmetric firms, market performance is no longer obviously related to cost pass-through. Pass-through $\kappa(n, \xi)$ reflects the number of competitors and demand conditions, while welfare losses $L(s_1^*, H^*, \xi)$ also depend on the details of the distribution of firms' market shares. Market performance can vary widely even for a fixed underlying rate of pass-through.

3.2 Endogenous Competitive Conduct in Two-stage Games

A significant body of empirical evidence shows that many industrial markets have a competitive intensity that is tougher than Cournot-Nash but falls short of perfect competition (Bresnahan, 1989). One way to model this, as in Section 2, is by adding an *exogenous* conduct parameter. Similarly, a widely used class of two-stage strategic games comes with an conduct

[14] Again, the demand parameters (α, β) do not play any role: the influence of α is subsumed in firms' market shares and β is merely a scale factor that does not affect *relative* welfare losses. (All else equal, doubling β halves both W^* and W^{fb} so their ratio is unchanged.)

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

parameter that is endogenously determined by the interaction of the two stages. It turns out that welfare losses in such models can be *much* lower than in the standard Cournot setup.

Consider the two-stage game introduced by Vickers (1985) and Fershtman and Judd (1987). Each firm delegates decision-making in the product market to a manager. Manager i receives an incentive contract that induces maximization of an objective function $\Omega_i = (1-\varphi_i)\pi_i + \varphi_i R_i$, where $R_i \equiv pq_i$ is the firm's sales revenue. In the first stage, each firm's shareholders choose the incentive weight φ_i to maximize their firm's profits π_i. In the second stage, each firm's manager chooses an output level q_i to maximize the objective Ω_i.

This setup reflects extensive evidence that managers across a wide range of industries appear to place significant emphasis on measures of their firm's size (Ritz, 2008, 2014a). This is particularly evident in competition for rankings in "league tables" that are based on firms' sales or market share, not profits – and play a prominent role, for example, in commercial and investment banking as well as in car and aircraft manufacturing.[15]

Firms can use their Stage 1 choice of the incentive contract as a commitment device to gain strategic advantage in the product market.[16] Higher values of φ_i constitute aggressive output-increasing behaviour since they correspond to placing less weight on costs. Aggressive behaviour is optimal when firms are competing in strategic substitutes since it induces a soft response from rivals. From the firms' viewpoint, this leads to a prisoner's dilemma: each firm individually has an incentive to engage in aggressive behaviour but this ends up making them collectively worse off.

Remark The exposition here focuses on a widely used two-stage model of delegation. Yet the same welfare conclusions apply to a range of other two-stage models that are strategically equivalent. This includes the seminal model of Allaz and Vila (1993) in which firms engage in forward trading of their production, hiring "overconfident" managers who overestimate the state of market demand, and models of strategic trade policy in which countries use output subsidies to commit their firms to aggressive behaviour. (See Ritz, 2008, 2014a for further discussion.)

The game is solved backward for the subgame-perfect Nash equilibrium. Manager i's first-order condition in Stage 2 is given by:

$$\frac{\partial \Omega_i}{\partial q_i} = (1-\varphi_i)\frac{\partial \pi_i}{\partial q_i} + \varphi_i \frac{\partial R_i}{\partial q_i} = \left[p(Q) + p'(Q)q_i - (1-\varphi_i)c_i\right] = 0.$$

This implicitly defines manager i's best response in the product market. Let $q_i^*(\varphi_1, \varphi_2, \ldots, \varphi_n)$ denote the Nash equilibrium output choice, as a function of all firms' incentive contracts. Given this, in Stage 1, each firm's shareholders choose their manager's incentive weight according to:

$$\frac{d\pi_i}{d\theta_i} = \left[p(Q^*) + p'(Q^*)q_i^*(1+\upsilon_{-i}) - c_i\right]\frac{dq_i^*}{d\varphi_i} = 0,$$

[15] There is also a large body of evidence that shows that executive compensation in manufacturing, service and financial industries often rewards measures of firm size in addition to profits.

[16] It is assumed that such commitment is credible; a sufficient condition for this is that managers' contracts are observable and cannot be renegotiated.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

where $\upsilon_{-i} \equiv (dQ_{-i}/dq_i)_{\{q_i^*\}_{i=1}^n} < 0$ is the aggregate response of rivals' $Q_{-i} \equiv \sum_{j \neq i} q_j$ and $dq_i^*/d\varphi_i > 0$. Combining the two first conditions, the contract places positive weight on sales revenue $\varphi_i^* > 0$ if and only if $\upsilon_{-i}^* < 0$. This corresponds to a conduct parameter for firm i's product-market behaviour; the only difference is that υ_{-i}^* is here determined endogenously in Stage 1.[17]

Ritz (2014a) shows that, with linear demand, $\upsilon_{-i}^* = -(n-1)/n < 0$ for all i, and equilibrium welfare losses are given by:

$$\widetilde{L}(n, s_1^*, H^*) = 1 - \frac{n(n + 2H^*)}{(n + s_1^*)^2}.$$

The market share of the largest firm and the Herfindahl index play similar roles as in Cournot-Nash ($\upsilon_{-i}^* \equiv 0$); the difference is that the number of firms now also plays a crucial role – because it determines the endogenous competitive intensity as per υ_{-i}^*.

With symmetric firms, welfare losses then become $\widetilde{L}(n) = 1/(n^2 + 1)^2$. Losses are now of order $1/n^4$, and thus vanish *extremely* quickly as the number of firms rises. In effect, n firms now behave like n^2 Cournot competitors; even in a duopoly, losses are only 4 per cent. The reason is that the conduct becomes endogenously more competitive with more firms; in addition to having "Cournot with more firms", "Cournot becomes more like Bertrand". Intuitively, there is more scope to manipulate rivals' behaviour if they are more numerous.

With asymmetric firms, the key point is that, given more intense competition, lower-cost firms capture larger market shares than under Cournot-Nash.[18] Turned on its head, this means that a weaker firm can sustain a given market share only if its cost disadvantage is less pronounced than under Cournot-Nash. This additional efficiency effect strongly limits welfare losses.

Ritz (2014a) shows that welfare losses now remain "small" (less than 5 per cent) for many empirically relevant market structures. A simple sufficient condition is that the market share of the largest firm is no larger than 35 per cent. Welfare losses are always small if firms are not too symmetric or are sufficiently numerous (both in contrast to Cournot-Nash). In the numerical example with $s_1^* = 40$ per cent, $s_2^* = 30$ per cent, $s_3^* = 20$ per cent, and $s_4^* = 10$ per cent, welfare losses are just below 5 per cent (instead of 18 per cent under Cournot-Nash). These insights also extend fairly widely to non-linear demand systems.[19]

---

17 Instead using a differentiated-products Bertand model in which prices are strategic complements would lead to firms choosing to place *negative* weight on sales revenue ($\varphi_i^* < 0$), which seems at odds with empirical observation. In related work, Miller and Pazgal (2001) show that the equilibrium outcomes (and hence welfare) under differentiated Cournot and Bertrand can be *identical* if delegation contracts instead take the form of relative profits, e.g., $\Omega_i = \pi_i - \gamma_i \pi_j$ (for a fixed $n = 2$). While competition is as such tougher under Bertrand, this is exactly offset by the "soft" equilibrium contract featuring $\gamma_i^* < 0$ – while $\gamma_i^* > 0$ under Cournot (strategic substitutes).

18 Boone (2008) pursues this logic to develop a novel measure of competition based on how the relative profits of an efficient and a less efficient firm diverge more strongly when competition is more intense. Also related, Aghion and Schankerman (2004) study the welfare impacts of policies designed to enhance competition, and the political economy of their support, in a differentiated-products model with asymmetric costs.

19 In related work on restructured electricity markets, Bushnell, Mansur and Saravia (2008) emphasize how retail market commitments by vertically integrated players play a similar role to forward sales in Allaz and Vila (1993) – and how such long-term commitments can substantially improve market performance.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

3.3 Discussion

The above welfare quantifications hold equally if firms' products are vertically differentiated in the eyes of consumers, due to (known) differences in quality. In particular, if firm i faced a demand curve $p_i = \phi_i + p(Q)$ where ϕ_i is a measure of vertical product differentiation, then the first-best has the firm with the highest "value-added", $\max_i \{\phi_i - c_i\}$, produce all output. At equilibrium, higher-quality firms tend to have too small market shares from a social viewpoint. However, like cost differentials, differences in product quality are fully captured in firms' observed market shares, allowing for welfare to be estimated.

Welfare losses, in practice, will be lower if the first-best outcome is not the relevant benchmark for comparison. For example, the most efficient firm may not apply to supply q^{fb} because of capacity constraints or the government intervention that would be required to achieve first-best itself causes other welfare-reducing distortions. Welfare losses relative to any second-best optimum will be smaller.

These models can also speak to merger analysis. For example, as long as the *post*-merger market structure is sufficiently symmetric under Cournot-Nash or the largest firm has a market share of less than 35 per cent with delegation, then welfare "losses" remain small even after one or several mergers.[20] In this sense, the welfare impact of the mergers is limited, and there may be little rationale for policy intervention. Note that this is a different perspective from the usual approach in merger analysis: instead of testing whether or not a merger *reduces* in welfare, it focuses on whether the *level* of welfare losses remains "small" post merger (regardless of the direction of change).

Conversely, welfare losses would be higher if either the mode of competition in the industry is (tacitly) collusive or if the approximate welfare standard is skewed more strongly toward consumers, e.g., $W_\lambda = \lambda\Pi + CS$ with $\lambda < 1$. Cournot-style equilibria with very concave demand (low-cost pass-through) often produce high W but only low CS – and hence possibly also low W_λ.[21] For example, Cournot-Nash equilibrium with linear demand yields $CS^*/W^{fb} = 1/(1 + s_1^*)^2$, so consumer losses due to market power will be substantial – and sometimes *very* large – unless the largest firm is itself small relative to market.[22]

Other strands of the literature develop related models with endogenous conduct that may have similar welfare properties that lie between Cournot-Nash and perfect competition. One example is supply function models in which firms choose a set of price–quantity pairs to supply rather than being restricted to price *or* quantity choices (Klemperer and Meyer, 1989; Green and Newbery, 1992). In more recent work, d'Aspremont, Dos Santos Ferreira and Gerard-Varet (2007) and d'Aspremont and Dos Santos Ferreira (2009) develop a related way of endogenizing conduct parameters. Although welfare results for some limiting cases and specific examples are known, I am not aware of any general welfare analysis for such models.[23]

---

20 Strictly speaking, this assumes that the underlying first-best welfare remains unaffected by the merger; this will be the case either if the most efficient firm is not involved in the merger, or if it does not experience any efficiency gains.

21 This may explain why policymakers often appear to have a distaste for low pass-through markets; while these often yield low deadweight losses, consumers typically capture only a small fraction of the gains from trade.

22 This formula can be obtained heuristically by setting $\xi = 0$ and $H^* = 0$ in the expression for $L(s_1^*, H^*, \xi)$; superimposing a zero Herfindahl in effect takes away industry profits.

23 Holmberg and Newbery (2010) study how deadweight losses vary with market structure, demand elasticity and capacity utilization in an application of the supply function approach to electricity markets.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

The finding that welfare losses due to market power are often quantitatively modest in Cournot-style models naturally leads to the question: what other market features could generate higher losses? One possibility is horizontal product differentiation, which confers additional market power on firms (Corchón and Zudenkova, 2009). Another possibility is welfare losses due to different forms of asymmetric information. Vives (2002) studies a symmetric (Bayesian) Cournot model in which firms have private information on their costs, and argues that informational losses can outweigh those due to classical market power. Its effect on deadweight losses is of order $1/n$, while that of market power is of order $1/n^2$. Put differently, a larger number of firms is more effective at curbing market power than reducing informational distortions. It would be interesting to know more about how such effects play out with (*ex ante*) firm heterogeneity.[24]

4 SOCIAL COSTS OF ENDOGENOUS ENTRY

Recent work provides several refinements to the classic result that, in symmetric oligopoly, there is a tendency towards "excess entry": more firms enter than would be chosen by a social planner (Mankiw and Whinston, 1986).[25]

In the long run, firms decide endogenously on whether to enter a market (at some cost, which is sunk). Amir, De Castro and Koutsougeras (2014) show for Cournot models that excess entry arises if and only if there is "business stealing": each entrant, to some degree, captures sales from incumbents rather than serving new customers; per-firm output satisfies $q'(n) < 0$.

Hence there is wedge between private and social incentives: some of an entrant's profits are a transfer from incumbent firms but yield no social gain; since entry is costly, this wedge matters for welfare. In free-entry equilibrium, each individual firm is too small from a social perspective.[26]

4.1 Quantifying Welfare Losses Due to Excess Entry

Most of the existing literature examines a second-best setting in which the social planner cannot influence post-entry pricing, and focuses on qualitative results. In more recent work, Corchón (2008) quantifies the welfare losses L arising in a symmetric Cournot free-entry

24 In recent work, Gabaix et al. (2016) highlight how price–cost margins (rather than welfare) under (symmetric) monopolistic competition can be much less sensitive to the number of firms than under Cournot. They show that, in a random utility model in which goods are homogeneous but consumers are affected by random Gaussian "taste" shocks, markups are asymptotically proportional to $1/\sqrt{\ln(n)}$. One interpretation is behavioural: "consumer confusion" not captured in standard models of imperfect competition may result in significantly higher prices – even in "large" markets.

25 Taken literally, the policy implication is that entry should be regulated or otherwise restricted. By contrast, under perfect competition the degree of entry by firms is welfare optimal; more entry is then always a good thing for society.

26 The same conclusion applies with a moderate degree of horizontal product differentiation, so each entrant adds only little extra variety of value to consumers. However, the result can be reversed, leading to "insufficient entry", if competition in the market is very tough (e.g., undifferentiated Bertrand), even though at most by one firm "too few" (Mankiw and Whinston, 1986). In recent work, Bertoletti and Etro (2016) unify many existing results from endogenous-entry models (with symmetric preferences and symmetric firms), covering Bertrand, Cournot, and monopolistic competition.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

equilibrium, relative to the first-best social optimum – in which a single firm enters and price equals marginal cost. Welfare losses under free entry are sometimes very large, even with symmetric firms, because of the further cost misallocation.

Similar to the previous section, the approach is based on observables as far as possible. Assuming a free-entry equilibrium, the number of firms n is observed from market data.[27] The fixed cost of entry K is obtained as follows. (This can also be thought of as a fixed investment cost or R&D outlay required for market entry.) Let $\pi(n)$ denote per-firm Cournot profits (where $\pi'(n) < 0$). Since the nth firm decided to enter, $K \leq \pi(n) \equiv K_{\max}$, while the $(n+1)$th firm staying out implies that $K > \pi(n+1) \equiv K_{\min}$. (This assumes a sufficiently large pool of potential entrants.) So the entry cost is bounded according to $K \in (K_{\min}, K_{\max}]$.

It is clear from Mankiw and Whinston (1986) that welfare losses are increasing in the size of the entry cost; indeed the social inefficiency disappears as the entry cost becomes small. Therefore, $L(n, K, \xi) \leq L(n, K_{\max}, \xi) \equiv L_{\max}$ and $L(n, K, \xi) > L(n, K_{\min}, \xi) \equiv L_{\min}$, where ξ is the familiar measure of (constant) demand curvature.

The limiting cases are instructive. First, with a large number of observed entrants in the industry, welfare losses tend to zero. In such cases, operating profits are driven down to almost zero, so the entry cost must have been tiny to have allowed so many firms to participate. Hence the outcome is essentially equivalent to perfect competition.

Second, with a very convex demand curve ($\xi \to 2$) industry profits are only a very small fraction of the overall surplus generated. Hence the entry cost sustaining n firms in the market cannot be very large, and so welfare losses are again tiny.

Third, and conversely, with a very concave demand curve ($\xi \to -\infty$), industry operating profits are very large relative to consumer surplus. So if some potential entrants nonetheless choose not to enter, then the fixed cost must be substantial – and so there is a lot of socially wasteful cost outlay. Indeed, if the fixed cost is large enough to wipe out all industry profits, then welfare losses tend to 100 per cent. Specifically, Corchón (2008) shows that $L_{\min} = (n-1)/n \geq 1/2$ while $L_{\max} = 1$.

This latter set of cases is interesting because it contrasts so strongly with a fixed number of firms. With exogenous n, welfare losses under symmetric Cournot tend to zero as $\xi \to -\infty$; the incentive for firms to withhold output disappears as they capture all surplus at the margin. By contrast, with endogenous n, the majority of this surplus is dissipated by fixed costs.

To get a feel for how welfare losses remain "large" in interior cases, consider the case with linear demand. Using the results in Corchón (2008), it is easy to check that $L_{\min}(4) \approx 21.8$ per cent while $L_{\max}(4)$ is just over 30 per cent. This is at least five to seven times as high as the loss of 4 per cent with an exogenous four firms. For a larger number of firms, the gap $[L_{\max}(n) - L_{\min}(n)]$ shrinks as per-firm profits decline. With ten firms, welfare losses are bounded by 13.5–16.0 per cent; they remain above 5 per cent until the number of firms exceeds 35.

4.2 Firm Heterogeneity and Endogenous Entry

More recent work on endogenous entry has relaxed the assumption that potential entrants are symmetric, allowing for differences in firms' marginal costs and in the timing of market

[27] This side-steps the problem of the integer constraint on n that arises when the number of firms is derived from market primitives on costs and demand.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

entry. By contrast, classic models of "excess entry" leave no room for competition to enhance productive efficiency via selection – and thus deprive it of one of its fundamental roles.

Vickers (1995) develops a simple Cournot example with unit-elastic demand (i.e., $p(Q) = S/Q$ with fixed industry revenue S) to illustrate how the adverse effects of entry may thus be overstated. Suppose that each firm discovers its unit cost (low or high) following entry; the industry already consists of three firms; the question is the welfare impact of a fourth entrant.

If the entrant ends up being high cost and at least two incumbents are low cost, then it finds it optimal not to produce, so the externality from entry is zero. Even if only one incumbent is low cost, the negative externality is less pronounced than under symmetry since business stealing mainly affects the high-cost incumbent.

The entry externality turns *positive* if two of the three incumbents are high cost; entry by a low-cost firm then induces one of the less efficient incumbents to quit, and the efficient incumbent again expands output post-entry.[28] Surprisingly the literature does not appear to have generalized this example to richer market structures or to different forms of competition.

Etro (2008) shows how a first mover facing endogenous entry of followers typically behaves "more aggressively" than under simultaneous moves, and how this is good for social welfare. This stands in contrast with Stackelberg leadership against a fixed number of firms, which is well known to be critically sensitive to the question of strategic substitutes (which leads to aggressive behaviour and a first-mover advantage) versus strategic complements (which yields a second-mover advantage).

Intuitively, endogenous entry means that the leaders' attention shifts away from the reactions of followers at the margin (are strategies substitutes or complements?) to how its behaviour affects entry, that is, their participation constraints. Since products are substitutes, more aggressive behaviour (more output or lower prices) always leads to a favourable response: rivals' non-entry (or exit) becomes more likely.

To illustrate, consider quantity competition with one leader and m potential followers. Demand is linear $p(Q) = 1 - Q$ and costs are zero – apart from the entry cost K.[29] The key point is that, with free entry of followers determined by a zero-profit condition, the number of *actual* entrants decreases with the leader's output. Etro (2006) shows that the equilibrium thus features strategic entry deterrence; the market leader produces $q_L = 1 - 2\sqrt{K}$, which prevents any entry, and the limit price is $p = 2\sqrt{K}$.

This simple example already has some interesting welfare implications. The price is higher than in the free-entry Cournot equilibrium (simultaneous moves), so consumers are worse off – contrary to the fixed-n Stackelberg logic. However, social surplus is nonetheless higher because of the profits made by the leader – which are associated with the saving on entry costs. The observed market structures are radically different: the market has flipped from n active firms with identical shares to a single quasi-monopolist.

Etro (2008) studies a general "aggregative game" in which each firm's profits depend on its own action and a summary statistic of those of its rivals combined, that is, firm i's payoff $\pi_i = \Pi(x_i, X_{-i}) - K$ where x_i is its own action (e.g., price or output) and $X_{-i} = \sum_{j \neq i} h(x_j)$ captures the "externalities" arising via the actions of other players, where $h(\cdot), h'(\cdot) > 0$. This

[28] The unit-elastic example is somewhat unusual in that an efficient incumbent regards rivals' outputs as a strategic complement.

[29] As is standard, the entry cost is assumed to be sufficiently low such that the market is not a natural monopoly. Both Nash and Stackelberg free-entry models converge to perfect competition as $K \to 0$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

setup nests as special cases quantity competition with differentiated products as well as price competition with logit and iso-elastic demand, amongst others. Typically the leader produces more than under simultaneous moves or prices lower than the followers; this achieves a Pareto improvement in the allocation of resources.

The general implication is that large market shares of leading firms in an industry can be good news for social welfare; this also restores the notion of a first-mover advantage that prevails under both price and quantity competition. The details of a Stackelberg free-entry equilibrium depend on firms' strategic variables (price or quantity), the nature of product differentiation, and the shape of their cost functions.[30]

Mukherjee (2012) builds on these insights to show that the "excess entry" result can be reversed in markets with leadership and endogenous entry. The model has a single leader that enjoys a unit-cost advantage relative to a tail of symmetric followers, and a linear homogeneous products demand curve. The analysis is again second-best in that the social planner chooses the number of followers taking as given that they will engage in Stackelberg competition with the leader post-entry.

The main novelty is a "business creation" effect: the leader's optimal response to an increase in the number of followers is to *raise* production, $dq_L/dn > 0$. The reason is that its output rises with the followers' cost, and does so more strongly if there are more of them. Intuitively, the leader meets more rivals with a "fighting response" that leverages its cost advantage. (More formally, the leader's optimal output is supermodular in its cost advantage and the number of follower-entrants.)

The key insight is that the excess entry result is reversed if the leader's cost advantage is sufficiently pronounced. Then the new business creation effect dominates the standard business-stealing effect (which still exists amongst the followers), and more followers than delivered by the market would be socially desirable.

4.3 Discussion

The welfare metric used in the literature on endogenous entry is social surplus, so that the productive inefficiency arising from excess entry counts. Instead, using consumer welfare, an extra entrant is always socially desirable as long as it reduces prices; the market, if anything, delivers insufficient entry.[31]

The additional welfare losses that arise with endogenous entry thus have a similar effect to placing less weight on consumer surplus in the social-welfare function. Lowering λ in $W_\lambda = \lambda\Pi + CS$ pays less attention to profits either for normative reasons or because these profits are dissipated in another way. Incorporating wasteful rent-seeking costs that firms incur in securing market power (Posner, 1975) has similar effects.

[30] Anderson, Erkal and Piccinin (2015) analyse the welfare impacts of changes that affect only a subset of firms in a market – such as a merger or a technology change – in a general aggregative-game setup. They show that the short-run impacts (e.g., a merger raises prices) of the change are often fully neutralized in the long run with endogenous entry (i.e., the merger has no impact on prices). The key condition is that the marginal entrants, who make zero profits, are not directly affected, e.g., by the merger – and their actions effectively pin down the behaviour of the aggregate (and hence prices) over the long run.

[31] Some exceptions to this baseline result are known. Chen and Riordan (2008) show how more firms can sometimes lead to higher prices in a discrete-choice model with product differentiation; see also Cowan and Yin (2008) who study a related Hotelling setup.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

A central conclusion is that surplus losses remain large with endogenous entry even with a considerable number of firms in the market. Again, this conclusion can be substantially altered in two-stage models of competition. When post-entry competition is more intense, the inferred entry cost for any given number of observed entrants is well below that of Cournot competition, and so the additional source of cost misallocation is also much smaller. Welfare losses already drop below 5 per cent whenever there are at least four to six observed entrants (Ritz, 2014a).

Models of excess entry make the (sometimes neglected) assumption that entry occurs sequentially. While this is often reasonable, there are other examples in which a potential entrant may not know what entry decision other firms have made. Cabral (2004) provides a second-best analysis in which entry is a simultaneous process and either happens immediately or takes time as in a war of attrition. If entry costs are fairly low, then the results from sequential models are fairly robust. However, with high entry costs, the details on the timing of the entry process become important and "insufficient entry" more likely: from a societal perspective, a firm may be too fearful of an "entry mistake" (more firms enter than the market can support).

Finally, Kremer and Snyder (2015) emphasize a related source of welfare losses that arises from under-entry. Suppose that there is only a single potential entrant and that, from a social standpoint, it would be efficient for this firm to invest. However, if this firm is unable to appropriate a sufficiently large fraction of the surplus as profits, it will choose not to enter; as a result, the market does not come into existence.[32] Kremer and Snyder (2015) provide worst-case bounds that take into account the possibility of such "zero entry", and argue that the resulting welfare losses can be large – for instance, in the pharmaceuticals industry in which consumers' valuations for products often vary widely.

5 CONCLUSION

The recent literature offers some new perspectives on a significant body of existing knowledge on oligopolistic competition and social welfare.

In a fairly broad class of oligopoly models, the division of surplus between firms and consumers is importantly determined by the rate of cost pass-through. Empirical estimates of pass-through across different markets thus offer indirect inference on welfare metrics. Yet pass-through is not a panacea in settings with firm heterogeneity, and the link between theory and the econometrics of pass-through still needs further tightening in future research.

The degree of welfare loss in widely used Cournot-style models is often surprisingly modest, even relative to first-best and with significant industry concentration. Under Cournot-Nash competition, losses can be significantly higher due to cost asymmetries between firms yet their adverse impact is strongly limited in two-stage models with tougher competitive conduct. Losses are also typically much higher under a consumer welfare standard. Future research could examine more closely the interaction between heterogeneity in firms' costs and asymmetric information.

[32] By contrast, the above models consider the welfare implications of endogenous entry where some entry has indeed occurred – so there exists an observed distribution of firms' market shares.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Market performance is similarly reduced in dynamic models featuring "excess entry" that dissipates a significant fraction of firm profits. Recent work has extended these results to allow for Stackelberg leadership as well as differences in firms' costs. Both can be good news for social welfare, especially if the market leader also enjoys a cost advantage. Future research may focus on how these results map onto the empirical study of specific markets.

REFERENCES

Aghion, Philippe and Mark Schankerman (2004). On the welfare effects and political economy of competition-enhancing policies. *Economic Journal* 114, 800–824.

Aiginger, Karl and Michael Pfaffenmayr (1997). Looking at the cost side of "monopoly". *Journal of Industrial Economics* 45, 245–267.

Allaz, Blaise and Jean-Luc Vila (1993). Cournot competition, forward markets and efficiency. *Journal of Economic Theory* 59, 1–16.

Amir, Rabah, Luciano de Castro and Leonidas Koutsougeras (2014). Free entry versus socially optimal entry. *Journal of Economic Theory* 154, 112–125.

Anderson, Simon P. and Régis Renault (2003). Efficiency and surplus bounds in Cournot competition. *Journal of Economic Theory* 113, 253–264.

Anderson, Simon P., Nisvan Erkal and Daniel Piccinin (2015). Aggregative oligopoly games with entry. Working Paper at the University of Virginia, June 2015.

Armstrong, Mark and John Vickers (2010). A model of delegated project choice. *Econometrica* 78, 213–244.

Bagnoli, Mark and Ted Bergstrom (2005). Log-concave probability and its applications. *Economic Theory* 26, 445–469.

Bertoletti, Paolo and Federico Etro (2016). Preferences, entry and market structure. *RAND Journal of Economics* 47, 792–821.

Boone, Jan (2008). A new way to measure competition. *Economic Journal* 118, 1245–1261.

Bresnahan, Timothy F. (1989). Empirical studies of industries with market power. In: *Handbook of Industrial Organization – Volume 2*, Richard Schmalensee and Robert Willig (eds.), 1011–1057. Amsterdam: North-Holland.

Bulow, Jeremy and Paul Klemperer (2012). Regulated prices, rent seeking, and consumer surplus. *Journal of Political Economy* 120, 160–186.

Bulow, Jeremy and Paul Pfleiderer (1983). A note on the effect of cost changes on prices. *Journal of Political Economy* 91, 182–185.

Bushnell, James B., Erin T. Mansur and Celeste Saravia (2008). Vertical arrangements, market structure, and competition: An analysis of restructured U.S. electricity markets. *American Economic Review* 98, 237–266.

Cabral, Luís M.B. (1995). Conjectural variations as a reduced form, *Economics Letters* 49, 397–402.

Cabral, Luís M.B. (2004). Simultaneous entry and welfare. *European Economic Review* 48, 943–957.

Chen, Yongmin and Michael H. Riordan (2008). Price-increasing competition. *RAND Journal of Economics* 39, 1042–1058.

Corchón, Luis C. (2008). Welfare losses under Cournot competition. *International Journal of Industrial Organization* 26, 1120–1131.

Corchón, Luis C. and Galina Zudenkova (2009). Computing welfare losses from data under imperfect competition with heterogeneous goods. *International Journal of Industrial Organization* 27, 646–654.

Cowan, Simon and Xiangkang Yin (2008). Competition can harm consumers. *Australian Economic Papers* 47, 264–271.

d'Aspremont, Claude and Rudolphe Dos Santos Ferreira (2009). Price–quantity competition with varying toughness. *Games and Economic Behavior* 65, 62–82.

d'Aspremont, Claude, Rudolphe Dos Santos Ferreira and Louis-André Gerard-Varet (2007). Competition for market share or for market size: Oligopolistic equilibria with varying competitive toughness. *International Economic Review* 48, 761–784.

Etro, Federico (2006). Aggressive leaders. *RAND Journal of Economics* 37, 146–154.

Etro, Federico (2008). Stackelberg competition with endogenous entry. *Economic Journal* 118, 1670–1697.

Farrell, Joseph and Michael Katz (2006). The economics of welfare standards in antitrust. *Competition Policy International* 2, 3–28.

Farrell, Joseph and Carl Shapiro (1990). Horizontal mergers: An equilibrium analysis. *American Economic Review* 80, 107–126.

Fershtman, Chaim and Kenneth L. Judd (1987). Equilibrium incentives in oligopoly. *American Economic Review* 77, 927–940.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Gabaix, Xavier, David Laibson, Deyuan Li, Hongyi Li, Sidney Resnick and Casper G. de Vries (2016). The impact of competition on prices with numerous firms. *Journal of Economic Theory* 165, 1–24.
Green, Richard J. and David M. Newbery (1992). Competition in the British electricity spot market. *Journal of Political Economy* 100, 929–953.
Harberger, Arnold C. (1954). Monopoly and resource allocation. *American Economic Review* 44, 77–87.
Holmberg, Pär and David Newbery (2010). Supply function equilibrium and its policy implications for wholesale electricity auctions. *Utilities Policy* 18, 209–226.
Kimmel, Sheldon (1992). Effects of cost changes on oligopolists' profits. *Journal of Industrial Economics* 40, 441–449.
Klemperer, Paul D. and Margaret A. Meyer (1989). Supply function equilibria in oligopoly under uncertainty. *Econometrica* 57, 1243–1277.
Kremer, Michael and Christopher M. Snyder (2015). Worst-case bounds on R&D and pricing distortions: Theory and disturbing conclusions if consumer values follow the world income distribution. Working Paper at Harvard University, January 2015.
Lahiri, Sajal and Yoshiyasu Ono (1988). Helping minor firms reduces welfare. *Economic Journal* 98, 1199–1202.
MacKay, Alexander, Nathan H. Miller, Marc Remer and Gloria Sheu (2014). Bias in reduced-form estimates of pass-through. *Economics Letters* 123, 200–202.
Mankiw, N. Gregory and Michael D. Whinston (1986). Free entry and social inefficiency. *RAND Journal of Economics* 17, 48–58.
Miller, Nolan H. and Amit I. Pazgal (2001). The equivalence of price and quantity competition with delegation. *RAND Journal of Economics* 32, 284–301.
Miller, Nolan H., Marc Remer and Gloria Sheu (2013). Using cost pass-through to calibrate demand. *Economics Letters* 118, 451–454.
Mukherjee, Arijit (2012). Social efficiency of entry with market leaders. *Journal of Economics & Management Strategy* 21, 431–444.
Posner, Richard A. (1975). The social costs of monopoly and regulation. *Journal of Political Economy* 83, 807–828.
Reiss, Peter C. and Frank A. Wolak (2007). Structural econometric modeling: Rationales and examples from industrial organization. In: *Handbook of Econometrics – Volume 6A*, James J. Heckman and Edward E. Leamer (eds.), 4278–4415. Amsterdam: North-Holland.
Ritz, Robert A. (2008). Strategic incentives for market share. *International Journal of Industrial Organization* 26, 586–597.
Ritz, Robert A. (2014a). On welfare losses due to imperfect competition. *Journal of Industrial Economics* 62, 167–190.
Ritz, Robert A. (2014b). A new version of Edgeworth's paradox of taxation. *Oxford Economic Papers* 66, 209–226.
Ritz, Robert A. and Ansgar Walther (2015). How do banks respond to increased funding uncertainty? *Journal of Financial Intermediation* 24, 386–410.
Rostek, Marszena, Marek Weretka and Marek Pycia (2009). Discriminatory or uniform? Design of divisible good auctions. Working Paper at the University of Wisconsin-Madison, February 2009.
Shapiro, Carl (1989). Theories of oligopoly behavior. In: *Handbook of Industrial Organization – Volume 2*, Richard Schmalensee and Robert Willig (eds.), 329–414. Amsterdam: North-Holland.
Shapiro, Carl (1996). Mergers with differentiated products. *Antitrust*, Spring 1996. 22–30.
Tirole, Jean (1988). *The Theory of Industrial Organization*. Cambridge, MA: MIT Press.
Vickers, John (1985). Delegation and the theory of the firm. *Economic Journal* 95, 138–147.
Vickers, John (1995). Concepts of competition. *Oxford Economic Papers* 47, 1–23.
Vives, Xavier (1999). *Oligopoly Pricing: Old Ideas and New Tools*. Cambridge, MA: MIT Press.
Vives, Xavier (2002). Private information, strategic behavior and efficiency in Cournot markets. *RAND Journal of Economics* 33, 361–376.
Weyl, E. Glen and Michal Fabinger (2013). Pass-through as an economic tool: Principles of incidence under imperfect competition. *Journal of Political Economy* 121, 528–583.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

PART II

DYNAMIC GAMES IN INDUSTRIAL ORGANIZATION

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

8. Dynamic games

Klaus Ritzberger

1 INTRODUCTION

To the interested reader of the *New York Times* (Passell, 1994) a game is about choosing *cooperate* or *defect* in a prisoner's dilemma, hence, an entirely static affair. This is a pity, because the issue at stake concerns non-cooperative games. And a game is *non-cooperative* if it has complete rules. An obvious example, and the inspiration for the discipline's nickname, is a parlor game. Yet, a parlor game is not a static one-shot affair. Or do you play chess by submitting a full strategy – a complete plan for all possible contingencies – to the umpire?

For most parlor games the interaction of players takes place in a sequence of moves as dictated by the rules. Hence, it is an inherently dynamic affair. When the rules are complete, every possible succession of attacks, counterattacks, and defenses that are consistent with the rules can be mapped out in a flow-chart. In fact, such a flow-chart represents a device to verify that the rules are complete, and is technically known as a "tree." Therefore, a game is non-cooperative if and only if it has such a dynamic representation by a tree, technically known as an "extensive form." In this sense *every* non-cooperative game is dynamic.

This may sound surprising given that the literature, somewhat imprecisely, refers to "dynamics" only in the presence of an explicit and common time dimension. Examples include repeated games, stochastic games (Shapley, 1953), or games in continuous time (also called "differential games"). Yet, in a non-cooperative game there is in principle no reason that players are aware of a common time axis, even if they perceive a personalized time-ordering of their own moves. (An example of this was already given by Kuhn, 1953, p. 199.) Still, their interaction is dynamic in the sense that it evolves over time.

For this reason this chapter will deal with games in *extensive form*. Of course, that does not exclude the narrower classes of repeated, stochastic, or differential games, as those are also games in extensive form. Indeed, differential games will be touched upon now and then, but no comprehensive overview will be provided. Instead, the development of extensive form game theory, historical and recent, will be outlined and its significance for applications will be explained. The emphasis will be on concepts that are peculiar to extensive form analysis, like the role of backwards induction and their strengths and weaknesses.

There is no intention to deny the virtues of normal form analysis. After all, the most basic solution concepts, like Nash equilibrium, rationalizable, or iteratively undominated strategies, rely on it. In fact, discovering the idea of a pure strategy has had a major impact on the development of game theory. It enabled the translation of the extensive form representation of a game (von Neumann, 1928) into the mathematically more transparent problem of a (payoff or utility) function on a product set (of pure-strategy combinations). Some have even argued that this *normal form* representation contains everything that is relevant to rational decision makers (e.g., Kohlberg and Mertens, 1986). Whether or not this is a valid argument will

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

be seen to depend on which decision-theoretic rationality notion one is willing to adopt – consistent with the idea that game theory is interactive decision theory (Aumann, 1987). Whatever view one holds about that, normal form games suppress the dynamics of strategic interaction. Extensive form games, on the other hand, live off this dynamics and, therefore, have an intuitive appeal that is often very informative in applications.

It is trivially true that every normal form game has an extensive form representation, one where all players simultaneously choose a strategy once and for all at the root of the tree. In this sense normal form games are indeed also non-cooperative games (even though I will raise some doubts about that later on). The key difference between normal and extensive form games rests with what are the *primitives*.

To highlight this, let us first agree that the rules of the game specify only outcomes, like win or lose, sums of money to be transferred, market shares, or the like, but no payoffs (or utilities). Such a pure representation of the rules is often called a *game form*. A game form turns into a *game* when preferences of players over outcomes are added. Then, fixing a set of players, in a normal form the two primitives are the (product) set of pure strategy combinations and an outcome function that maps each strategy combination into an outcome. By contrast, in an extensive form the two primitives are the tree (that incorporates the outcomes) and the players' sets of choices that determine information sets. In particular, for the extensive form representation strategies are derived objects, not primitives.

This distinction determines on which objects assumptions can be imposed that give rise to a particular application. In the normal form the assumptions are put on strategies or on the outcome function. In the extensive form the assumptions are put on the tree or the players' choices. Since the latter formalize the evolution of the players' information structure, it is an explicitly dynamic model of strategic interaction that gives rise to observable implications. This is not true for the normal form as strategies can differ off the equilibrium path, at counterfactual events that cannot be observed. Therefore, the dynamic model of the extensive form is better suited to studying applications that can be taken to the data – and this is one reason for its popularity in industrial organization.

The chapter is organized as follows. Section 2 introduces three possible definitions of game trees and extensive forms and discusses their virtues and vices. Section 3 comes back to how the normal form is related to the extensive form and explains the concept of perfect recall and its significance for applications (Kuhn's theorem). Section 4 discusses two applications.

2 EXTENSIVE FORMS

The idea of how to represent a game in extensive form by a tree goes back to an early paper by von Neumann (1928), but it was not until the *Theory of Games and Economic Behavior* (von Neumann and Morgenstern, 1944) that it was recognized beyond the initiated circles. The formalizations proposed since differ in details, but share at least three ingredients: first, a set of *players*, throughout denoted by I; second, a *tree* that captures how an omniscient outside observer sees the evolution of play; and third, a model for the evolution of the players' private information, typically comprising choice sets and/or information sets. The set of players $i \in I$ will not pose any problems, except that it will mostly be taken to be finite. The various incarnations of the other two ingredients will now be described more formally.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

2.1 Trees

In their foundational book, von Neumann and Morgenstern proposed to formalize a game tree, in line with statistics and decision theory, as a collection of subsets of a common underlying set of plays, outcomes, or states (1944, p. 60). Yet, for reasons that are elusive they imposed restrictive assumptions, for instance, that the number of moves that come before any node in a given information set is constant (1944, p. 73). This limitation caused Kuhn (1953) to generalize the model and recast the tree as a graph. His remained the canonical formalization until the turn of the twenty-first century (see, e.g., Selten, 1975; Kreps and Wilson, 1982). The only exception appeared first in a paper by Harris (1985) and was then taken up in a textbook by Osborne and Rubinstein (1994). They perceived a tree as a collection of sequences. This century, the original formalization of an event tree launched a comeback in a series of papers by Alós-Ferrer and Ritzberger (2005a, 2005b, 2008, 2013, henceforth AR1a, AR1b, AR2, and AR3). Each of those three representations will now be discussed in turn, beginning with von Neumann and Morgenstern's original formulation in its modern incarnation.

2.1.1 Event trees

Throughout, W will denote a fixed set of outcomes or plays. A play is a maximal history of the game, from the beginning to the end – if there is an end. Since a play keeps track of everything that happens along it, the set W of plays is the appropriate domain for the players' preferences. Statisticians may also want to think of W as a state space. In the early literature W is typically taken to be finite, but this will not be necessary, that is, W may well be a continuum. The following definition captures what AR3 call a *discrete game tree*.

Definition 1 A **discrete event tree** is a pair $T = (W, N)$ consisting of a non-empty set W of *plays* and a collection of non-empty subsets $x \in N$ (the *nodes*) of W, partially ordered by set inclusion, such that $W \in N$, $\{w\} \in N$ for all $w \in W$, and
(GT1) $h \subseteq N$ is a chain if and only if there is $w \in W$ such that $w \in \cap_{x \in h} x$,[1]
(GT2) every chain in the set $X \equiv N \setminus \{\{w\}\}_{w \in W}$ of *moves* has a maximum and either an infimum in the set $E \equiv \{\{w\}\}_{w \in W}$ of *terminal nodes* or a minimum.[2]

In statistics an *event* is a non-empty subset of the state space. The same applies here. It begins at the *root* $W \in N$ where nothing has been decided yet, so everything is still possible. At this all-encompassing event someone takes a decision and thereby narrows down the possible outcomes. Continuing inductively, suppose that a move $x \in X \equiv N \setminus \{\{w\}\}_{w \in W}$ has materialized. At x a further decision discards more outcomes in x and, hence, a smaller node $y \in N$ with $y \subset x$ is reached. This continues until a terminal node $\{w\} \in N$ finally pins down the realized play – which follows from the only-if part of (GT1). By the if part of (GT1) every two events (nodes) $x, y \in N$ must either be disjoint or one must contain the other; hence, succession translates into a refinement of information.

1 A subset of a partially ordered set is a *chain* if the order relation restricted to the subset is complete.
2 Maximum, minimum, supremum, and infimum are with respect to set inclusion. Henceforth $\subset$ denotes proper inclusion and $\subseteq$ denotes inclusion or equality. Note that AR3 call an event tree a *(discrete) game tree*, hence (GT1) and (GT2). They also denote the tree by $(N, \supseteq)$, which is justified by AR1, Theorem 1, p. 775.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

For each node $x \in N$ let $\uparrow x = \{y \in N \,|\, x \subseteq y\}$ (the past) and $\downarrow x = \{y \in N \,|\, y \subseteq x\}$ (the future). Then, nodes in an event tree can be classified as follows. A node $x \in N \setminus \{W\}$ is *finite* if $\uparrow x \setminus \{x\}$ has a minimum and *infinite* if $x = \inf \uparrow x \setminus \{x\}$. All infinite nodes must be terminal (by Theorem 1(b) of AR3). Let $F(N) \subseteq N$ denote the set of finite nodes together with the root $W \in N$. Its complement, the set of infinite nodes, consists of those that are reached only after infinitely many decisions – thus justifying its name. Whether or not infinite terminal nodes are included in the tree does not matter (AR3, Proposition 4). If they are, the tree is called *complete*. For a finite node $x \in F(N) \setminus \{W\}$ let $p(x) = \min \uparrow x \setminus \{x\} \in N$ denote its *immediate predecessor*.

Definition 1 can be shown to represent the most general notion of a game tree in the literature, save for one aspect incorporated in (GT2). Condition (GT2) demands that every chain of moves has a maximum and either also a minimum or an infimum in the set of singletons (terminal nodes). The first part, existence of a maximum or "up-discreteness," corresponds to the existence of immediate successors. This can be shown to be necessary. Without it, strategies may induce no outcome at all or a whole continuum thereof (AR2, Theorems 2 and 3). The second part, existence of a minimum or a terminal infimum (or "down-discreteness"), corresponds to the existence of immediate predecessors. This one, by contrast, is a convenience property. It holds in all known applications, with the single exception of differential games. Yet, it simplifies working with the most general definition (given as Definition 4 of AR1b) that also encompasses continuous-time games.[3] Hence, for the present purpose Definition 1 will do. In particular, note that it allows for infinitely many choices at a move as well as for an infinite horizon.

2.1.2 Graph trees

Since graphs are more abstract, their introduction takes some preparation. A *graph* is a pair (N, R) consisting of a non-empty set N, whose elements $x \in N$ are called nodes or vertices, together with a binary relation $R \subseteq N \times N$, called edges or incidence. If the relation R is symmetric, the graph is called *undirected*.[4] Otherwise it is *directed* (or a digraph).

A key notion for graphs is as follows: A finite subset $\{x_1, \ldots, x_k\} \subseteq N$ for some $k \in \mathbb{Z}_{++}$ is a *path* if $(x_\ell, x_{\ell+1}) \in R$, no edge $(x_\ell, x_{\ell+1})$ and no node x_ℓ being repeated, for all $\ell = 1, \ldots, k-1$. Node x_k may be repeated, though. If $x_k = x_\ell$ for some $\ell = 1, \ldots, k-1$, then the path $\{x_\ell, \ldots, x_k\}$ is called a *cycle* if $\ell < k-1$ or a *loop* if $\ell = k-1$. A graph that contains no cycles or loops is called *acyclic*.

The notion of a path gives rise to a new relation on N, defined by taking the transitive closure.[5] In particular, say that $x \in N$ is *connected to* node $y \in N$ if there is a path $\{x_1, \ldots, x_k\}$ with $x_1 = x$ and $x_k = y$. Clearly, if R itself is transitive, then its transitive closure agrees with R, otherwise it is larger than R. A graph in which for all distinct $x, y \in N$ either x is connected to y or y is connected to x is a *connected graph*.

For an acyclic directed graph (N, R) the transitive closure of R is irreflexive and transitive and denoted by $>$. Transitivity is by construction and irreflexivity by the absence of loops. Taking the union with equality results in a reflexive, transitive, and antisymmetric

---

[3] Strictly speaking there is an even more general formulation, proposed by AR1a. The extra generality, however, concerns only allowing trivial structures, like nodes with only one successor.

[4] A binary relation R is *symmetric* if $(x, y) \in R$ implies $(y, x) \in R$, *reflexive* if $(x, x) \in R$, *transitive* if $(x, y) \in R$ and $(y, z) \in R$ imply $(x, z) \in R$, and *antisymmetric* if $(x, y) \in R$ and $(y, x) \in R$ imply $x = y$, for all $x, y, z \in N$.

[5] The transitive closure of a relation is the smallest transitive relation that contains the original one.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

relation $\geq$, i.e., $x \geq y$ if and only if either $x > y$ or $x = y$ is a partial ordering. Reflexivity and transitivity follows from the definition. If $x \geq y$ and $y \geq x$ would hold with either $x > y$ or $y > x$, then $x > x$ by transitivity would contradict irreflexivity of the transitive closure; hence, $x \geq y$ and $y \geq x$ must imply $x = y$, viz. antisymmetry. This union of equality with the transitive closure, $\geq$, is called the *induced (partial) ordering*. The following definition refers to this induced ordering:

Definition 2 A **graph tree** (N, R) is an acyclic directed graph that has a maximum $x_0 \in N$ (called the root) with respect to the induced ordering $\geq$ of R such that
(ST1) for all $x \in N$, the set $\uparrow x \equiv \{y \in N \,|y \geq x\}$ is a path with $x_0 \in \uparrow x$, and
(ST2) for all $x, y \in N$, if $x > y$ then there is $z \in N$ such that $x > z$ and neither $y \geq z$ nor $z \geq y$, where $>$ denotes the transitive closure of R.[6]

A concept closely related to this definition is as follows. A directed graph (N, R) is an *arborescence* if there is a distinguished node $x_0 \in N$ (the root) such that every other node $x \in N$ is connected to x_0 by a unique path. Hence, an arborescence satisfies (ST1). For instance, $N = \{x, y, z\}$ together with $R = \{(x, z), (z, y), (y, x)\}$ is an arborescence, if one of the three nodes is declared the root (because a path cannot repeat edges). Thus, an arborescence may look like a cycle, which a graph tree cannot. Another example of an arborescence is $(\mathbb{Z}_{++}, R)$ with $R = \{(n, n-1) \,|n > 1\}$ and root $x_0 = 1$. For this example the transitive closure $>$ of R is simply the reversed natural ordering of the positive integers $n \in \mathbb{Z}_{++}$. This is an example of an arborescence that fails (ST2) as the latter is a non-triviality requirement that excludes pure chains.

A peculiarity of a graph tree is that an infinite play does not end at a terminal node; instead it never ends. The corresponding play is nevertheless well defined – as a maximal chain with respect to $\geq$, where a chain $c \subseteq N$ is maximal if there is no $x \in N \setminus c$ such that $c \cup \{x\}$ is a chain.[7] Hence, the set W of plays is the set of maximal chains and again provides the appropriate domain for the players' preferences.

For a graph tree (N, R) denote by W its set of plays, and for each $x \in N$ let $W(x) = \{w \in W \,|x \in w\}$ denote the set of plays passing through node x. This provides a way to make an equivalence between graph trees and event trees precise.

Proposition 1 *(a) If (N, R) is a graph tree, then its set of plays W together with the set $\{W(x) \,|x \in N\} \cup \{\{w\}\}_{w \in W}$ is an event tree.*

(b) If (W, N) is an event tree, then the set $F(N)$ of finite nodes together with the binary relation $R = \{(x, y) \in F(N) \times F(N) \,|y = \min \uparrow x \setminus \{x\}\}$ is a graph tree.

Proof (a) First, observe that for all $x, y \in N$ it holds that $W(y) \subseteq W(x)$ if and only if $x \geq y$, i.e., the partially ordered set $(N, \geq)$ and $(\{W(x) \,|x \in N\}, \supseteq)$ are order isomorphic.[8] To verify (GT1) let $h' \subseteq N' \equiv \{W(x) \,|x \in N\}$ be a chain, that is, for all $x', y' \in h$ either $x' \subseteq y'$ or $y' \subset x'$. By order isomorphism there is a chain $h \subseteq N$ such that $h' = \{W(x) \,|x \in h\}$. By the

[6] AR3, Definition 7, calls a graph tree a "simple tree," hence (ST1) and (ST2).

[7] In a partially ordered set every chain is contained in a maximal chain by the Hausdorff maximality principle – a version of the axiom of choice (see, e.g., Hewitt and Stromberg, 1965, Ch. 1).

[8] An order isomorphism between two partially ordered sets is a bijection that preserves the partial orderings on domain and codomain.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Hausdorff maximality principle h is contained in a maximal chain $w \in W$. It follows that $w \in W(x)$ for all $x \in h$, as required by the only-if part of (GT1). To see the if part, let h' be a subset of N' such that there is some $w \in W$ with $w \in W(x)$ for all $W(x) \in h'$. It follows that the set $h \subseteq N$ such that $h' = \{W(x) \,|x \in h\}$ is a chain, thus h' is also a chain.

To see (GT2), let $w, \hat{w} \in W$ with $w \neq \hat{w}$. Since w and $\hat{w}$ are maximal chains in N, there are $x, \hat{x} \in N$ such that $x \in w$ and $\hat{x} \in \hat{w}$, but $x \notin \hat{w}$ and $\hat{x} \notin \hat{w}$. It follows that $w \in W(x) \setminus W(\hat{x})$ and $\hat{w} \in W(\hat{x}) \setminus W(x)$. Hence, $(\{W(x) \,|x \in N \,, \supseteq\})$ is a game tree in the sense of Definition 4 of AR1b. By the existence of a maximum $x_0 \in N$ with respect to $\geq$, this game tree is rooted. By (ST1) and order isomorphism $\uparrow W(x)$ is finite and, hence, $W(x)$ is a finite node in this game tree for all $x \in N$. Therefore, $(\{W(x) \,|x \in N\}, \supseteq)$ is regular (AR2, Definition 2.2(c)) and rooted, hence, it is discrete by Theorem 1 of AR3. Adding the singletons $\{\{w\}\}_{w \in W}$ can only add infinite terminal nodes.[9]

(b) Let $R = \{(x, y) \in F(N) \times F(N) \,|y = p(x)\}$, denote by $> \subseteq F(N) \times F(N)$ the transitive closure of R, and by $\geq \subseteq F(N) \times F(N)$ the union of equality (the diagonal) with $>$. Then $\geq$ is simply (weak) set inclusion and, therefore, a partial ordering. If there were a cycle $\{x'_1, \ldots, x'_k\} \subseteq F(N)$ with $x'_1 = x'_k$, then $\geq$ cannot be antisymmetric, because for any two distinct $x', y' \in \{x'_1, \ldots, x'_k\}$ it would be the case that $x' \geq y'$ and $y' \geq x'$, even though $x' \neq y'$. Likewise, if there were a loop $\{x', x'\}$, then $x' > x'$ would contradict that $>$ (viz. proper set inclusion) is irreflexive. Hence, that $\geq$ is a partial ordering implies that $(F(N), R)$ is acyclic.

As $W = x_0 \geq x'$ for all $x' \in N$, it follows that $x_0 \in \uparrow x' = \{y' \in F(N) \,|y' \geq x'\}$ for all $x' \in F(N)$. Because in the event tree $\uparrow x = \{y \in N \,|y \supseteq x\}$ is finite for all $x \in F(N)$ by Theorem 1(c) of AR3, so is $\uparrow x'$ for all $x' \in F(N)$; thus, $\uparrow x$ is a path and (ST1) holds. (ST2) follows from Proposition 5(b) of AR3. □

By this result Definitions 1 and 2 are equivalent up to the inclusion or exclusion of infinite terminal nodes. Therefore, even though Definition 2 is again not the most general definition, it suffices for the present purpose. In particular, infinitely many choices are available at a move and a potentially (countably) infinite horizon are still covered. To a certain extent it is a matter of taste which definition is used.

It should be noted, though, that the graph formulation does not admit a generalization as event trees do. The reason is most easily seen for differential games, i.e., continuous-time problems. Because the real numbers are not well ordered, for a given node at "time" $t \in \mathbb{R}$ there is simply no hope of identifying immediate successors or predecessors. Hence, the statement $(x, y) \in R$, which is to be read as "$y \in N$ immediately precedes $x \in N$," has no meaning. Yet, as pointed out before, such problems appear to be outside the realm of dynamic game theory regardless.

2.1.3 Sequence trees

There is a third definition of a tree, introduced by Harris (1985) and later popularized by Osborne and Rubinstein (1994, p. 200) in their textbook. The present exposition will follow the latter source, because Harris's paper is concerned with perfect information games. The formulation starts from a universal set of *actions* A, like "move left," "pay a certain amount of money," "shoot," "stop producing," etc., which lists everything that players may do in the course of a game after all sorts of histories. Even though in many applications A may

[9] This proof of statement (a) is almost identical to the proof of Proposition 5(a) of AR3.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

be finite, there is once again no reason for that; hence, A may well have the cardinality of the continuum. The second ingredient is a subset H of the infinite product A^∞.

Definition 3 A **sequence tree** is a pair (A, H) consisting of a set A of actions and a set $H \subseteq A$ of sequences (finite or infinite) from A such that
(OR1) the empty sequence belongs to H, i.e., $\emptyset \in H$,
(OR2) if $(a_k)_{k=1}^K \in H$ (where K may be infinity) and $L < K$, then $(a_k)_{k=1}^L \in H$,
(OR3) if an infinite sequence $(a_k)_{k=1}^\infty$ from A satisfies $(a_k)_{k=1}^L \in H$ for every positive integer $L \in \mathbb{Z}_{++}$, then $(a_k)_{k=1}^\infty \in H$.

Each element of H is a *history*, and each coordinate a_k of a history is an action taken by some player from A. A history $(a_k)_{k=1}^K \in H$ is *terminal* (or maximal) if it is infinite or if there is no $a_{K+1} \in A$ such that $(a_k)_{k=1}^{K+1} \in H$. Otherwise it is non-terminal. The set of terminal histories corresponds to the set of plays and is denoted by W. The set of actions available at any non-terminal history $x \in H \setminus W \equiv X$ is the set $A(x) = \{a \in A \,|\, (x, a) \in H\}$, which must be non-empty by the definition of X (the set of moves or non-terminal histories). For each $t = 1, 2, \ldots$ denote by $A_t = \left\{a \in A \,\middle|\, \exists\, (a_k)_{k=1}^t \in H : a_t = a\right\}$ the set of actions available at stage t.

Once again, for a given non-terminal history $x \in X$ the set $A(x)$ of available actions may be infinite; and by construction the time horizon may also be (countably) infinite. In fact, Definition 3 is slightly more general than the other two, because it covers the arborescence $(\mathbb{Z}_{++}, R)$ with $R = \{(n, n-1) \,|\, n > 1\}$ and with root $x_0 = 1$. That is, a sequence tree may contain trivial structures, where every non-terminal history $x \in X$ has only a single successor $y = (x, a) \in H$, and nothing else can happen between x and y. Therefore, a result like Proposition 1 is infeasible. Yet, if Definition 3 were amended by a non-triviality condition, then it can again be shown to be equivalent to Definition 1 in a precise sense (see Example 10 of AR3). Since pure chains are uninteresting for game theory, all three definitions can be taken as equivalent for practical purposes.

Unlike event trees (Definition 1), sequence trees do not admit a generalization to continuous-time problems. This is because a sequence is a function from the natural numbers to some set and, therefore, can take at most countably many values. Yet, since differential games pose problems anyway, this amounts to a minor loss of generality.

2.2 The Problem of Information

The tree represents the evolution of information of an omniscient outside observer or an umpire. The second ingredient of an extensive form representation, therefore, has to concern the players' private information – this is why the phrase "the problem of information" occurs in the title of Kuhn's (1953) seminal paper. With regard to formalizing that problem the dividing line between approaches is somewhat different, as here Kuhn's approach agrees with von Neumann and Morgenstern's. They both emphasize information sets. It was not until the recent revival of event trees (as in Definition 1) that choices were discovered as the true primitive.

Before turning to formal definitions, let us agree on two conventions. First, chance moves will not be treated explicitly; instead chance (or nature) is seen as one of the players and, if present, assigned the index $0 \in I$. Personal players will mostly be denoted by $i = 1, \ldots, n \in I$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

In any case, the player set will remain fixed. The reason not to treat chance separately is that this would require a specification of some randomized strategy for nature. Yet, in the absence of any cardinality assumptions this poses problems (see, e.g., Aumann, 1961, 1963, or Feldman and Gilles, 1985). For instance, how would you feel about having to simulate a uniform distribution on the unit interval, if all that you have at your disposal is a two-sided coin that you can flip only once?

Second, because any given game could be played by many different player sets (with the same cardinality), preferences (or payoffs or utilities) will not figure in the definitions. They will be added in the end, when a game form is turned into a game. To emphasize that, the term "game" will not appear in the definition of a "form."

2.2.1 Discrete extensive forms

As for notation, given an event tree (W,N) as in Definition 1 and a subset $a \subseteq W$ that is a union of nodes, define $\downarrow a = \{x \in N \,|x \subseteq a\}$ and

$$P(a) = \{x \in N \,|\exists y \in\downarrow a : \uparrow x =\uparrow y \setminus \downarrow a\}, \tag{8.1}$$

where the latter is the set of nodes at which a is *available*. This function will figure in the following definition that adds choices to an event tree:

Definition 4 A **discrete extensive form** (DEF) with player set I is a pair (T,C), where $T = (W,N)$ is an event tree (Definition 1) and $C = (C_i)_{i\in I}$ consists of collections C_i (the set of player i's choices) of non-empty unions of nodes for all $i \in I$, such that
(DEF1) if $P(c) \cap P\left(c'\right) \neq \emptyset$ and $c \neq c'$, then $P(c) = P\left(c'\right)$ and $c \cap c' = \emptyset$, for all $c, c' \in C_i$ and all $i \in I$, and
(DEF2) $p^{-1}(x) = \left\{x \cap \left(\cap_{i\in J(x)} c_i\right) \,\middle|(c_i)_{i\in J(x)} \in A(x)\right\}$, for all $x \in X$,
where $A(x) = \times_{i\in J(x)} A_i(x)$, $A_i(x) = \{c \in C_i \,|x \in P(c)\}$ are the choices available to $i \in I$ at $x \in X$, and $J(x) = \{i \in I \,|A_i(x) \neq \emptyset\}$ is the set of decision makers at x, which is required to be non-empty for all $x \in X$.

The new insight in this is that only *choices* are needed to formalize an extensive form, because information sets can be reconstructed from choices (but not vice versa). Unlike the "actions" in Definition 3, choices keep track of their histories, that is, "turn left" after two distinct histories shows up as two distinct choices. By this token (formally because choices are sets of plays) a player's choices pin down the information sets at which they are available.

Condition (DEF1) ensures that, once at an information set $P(c)$ with $c \in C_i$, player i cannot infer from the available menu of choices at which node in her information set she is called upon to choose. Condition (DEF2) states that the combined decisions of all players choosing at move $x \in X$ give rise to an immediate successor of x. This in fact incorporates a second innovation. For, Definition 4 allows several players to decide at the same move, thus formalizing truly simultaneous decisions. This has not been the case in earlier definitions.

Another noteworthy property of a DEF is that, because choices are sets of plays or outcomes, it is impossible for a play to pass through an information set more than once (Proposition 13 of AR1b). This condition had to be added as part of the definition of an extensive form by Kuhn (1953, p. 219) and his followers. Here it follows naturally from the set-up.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Since Definition 4 is based on discrete game trees, it is not the most general definition of an extensive form in the literature. Yet, it is the specialization of the most general definition (viz. Definition 3.1 of AR2) to the case of discreteness. The latter in fact encompasses even games in continuous time, hence it allows for plays of uncountable length.

2.2.2 Graph-based extensive forms

Kuhn's (1953) seminal definition employed a strict finiteness assumption, as did Selten (1975) and Kreps and Wilson's (1982). Though well justified back then, from today's viewpoint this appears unwarranted. Because of that, a modified definition will be given here. It follows Selten (1975) as closely as possible, save for the cardinality assumptions.

As for notation, given a graph tree (N,R) as in Definition 2, let $X = \{x \in N \,|\, \exists y \in N : (y,x) \in R\}$ denote the set of moves. For each node $x \in N$ denote its immediate predecessor by $\tilde{p}(x) \in X$. This gives a function $\tilde{p} : N \to X$, because $\tilde{p}(x) = \min \uparrow x \setminus \{x\}$ by condition (ST1). Let $\tilde{p}^{-1}(x) = \{y \in N \,|\, \tilde{p}(y) = x\}$ denote the immediate successors, for all moves $x \in X$. Recall that $\geq$ denotes the union of the transitive closure of R with equality.

Definition 5 A **graph-based extensive form** with player set I is a triplet $(\mathfrak{T}, \mathfrak{X}, (D, C))$, where

- $\mathfrak{T} = (N,R)$ is a graph tree as in Definition 2;
- $\mathfrak{X} = (X_i)_{i \in I}$ is a partition of the set X of moves into cells X_i, where the moves $x \in X_i$ are player i's *decision points*;[10]
- the pair $(D,C) = ((D_i, C_i))_{i \in I}$ consists, for each player $i \in I$, of one partition D_i of the set X_i of player i's decision points into *information sets* $d \in D_i$ together with a collection $C_i = (C_{id})_{d \in D_i}$ of partitions C_{id} of the sets $\{(y,x) \in R \,|\, x \in d\}$ of edges incident at (moves in) $d \in D_i$, one *choice partition* for each information set $d \in D_i$, such that:
  (KS1) $d \neq \emptyset$ for all $d \in D_i$ with $i \neq 0$,
  (KS2) if $x, y \in d \in D_i$ and $x \geq y$ then $x = y$, and
  (KS3) for all information sets $d \in D_i$, each move $x \in d$, and all *choices* $c \in C_{id}$ there is a unique $y \in \tilde{p}^{-1}(x)$ such that

$$c \cap \{(z,x) \in R \,|\, \tilde{p}(z) = x\} = \{(y,x)\}. \tag{8.2}$$

In this definition choices are sets of edges emanating from an information set, rather than sets of plays. It is, therefore, not surprising that condition (KS2) has to be added, demanding that every play intersects an information set at most once. The significance of this condition will be discussed in the next subsection. Condition (KS3) is needed, because information sets are treated as primitives, rather than as a derived object as in a DEF. If an information set $d \in D_i$ contained one move $x \in d$ with a continuum of immediate successors and another move $y \in d$ with only two immediate successors, (8.2) is bound to fail. Therefore, the consistency condition (KS3) is needed. In the finite case (KS3) would say that the number of edges emanating from a move in an information set $d \in D_i$ is constant across all moves in d, and that every choice $c \in C_{id}$ contains exactly one edge $(y,x) \in R$ emanating from x for

[10] A *partition* of a non-empty set X is a collection of pairwise disjoint subsets whose union is all of X.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

every move $x \in d$. Hence, the successor reached, after the player owning the move has made her decision, is unique.

Because $\mathfrak{X} = (X_i)_{i \in I}$ is a partition, at each move $x \in X$ there is exactly one player who decides (possibly chance). Hence, truly simultaneous decisions (as in a DEF) are excluded. Yet, there may well be informationally simultaneous decisions, if for all $x \in d \in D_i$ all $y \in \tilde{p}^{-1}(x)$ belongs to the same information set $d' \in D_j$ of player j (where $j \in I$ may or may not be distinct from $i \in I$). That is, even though player i has made her choice when j gets to choose, j cannot distinguish what i has done, nor can she distinguish between plays that i could not distinguish.

Whether or not this informational simultaneity is sufficient for modeling simultaneous decisions is a matter of the underlying decision theory. If, in particular, one adopts the expected utility axioms (see, e.g., von Neumann and Morgenstern, 1944, p. 26, or Fishburn 1988, p. 10), then any compound lottery is indifferent to its reduced lottery by the independence axiom. In such a case it does not matter for decision makers who moves first; all that matters is informational simultaneity. Under decision theories that are characterized by an independence axiom, therefore, this model of "cascading information sets" is enough. Yet, imagine how large the tree would have to be to capture simultaneous decisions of infinitely many players.

The absence of truly simultaneous decisions entails a loss of generality for graph-based extensive forms in comparison with DEFs. On the other hand, a DEF also allows for cascading information sets, since it does not insist on modeling simultaneous decisions as taking place at one node. If one assumes that simultaneous decisions are represented by cascading information sets, graph-based extensive forms and DEFs are equivalent in the following sense:

Proposition 2 *(a) If* $(\mathfrak{T}, \mathfrak{X}, (D, C))$ *is a graph-based extensive form with player set I, then* $\left(T, C'\right)$ *is a DEF with player set I whenever*

$$T = \left(W, \{W(x) | x \in N\} \cup \{\{w\}\}_{w \in W}\right),$$

where W is the set of plays of $\mathfrak{T} = (N, R)$*, and* $C' = \left(C'_i\right)_{i \in I}$*, where,* $\forall i \in I$*,*

$$C'_i = \{\{w \in W \,| \exists y \in w : (y, \tilde{p}(y)) \in c, \, \tilde{p}(y) \in d\} \,| c \in C_{id}, \, d \in D_i\}.$$

(b) If (T, C) *is a DEF with player set I such that* $J(x) = \{i\}$ *for some* $i \in I$ *for all* $x \in X$*, then* $\left(\mathfrak{T}, \mathfrak{X}, \left(D, C'\right)\right)$ *is a graph-based extensive form with player set I whenever* $\mathfrak{T} = (F(N), R)$*, where* $F(N)$ *is the set of finite nodes of* $T = (N, \supseteq)$ *and* $R = \{(x, y) \in F(N) \times F(N) \,| p(x) = y\}$*,* $\mathfrak{X} = (X_i)_{i \in I}$ *is the partition of X given by* $x \in X_i \Leftrightarrow J(x) = \{i\}$*,* $D = (D_i)_{i \in I}$ *is the collection of partitions* D_i *of* X_i *given by* $x \in d \in D_i \Leftrightarrow x \in P(c)$ *for some* $c \in C_i$*, and* $C' = \left(C'_i\right)_{i \in I}$ *consists of collections of partitions* $C'_i = \left(C'_{id}\right)_{d \in D_i}$ *of sets of edges in R, one partition* C'_{id} *for each* $d \in D_i$*, given by* $(y, x) \in c' \in C'_{id} \Leftrightarrow y = x \cap c$ *for some* $c \in C_i$ *with* $x \in P(c)$*, for all* $i \in I$*.*

Proof (a) By Proposition 1(a) it is enough to demonstrate (GT1) and (GT2) for the construction in the statement. (GT2) follows directly from the fact that $\mathfrak{X} = (X_i)_{i \in I}$ partitions X and that an edge is a pair of nodes, where the second is the immediate predecessor of the first. To see (GT1), let $c'_1, c'_2 \in C'_i$ be such that $P\left(c'_1\right) \cap P\left(c'_2\right) \neq \emptyset$, where P is derived from T. Then

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

there are $d \in D_i$ and $c_1, c_2 \in C_{id}$ such that $c_1' = \{w \in W \mid \exists y \in w : (y, \tilde{p}(y)) \in c_1, \tilde{p}(y) \in d\}$ and $c_2' = \{w \in W \mid \exists y \in w : (y, \tilde{p}(y)) \in c_2, \tilde{p}(y) \in d\}$. If $c_1 \neq c_2$, then also $c_1' \cap c_2' = \emptyset$, because C_{id} partitions the edges emanating from d. Likewise, since a partition is exhaustive, $P(c_1') = P(c_2')$, which completes the verification of (GT1).

(b) By Proposition 1(b) and the hypothesis that each move belongs to precisely one player it is enough to verify (KS1)–(KS3). The first is obvious, and (KS2) follows from Proposition 13 of AR1b. Finally, (KS3) follows from (GT2) specialized to the case $X_i \cap X_j \neq \emptyset \Rightarrow i = j$ for all $i, j \in I$. □

By this proposition Definition 5 is as good as a DEF. In particular, plays may be of (countably) infinite length and decision makers may well choose from continua. Yet, because the relation R in the definition of a graph tree (Definition 2) relies on the identification of immediate predecessors, graph-based extensive forms do not admit a generalization to continuous-time problems.

2.2.3 Sequence-based extensive forms

The third variant concerns sequence trees. In those, the empty history serves as the root where all plays begin. The other ingredients are similar to their analogues in the other two definitions.

Definition 6 A **sequence-based extensive form** with player set I is a triplet $(\mathcal{T}, \pi, D)$, where:

- $\mathcal{T} = (A, H)$ is a sequence tree as in Definition 3;
- $\pi : X \to I$ is a function that assigns to each non-terminal history one member of the player set I, i.e., $\pi(x) \in I$ is the player (possibly chance, $0 \in I$) who takes an action at $x \in X = H \setminus E$; and
- $D = (D_i)_{i \in I}$ is a family of partitions, one partition D_i for each of the sets $X_i \equiv \pi^{-1}(i) = \{x \in X \mid \pi(x) = i\}$ of player i's moves, such that

$$\text{if } x, y \in d \in D_i \text{ then } A(x) = A(y) \tag{8.3}$$

  for all $d \in D_i$ and all $i \in I$, where $A(x) = \{a \in A \mid (x, a) \in H\}$.

A cell d in the *information partition* D_i of player $i \in I$ is an *information set* for player i. The histories contained in some $d \in D_i$ are meant to be indistinguishable to player i, when she is called upon to decide after some history $x \in d$. Therefore, condition (8.3) is needed to ensure that player i cannot infer from the available menu of actions $A(x)$ at which particular $x \in d$ she is called upon to choose. Because π is a function defined on all of $X = H \setminus W$, every non-terminal history is assigned to exactly one player. Hence, once again there are no truly simultaneous decisions.

Strictly speaking Definition 6 is not proper, because the tree and the assignment $D = (D_i)_{i \in I}$ of information sets to players interact. In particular, condition (8.3) restricts what the tree can be when D is given, or restricts what information sets $d \in D_i$ can be when the tree is given. Osborne and Rubinstein (1994, p. 200) avoid this glitch by merging Definitions 3 and 6 into a single one.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Choices do not show up explicitly in Definition 6. They may be reconstructed, though, by pairing non-terminal histories $x \in H$ with actions $a \in A$ that are available at x in the sense that $(x, a) \in H$ and taking the union over all decision points $x \in d$ in a particular information set $d \in D_i$ for fixed $a \in A$. That this is possible is guaranteed by (8.3). Hence, in a sense a sequence-based extensive form is a mirror image of a DEF. While in the latter choices are primitives, in the former information sets are primitives.

A peculiarity of Definition 6 is that it admits the possibility that an information set $d \in D_i$ for some player $i \in I$ contains two histories $x, y \in d$ with $y = \left(x, a^1, \ldots, a^k\right)$ for some sequence $\left(a^1, \ldots, a^k\right)$ of actions. Hence, what (KS2) rules out for the graph-based case, and is automatically excluded for a DEF, is allowed here. That is, a path may cross the same information set more than once. This has been dubbed *absent-mindedness* and, ironically, spurred a whole literature.[11] Yet, it is inconsistent with any known decision theory. For, when a player decides at one of her information sets, the histories contained in it figure as "states"; and in all our decision theories states need to be *independent* of what the decision maker does – after all "states" are supposedly chosen by nature or by an opponent. With absent-mindedness they are not.

Example 1 The well-known "drunken driver" example (Piccione and Rubinstein, 1997) illustrates this point. This is a single-player game, $I = \{1\}$, with binary action set $A = \{0, 1\}$, where 0 stands for "continue" and 1 for "exit." The extensive form is obtained by specifying the tree by $H = \{\emptyset, (1), (0), (0, 1), (0, 0)\}$, assigning all non-terminal histories to the same player, $\pi(\emptyset) = \pi((0)) = 1$, and the information partition $D_1 = \{\{\emptyset, (0)\}\}$ with a single cell. The interpretation is that a drunken driver wishes to go home and to do so has to take the second exit from the freeway. Since she is absent-minded, she knows that, once at the second exit, she will have forgotten that she passed the first.

In her single information set the two histories $\emptyset \in H$ and $(0) \in H$ should be like "states." But by taking the action $1 \in A$ the driver precludes the second state, $(0) \in d$, from materializing. This implies that no pure strategy can take our driver home.

For, suppose for an information set $d \in D_i$ of some player $i \in I$ there are $x, y \in d$ such that $y = \left(x, a^1, \ldots, a^k\right)$ for some $k = 1, 2, \ldots$ and $A(x)$ contains at least two elements. Then by (8.3) there is $(y, a) \in H$ such that $a \in A \setminus \left\{a^1\right\}$. Let $h \in E$ be a terminal history with $h = (y, a, \ldots)$. Every pure strategy of player i that picks $a^1 \in A(x)$ at $x \in d$ must also pick $a^1 \in A(y)$ at $y \in d$ and, therefore, cannot reach $h \in E$. On the other hand, every pure strategy of i that does not choose $a^1 \in A(x)$ at $x \in d$ cannot reach $y \in d$ and, therefore, also not $h \in E$. Hence, h is unreachable by any pure strategy. $\Diamond$

Other than that, sequence-based extensive forms are as general as the other two, even though an equivalence result like Proposition 2 cannot be had, due to trivial structures in the tree and the potential presence of absent-mindedness. With appropriate restriction, however, Kline and Luckraz (2016) establish an equivalence between sequence-based extensive forms and graph-based ones. On the other hand, because a sequence tree is bound to have at most countable depth (maximal length of a play), it does not admit a generalization to continuous time.

[11] Among those papers, Gilboa (1997) points out that the phenomenon of absent-mindedness also has a representation that satisfies (KS2).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Even though sequence-based extensive forms are fairly general, this does not mean that they are easy to work with. On the contrary, the implicit product construction may induce "spurious" dimensions in the outcome space, as the following example illustrates:

Example 2 Consider a two-player perfect information game, where at the root player 1 picks either a pair $(a, b) \in [0, 1]^2$ with $a < 1$ or sets $a = 1$ and gives the move to player 2. In the former case ($a < 1$) the game ends. In the latter case ($a = 1$) player 2 is called upon to choose $b \in [0, 1]$ and, once b has been selected, the game ends. With an event tree the set of plays would be $W = [0, 1]^2$ (hence two-dimensional) and the nodes would simply be $N = \{W, \{1\} \times [0, 1], (\{w\})_{w \in W}\}$.

In the sequence approach, on the other hand, action sets for each "stage" need to be constructed. For player 1, who moves at the root, the obvious choice is the action set $A_1 = [0, 1]^2$, as she may (but need not) choose both a and b. For player 2 the action set is clearly $A_2 = [0, 1]$. Since the game has only two "stages," the appropriate product for the outcome space is hence $A = A_1 \times A_2 = [0, 1]^3$. That is, by taking the product of action sets an extra dimension has crept in through the back door! To retrieve the set of plays in this large set A, two classes of histories need to be considered. The first are those where player 1 has fixed both a and b (with $a < 1$) and player 2 does not get to choose, $H_1 = \{((a, b_1), b_2) \,|b_1 = b_2, a < 1\}$. The fact that player 2 is not called upon is incorporated in H_1 by the restriction that $b_2 = b_1$. The second class includes the plays where player 1 has set $a = 1$ and player 2 chooses $b \in [0, 1]$, viz. $H_2 = \{((a, 0), b) \,|a = 1, b \in [0, 1]\}$. The coordinate 0 in $(a, 0)$ is merely an arbitrary marker indicating that player 1 does not choose b. Any other marker would also do, but it would be incorrect to write (a, b), because if $a = 1$, player 2's choice of b is unconstrained by 1's decision. Ultimately, the set of plays in the sequence approach is the union $H_1 \cup H_2$ as a subset of the three-dimensional cube, $H_1 \cup H_2 \subset A = [0, 1]^3$. $\Diamond$

That the sequence approach may blow up the dimension of the space can lead to problems. For instance, traditional existence theorems for subgame perfect equilibria in perfect information games (e.g., Harris, 1985) may not be applicable, even if the game does have a subgame perfect equilibrium, purely because the set H of histories may not be closed as a subset of the product space A (see Example 7 in Alós-Ferrer and Ritzberger, 2016, or Example 3 in Alós-Ferrer and Ritzberger, 2017a) or payoffs may not be continuous in the product topology (see Example 2 in Alós-Ferrer and Ritzberger, 2017b).

2.3 Games

Given an extensive form, in any of the representations above, the step to obtaining a game involves two specifications. First, adding preferences for all the players and, second, if chance (player $0 \in I$) is present (i.e., $X_0 \neq \emptyset$), pinning down what it does. The first is achieved by specifying a profile $\precsim = (\precsim_i)_{i \in I}$ of reflexive, transitive, and complete binary relations $\precsim_i$ on the set W of plays, one *preference relation* for each player $i \in I$. Typically it is assumed that preferences are representable by a utility function $u = (u_i)_{i \in I} : W \to \mathbb{R}^{|I|}$. With that comes an extension of preferences to uncertain prospects on W that is needed to cope with chances moves (see below) and potentially randomized strategies by other players. This extension is almost invariably *expected utility* (resp. its three axioms) (see von Neumann and Morgenstern, 1944, p. 26, or Fishburn, 1988, p. 10).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

The second ingredient is easy in the finite case. It consists of specifying probability distributions over player 0's choices for each of her information sets. In fact, it may be assumed that all of the information sets of chance are singletons. Then, in the finite case, it is enough to specify a positive vector, whose coordinates sum to one, for each of 0's decision points. Such a vector is then seen as the vector of conditional probabilities of the choices of chance given that the decision point, where these choices are available, has been reached.

In the general case this is not so straightforward. The reason is that the cardinality of the set of choices available at any one of the decision points of chance may be too large for a σ-additive probability measure to apply. In that case it is wiser to proceed as follows (see Aumann, 1961, 1963, 1964, and Remark 1 below). Let $(\Omega, \Sigma, \lambda)$ be an extraneous standard probability space,[12] where Ω is the sample space, Σ a σ-algebra on Ω, and λ a probability measure, e.g., Ω the unit interval $[0, 1]$, Σ the Borel σ-algebra, and λ the Lebesgue measure – think of that as a roulette wheel. Endow the set C_0 of 0's choices with a σ-algebra $\mathscr{C}_0$. Assume again that all of 0's information sets are singletons, hence, $C_0 \subseteq N$. Then, a specification for chance is a function $f : X_0 \times \Omega \to C_0$ such that $f(x, \omega) \in p^{-1}(x)$ for all $(x, \omega) \in X_0 \times \Omega$ and the evaluation function $f_x : \Omega \to C_0$, defined by $f_x(\omega) = f(x, \omega)$ for all $(x, \omega) \in X_0 \times \Omega$, is $(\Sigma, \mathscr{C}_0)$-measurable for all $x \in X_0$.[13]

This construction replaces the probability distribution from the finite case with a random variable. While this finesses problems with the cardinality of nature's choices, its drawback is that not all such choices may have positive probability. For instance, if the sample space is $\Omega = \{0, 1\}$, i.e., a coin flip, and for some $x \in X_0$ the set $p^{-1}(x)$ of successors has the cardinality of the continuum, then for each f only two points from the continuum can be chosen. Since probability theory cannot go beyond spaces that resemble the unit interval, this is unavoidable. On the other hand, in applications [0, 1] will most likely suffice to capture chance moves. Hence, for practical purposes the random variable approach sketched above will do.

3 STRATEGIES AND NORMAL FORMS

Having explained the details of formalizing an extensive form, from this point onwards the notions incorporated in Definitions 1 and 4 will be employed. It has been shown that this does not make a difference, and using event trees and DEFs makes the exposition more compact.

3.1 Strategies

Putting things together, the primitives of an extensive form representation are the tree and choices or information sets. The basic assumption of non-cooperative game theory is *complete information*, meaning that this representation and all players' preferences are common knowledge among the players.[14] In particular, the extensive form is common knowledge among the players. Therefore, each player can reason through all possible plays of the game

---

[12] The sample space Ω is *standard* if it is either finite or countable with the discrete σ-algebra or isomorphic to the unit interval.

[13] The random variable $f_x : \Omega \to C_0$ is (Σ, Σ_0)-measurable if $f_x^{-1}(\vartheta) \in \Sigma$ for all $\vartheta \in \Sigma_0$.

[14] The literature is somewhat unfortunately haunted by terms like "incomplete information game." This is not to be taken literally. It refers to the complete information game that is obtained after all uncertainty that players may entertain about the rules or the preferences has been encoded into "types" – the Harsanyi transformation.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

and come up with a plan for all contingencies. Such a complete plan for all possible events is known as a pure strategy.

In a DEF a *pure strategy* for player $i \in I$ is a function $s_i : X_i \to C_i$ such that

$$s_i^{-1}(c) = P(c) \text{ for all } c \in s_i(X_i), \tag{8.4}$$

where $s_i(X_i) = \{s_i(x) \in C_i \,|\, x \in X_i\}$ and $s_i^{-1}(c) = \{x \in X_i \,|\, s_i(x) = c\}$. That $s_i^{-1}(c) \subseteq P(c)$ says that choice $c \in C_i$ can only be taken where it is available in the sense that $x \in P(c)$. That $s_i^{-1}(c) \supseteq P(c)$ says that at all moves, where i has choice $c \in C_i$ available, she must take the same choice. That is, a pure strategy is measurable with respect to the player's information partition.

A pure strategy is the direct generalization to dynamic interactive decision theory of Savage's (1954) notion of an "act." In a static decision problem under uncertainty an *act* is a function from states to consequences. At an information set the moves contained in it perform the role of states, and choices the role of consequences. Hence, a strategy maps decision points into choices. But it does so under the constraint that it cannot use more information than what the player has – hence measurability. A pure strategy is therefore a derived object, not a primitive.[15]

For each player $i \in I$ denote by S_i the set of pure strategies, i.e., the set of all functions satisfying (8.4). The product $S = \times_{i \in I} S_i$ is the set of all *pure strategy combinations*. A pure strategy combination $s = (s_i)_{i \in I} \in S$ *induces* the play $w \in W$ if

$$w \in \bigcap \{s_i(x) \,|\, w \in x \in X,\ i \in J(x)\}, \tag{8.5}$$

that is, if w is a fixed point of the correspondence (from W to W) defined by the right hand side of (8.5).

With the notion of a pure strategy at hand the somewhat cryptic remarks about continuous-time problems can now be made more precise. This is done in the following example (viz. Example 10 of AR2, p. 229):

Example 3 Let W be the set of all functions $f : \mathbb{R}_+ \to A$, where A is some non-empty set, say, $A = \{0, 1\}$. To define the set of nodes let $N = \{x_t(g) \,|\, g \in W,\ t \in \mathbb{R}_+\}$, where $x_t(g) = \{f \in W \,|\, f(\tau) = g(\tau)\ \forall \tau \in [0, t)\}$, for any $g \in W$ and $t \in \mathbb{R}_+$. Intuitively, at each point in time t a decision $a_t \in A$ is taken. The "history" of all decisions in the past up to, but exclusive of time t is a function $f : [0, t) \to A$, i.e., $f(\tau) = a_\tau$ for all $\tau \in [0, t)$. A node at time t is the set of all functions that coincide with f on $[0, t)$, all possibilities still open for their values thereafter. It can be shown that this is a game tree, in its most general incarnation, albeit not a discrete one (see Example 14 of AR1b). To turn it into an extensive form, define first[16]

$$\gamma(x, w) = \bigcup \{z \in N \,|\, w \in z \in \downarrow x \setminus \{x\}\} \tag{8.6}$$

[15] One of his students attributes the following stark expression of this principle to Pierpaolo Battigalli: "Strategies cannot be chosen."

[16] The function γ can be shown to provide choices under perfect information in any (general) game tree; see Theorem 1 of AR2, p. 226.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

for each $x \in X = N \setminus \{\{w\}\}_{w \in W}$ and all $w \in x$ and, second, define choices available at $x_t(f) \in X$ by $c_t(f,a) = \{g \in x_t(f) \,|\, g(t) = a\} = \gamma(x_t(f), g) \subset x_t(g)$ for any $g \in x_t(f)$ with $g(t) = a \in A$. Let C_1 denote the set of all such choices. Now assume that there is a single player, who has perfect information. With a single player and perfect information a pure strategy is simply a function $s_1 : X \to C_1$ such that $s_1^{-1}(\gamma(x_t(f), g)) = \{x_t(f)\}$ for all $x_t(f) \in X$.

One such strategy is $s_1(W) = c_0(f, 1)$, $s_1(x_t(f)) = c_t(f, 0)$ if $f(\tau) = 1$ for all $\tau < t$, and $s_1(x_t(f)) = c_t(f, 1)$ otherwise, for any $t > 0$ and $f \in W$. Intuitively, this strategy starts with choosing $1 \in A$ at time $t = 0$ and then, for any positive $t \in \mathbb{R}_{++}$, chooses $0 \in A$, provided that $0 \in A$ has not been chosen before, in which case it continues with $1 \in A$. Even though this is a perfectly valid strategy, it induces no outcome (play).

For, the constant function $\mathbf{1}$ (i.e., $\mathbf{1}(t) = 1$ for all $t \geq 0$) cannot be a fixed point as in (8.5), because $\mathbf{1}(\tau) = 1$ for all $\tau < t$ for any $t > 0$, so that by the construction of s_1 it would follow that $\mathbf{1}(t) = 0$, a contradiction. Suppose there is a fixed point $f \in W$ as in (8.5). Then $f(0) = 1$, but, since $f \neq \mathbf{1}$, there exists $t > 0$ such that $f(t) = 0$. Therefore, the set of real numbers $\{t \geq 0 \,|\, f(t) = 0\}$ is non-empty and bounded below by $0 \in \mathbb{R}_+$. By the supremum axiom this set has an infimum t^*. If $t^* > 0$, consider $t' = t^*/2 > 0$. Then $f(t') = 1$, but also $f(\tau) = 1$ for all $\tau < t'$. By the definition of s_1 it would follow that $f(t') = 0$, a contradiction. Therefore, $t^* = 0$. But then consider any $t > 0$. By the definition of an infimum there exists $\tau \in (0, t)$ such that $f(\tau) = 0$. By the definition of s_1 it follows that $f(t) = 1$. Since $t > 0$ was arbitrary, it follows that f must be identically 1, i.e., $f = \mathbf{1}$, a contradiction. $\diamondsuit$

The problem illustrated in this example defies the whole purpose of game theory. For instance, recall that a Nash equilibrium is defined as a strategy combination, namely one from which no player has an incentive to deviate. If there are strategy profiles that induce no outcome, what sense does such a notion make? How is a deviation to be evaluated, if it induces no outcome at all?

This is not a matter of existence of equilibrium, but a conceptual problem. The very idea of a non-cooperative game, i.e., a game with complete rules, is a thought-experiment in which all decisions are left to the players. But players can only evaluate objects in the domain of their preferences: plays or outcomes. If a strategy results in no play at all, players cannot decide and the thought-experiment has miserably failed. This also applies to the next example (viz. Example 12 of AR2, p. 239):

Example 4 Take the same example as before, the single-player differential game. Now consider the strategy $s_1 \in S_1$ defined by $s_1(x_t(f)) = c_t(f, 1)$ if $f \in x_t(\mathbf{1})$ and $s_1(x_t(f)) = c_t(f, 0)$ otherwise, for all $x_t(f) \in X$, where $\mathbf{1}$ again denotes the constant function, $\mathbf{1}(t) = 1\ \forall t$. Intuitively, this strategy begins at $t = 0$ with $1 \in A$ and then, for any $t > 0$, sticks to $1 \in A$ except when $0 \in A$ has occurred in the past, in which case it chooses $0 \in A$. Once again this is a perfectly valid strategy.

This strategy induces quite a number of plays. In fact, for every $r > 0$ the function $f_r \in W$ such that $f_r(t) = 1$ for all $t \in [0, r]$ and $f_r(t) = 0$ for all $t > r$ is a fixed point as in (8.5), because

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

$$f_r \in \bigcap_{g \in x_t(f_r)} s_1\left(x_t\left(g\right)\right) = \left[\bigcap_{t \leq r} c_t\left(\mathbf{1}, 1\right)\right] \cap \left[\bigcap_{r<t} \left\{h \in x_t\left(f_r\right) \left| h\left(t\right) = 0\right.\right\}\right]$$
$$= \left[\bigcap_{t \leq r} c_t\left(\mathbf{1}, 1\right)\right] \cap \left[\bigcap_{r<t} \left\{h \in x_t\left(f_r\right) \left| h\left(\tau\right) = 0 \, \forall \tau \in \left(r, t\right]\right.\right\}\right] = \{f_r\}.$$

In other words, this strategy s_1 induces a whole continuum of plays, as f_r is a fixed point for every $r > 0$. ◊

These two examples show that continuous time is inappropriate for the purposes of game theory. The underlying insight goes beyond examples, though. AR2 (Theorem 6, p. 242) provides a characterization theorem for the class of game trees that are such that (a) every strategy profile induces an outcome, and (b) if a strategy profile induces an outcome, this outcome is unique. The class thus characterized is slightly more general than DEFs,[17] but it definitely excludes continuous-time problems.

There is a third property that one may ask for: (c) every play is induced by some strategy combination. This holds true in a DEF (by Theorem 4 of AR2, p. 238). But it may fail in the absence of a condition like (KS2), that every play intersects an information set at most once. Hence, it may fail in sequence-based extensive forms (Definition 6), as illustrated in Example 1.

3.1.1 Randomized strategies

Strategies are functions, so the set of all pure strategies of a player is a function space. Therefore, in general this space is huge. Take, for instance, a two-player game with perfect information,[18] where player 1 first chooses from the unit interval, player 2 sees 1's choice and then also chooses from the unit interval. The set of player 2's pure strategies is the the set of all functions from the unit interval to the unit interval, $[0, 1]^{[0,1]}$. With spaces of that size there is no guarantee that probability distributions on strategies do result in probability distributions on outcomes (plays).

Example 5 There are two players, 1 and 2, engaged in ultimatum bargaining. Player 1 proposes a split of a unit surplus, which is any number from the interval $[0, 1]$. The set $[0, 1]$ hence coincides with the set S_1 of pure strategies for player 1. Player 2 observes the proposal and responds by either accepting (1) or rejecting (0) the proposed split. Thus, the set of pure strategies of player 2 is the set S_2 of all functions of the form $s_2 : [0, 1] \to \{0, 1\}$. The set of plays is $W = [0, 1] \times \{0, 1\}$. A minimal requirement on a σ-algebra $\mathscr{W}$ on W is that singletons and sets of the form $\{r\} \times \{0, 1\}$ for $r \in [0, 1]$ are measurable. This is fulfilled if one takes, for instance, the product of the Borel σ-algebra on $[0, 1]$ and the discrete σ-algebra on $\{0, 1\}$. The outcome function ϕ is explicitly given by $\phi\left(s_1, s_2\right) = \left(s_1, s_2\left(s_1\right)\right)$.

Strikingly, problems already arise if player 1 randomizes uniformly and player 2 uses a pure strategy. Take a non-Borel set A of $[0, 1]$ and consider the indicator function $\mathbf{1}_A \in S_2$. This is

[17] Specifically, only up-discreteness and regularity ($\uparrow x \setminus \{x\}$ has an infimum for all $x \in N$) are necessary and sufficient, but not down-discreteness.

[18] Formally, perfect information for a DEF is defined by $c \in N$ for all choices $c \in C_i$ and all players $i \in I$, i.e., by all choices being nodes.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

a pure strategy for player 2. Thus, for a pure decision theorist, nothing can rule it out. Now suppose player 1 randomizes uniformly over $S_1 = [0,1]$. What is the induced distribution over outcomes? Clearly, the set $[0,1] \times \{1\}$ should be measurable for any reasonable model of the game. But $\phi_2^{-1}([0,1] \times \{1\}) = A$, which is *not* measurable by construction. Thus the uniform randomization of player 1 (which is a well-defined random variable) does not induce a distribution over the set of outcomes. $\diamondsuit$

It may be tempting in this example to replace the Borel σ-algebra by the discrete σ-algebra. But, of course, it is known that on the class of all subsets of the unit interval there exists no probability measure that assigns the value zero to all singletons (see Birkhoff, 1961, p. 187; Billingsley, 1986, p. 41).

Remark 1 The problem is in fact deeper and concerns the following general situation. Suppose you are given a probability space (Ω, Σ, μ) and a measurable space $(W, \mathscr{W})$, say, the first capturing chance or other players and the second outcomes. Your own strategy space S consists of all $(\Sigma, \mathscr{W})$-measurable functions $s : \Omega \to W$. If you wish to randomize your strategy choice, you will be interested in the subsets $F \subseteq S$ such that there is a σ-algebra $\mathscr{S}_F$ on F for which the "outcome function" $\varrho_F : F \times \Omega \to W$, defined by $\varrho_F(s, \omega) = s(\omega)$ for all $(s, \omega) \in F \times \Omega$, is jointly measurable. For, if you randomize over F according to the probability measure ν (on $\mathscr{S}_F$) and $B \in \mathscr{W}$ can be assigned a probability, then $\Pr(s(\omega) \in B) = (\mu \times \nu)\left(\varrho_F^{-1}(B)\right)$, hence $\varrho_F^{-1}(B)$ must belong to the product σ-algebra on $F \times \Omega$.

Aumann (1961) has proved a characterization theorem for such subsets $F \subseteq S$ (under the hypothesis that both (Ω, Σ) and $(W, \mathscr{W})$ have countable generating families). This theorem relates the "sizes" of (Ω, Σ) and $(W, \mathscr{W})$ to how large F can be. If both Ω and W are finite, then randomization can apply to all strategies, that is, $F = S$ can actually hold. On the other hand, if both Ω and W are continua, then the best that can be achieved is a proper subset, $F \subset S$. In particular, if both Ω and W are like unit intervals (with the Borel algebra), it is impossible to randomize over *all* functions from the unit interval to itself.

To bypass this problem, Aumann (1964) proposed using the following random variable approach.[19] Fix a standard probability space $(\Omega, \Sigma, \lambda)$, as with chance moves. For each player $i \in I$ endow the space S_i of pure strategies with a σ-algebra $\mathscr{S}_i$.[20] A *mixed strategy* for player $i \in I$ is a $(\Sigma, \mathscr{S}_i)$-measurable function $\sigma_i : \Omega \to S_i$. Denote by M_i the set of all mixed strategies of player $i \in I$ and by $M = \times_{i \in I} M_i$ the space of all mixed strategy profiles. The interpretation of a mixed strategy of player $i \in I$ is that i picks the set $\vartheta \in \mathscr{S}_i$ of pure strategies with probability $\lambda\left(\sigma_i^{-1}(\vartheta)\right)$. By varying the function σ_i the player chooses this probability. Of course, the same caveat as before applies. If S_i is too large as compared to Ω, only "few" pure strategies can be chosen with positive probability.

19 In another paper (Aumann, 1963) he actually shows that the random variable approach is essentially equivalent to using the largest F satisfying the joint measurability property described in Remark 1.

20 Strictly speaking $\mathscr{S}_i$ is not entirely arbitrary. To make things work, the first object is a σ-algebra $\mathscr{W}$ on the set of plays W that contains at least all nodes $x \in N$. The second is a σ-algebra $\mathscr{S}$ on the product space S such that the function $\phi : S \to W$ (whose existence follows from Theorems 4 and 6 of AR2) is $(\mathscr{S}, \mathscr{W})$-measurable. The σ-algebra $\mathscr{S}_i$ is then the projection σ-algebra on S_i, for each $i \in I$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

A conceptually different type of randomized strategy derives from a piecemeal approach. Suppose each player $i \in I$ is split into "agents," one agent for each of i's information sets. The only coordination device between the agents is the fact that all agents of the same player have the same preference relation on plays. The correlation between decisions at different information sets (of the same player), which is implicit in the notion of a mixed strategy, disappears.

For each player $i \in I$ endow the set C_i of choices with a σ-algebra $\mathscr{C}_i$ and let B_i denote the set of all $(\Sigma, \mathscr{C}_i)$-measurable functions $b : \Omega \to C_i$. A *behavior strategy* for player $i \in I$ is a function $\beta_i : X_i \to B_i$, whose values are denoted $b_{ix} = \beta_i(x) : \Omega \to C_i$ for all $x \in X_i$, such that, for all $x, y \in X_i$, (a) $b_{ix}(\Omega) \subseteq A_i(x)$, (b) if $y \in P(b_{ix}(\omega))$ for some $\omega \in \Omega$, then $b_{iy} = b_{ix}$, and (c) if there is no $c \in C_i$ with $x, y \in P(c)$, then $\lambda\left(b_{ix}^{-1}(\vartheta) \cap b_{iy}^{-1}(\vartheta')\right) = \lambda\left(b_{ix}^{-1}(\vartheta)\right)\lambda\left(b_{iy}^{-1}(\vartheta')\right)$ for all $\vartheta, \vartheta' \in \mathscr{C}_i$.

Condition (a) states that if $c \in b_{ix}(\Omega) = \{b_{ix}(\omega) \in C_i \,|\, \omega \in \Omega\}$, then $c \in A_i(x) = \{c \in C_i \,|\, x \in P(c)\}$, i.e., $x \in P(c)$, for all $x \in X_i$; that is, it ensures that the random variable b_{ix} is supported on choices that are available at $x \in X_i$. Condition (b) demands that the same random variable b_{ix} is assigned to all moves y in the information set that contains x; hence, the behavior strategy ρ_i does not use more information than the player has. Finally, condition (c) imposes independence on decisions at distinct information sets. Denote by $\mathcal{B}_i$ the set of all behavior strategies of player $i \in I$, and by $\mathcal{B} = \times_{i \in I}\mathcal{B}_i$ the space of all behavior strategy profiles.

The interpretation of the probability $\lambda\left(b_{ix}^{-1}(\vartheta)\right)$ is as the *conditional* probability that player i takes a choice in the set $\vartheta \in \mathscr{C}_i$ *given* that move $x \in X_i$ has materialized. Player i decides on this conditional probability by choosing the function (random variable) $b_{ix} \in B_i$. By condition (b) these decisions are perfectly correlated across all moves in the information set that contains x, but independent across different information sets by condition (c). Hence, while mixed strategies pick functions from decision points to choices potentially at random, behavior strategies pick choices at each decision point (again potentially at random), independently across different information sets. Because of the independence inherent in behavior strategies, they are, in general, less powerful than mixed strategies are. To see that, endow the set W of plays with a σ-algebra $\mathscr{W}$ that contains at least all nodes $x \in N$. Alós-Ferrer and Ritzberger (2017a) then prove the following result:

Proposition 3 *For any DEF (T, C): If the behavior strategy profile $\beta \in \mathcal{B}$ induces the probability measure $\mu : \mathscr{W} \to [0, 1]$ on the measurable space $(W, \mathscr{W})$, then there exists a mixed strategy combination $\sigma \in M$ that also induces μ.*

Proof To begin with, define for the behavior strategy profile $\beta \in \mathcal{B}$ and each player $i \in I$ the function $f_i : X_i \times \Omega \to C_i$ by $f_i(x, \omega) = b_{ix}(\omega) = \beta_i(x)(\omega)$ for all $(x, \omega) \in X_i \times \Omega$, and let $f = (f_i)_{i \in I}$ denote the associated profile. Let $\phi : S \to W$ be the surjection that assigns to each pure strategy combination $s \in S$ the play that it induces, as in (8.5). This function exists by Theorems 4 and 6 of AR2. Observe that for each fixed $\omega \in \Omega$ the function $f(\cdot, \omega) : X_i \to C_i$ is a pure strategy combination, i.e., $f(\cdot, \omega) \in S$, for (a) guarantees that $f_i(\cdot, \omega)^{-1}(c) \subseteq P(c)$ and (b) ensures that $f_i(\cdot, \omega)^{-1}(c) \supseteq P(c)$ for all $c \in C_i$ and all $i \in I$. Hence, β induces μ if $\mu(V) = \lambda(\{\omega \in \Omega \,|\, \phi(f(\cdot, \omega)) \in V\})$ for all $V \in \mathscr{W}$, for fixed $\omega \in \Omega$ a mixed strategy

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

profile $\sigma \in M$ is a pure strategy combination by definition, $\sigma(\omega) \in S$. Hence, σ induces μ if $\mu(V) = \lambda(\{\omega \in \Omega \,|\phi(\sigma(\omega)) \in V\})$ for all $V \in \mathscr{W}$.

Given the behavior strategy profile $\beta \in \mathcal{B}$, construct a mixed strategy profile $\sigma \in M$ by setting $\sigma(\omega) = f(\cdot, \omega) \in S$ for each $\omega \in \Omega$. Then by construction, if β induces μ, then $\sigma \in M$ also does, and the statement is verified. □

The proof of Proposition 3 implicitly relies on the fact that every DEF satisfies no-absent-mindedness, i.e., Example 1 is excluded. It can be shown that, at least in the finite case, no-absent-mindedness is in fact equivalent to the statement of Proposition 3 (see Ritzberger, 2002, Theorem 3.2, p. 122).

An analogous result like Proposition 3 with the roles of behavior and mixed strategies reversed does not hold, because of the independence (condition (c)) in the definition of behavior strategies. This can lead to problems in applications, for in applications one often determines optimal choices locally, at each information set separately, and pastes together a solution to the overall game from these local solutions – the essence of dynamic analysis as incorporated in the notion of backwards induction. That is, in practice, equilibria are frequently determined in behavior strategies. But, if mixed strategies are indeed more powerful, then there could be profitable deviations in mixed strategies, even if there are no such deviations in behavior strategies. Hence, in general, an "equilibrium in behavior strategies" may not be an equilibrium at all.

This calls for a condition that puts mixed and behavior strategies on an equal footing. And indeed, such a condition has already been provided by Kuhn (1953) in his seminal paper. This condition will be discussed next.

3.2 Perfect Recall

The condition that renders mixed and behavior strategies equally powerful is known as "perfect recall." Kuhn (1953, p. 213) describes it as "equivalent to the assertion that each player is allowed by the rules of the game to remember everything he knew at previous moves and all of his choices at those moves." Strictly speaking it does a little more. In particular, it also implies that for each player separately there is something like a time axis, formally each player's information sets can be partially ordered (see Ritzberger, 1999). And this is needed, because otherwise notions like "the past" or "memory" don't make sense.

This description reads more like a rationality condition than a property of different types of randomized strategies. But indeed Kuhn's theorem (1953, p. 214) states the following remarkable characterization: mixed and behavior strategies are equivalent *if and only if* the game satisfies perfect recall.[21] By the only-if part of this statement perfect recall is a necessity in a dynamic analysis of games. Hence, the concept will now be introduced formally.

Given a DEF (T, C) and a move $x \in X_i$ of player $i \in I$, say that x is *possible* under $s_i \in S_i$ for player i, denoted $x \in \text{Poss}(s_i)$, if there is $s_{-i} \in \times_{j \neq i} S_j \equiv S_{-i}$ such that $\phi(s_i, s_{-i}) \in x$, where $\phi : S \to W$ denotes the surjection that assigns to each pure strategy combination the play that it induces. Similarly, an information set $P(c)$ for $c \in C_i$ for player $i \in I$ is *relevant*

[21] In the light of Proposition 3 Kuhn's theorem may be rephrased as follows: for every mixed strategy profile there exists a behavior strategy profile that induces the same probability measure on plays (as the mixed strategies do) if and only if the game satisfies perfect recall.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

under $s_i \in S_i$ for player i, denoted $P(c) \in \mathrm{Rel}(s_i)$, if $P(c) \cap \mathrm{Poss}(s_i) \neq \emptyset$. The following definition is directly from Kuhn (1953, p. 213):

Definition 7 A DEF (T, C) satisfies **perfect recall** if $P(c) \in \mathrm{Rel}(s_i)$ implies $P(c) \subseteq \mathrm{Poss}(s_i)$ for all $c \in C_i$, all $s_i \in S_i$, and all $i \in I$.

Clearly, perfect recall could be defined for each player separately by dropping the last quantifier. Furthermore, the definition may also be rewritten in terms of the function ϕ as follows: player i's choices in a DEF satisfy *perfect recall* if, for all $x \in P(c)$,

$$x \cap \phi(s_i, S_{-i}) \neq \emptyset \Rightarrow y \cap \phi(s_i, S_{-i}) \neq \emptyset, \ \forall y \in P(c) \tag{8.7}$$

for all $c \in C_i$ and all $s_i \in S_i$, where $\phi(s_i, S_{-i}) = \{\phi(s_i, s_{-i}) \,|\, s_{-i} \in S_{-i}\}$. This follows from, for all $s_i \in S_i$,

$$\begin{aligned} \mathrm{Poss}(s_i) &= \{x \in X_i \,|\, x \cap \phi(s_i, S_{-i}) \neq \emptyset\} \ \text{ and} \\ \mathrm{Rel}(s_i) &= \{P(c) \,|\, c \in C_i, \exists x \in P(c) : x \cap \phi(s_i, S_{-i}) \neq \emptyset\}. \end{aligned}$$

A drawback of the original definition is that it refers to derived objects, namely pure strategies. Selten (1975) has provided an equivalent definition that uses only primitives: player i's choices in a DEF satisfy perfect recall if, for all $c, c' \in C_i$, that there is $x \in P(c)$ with $x \subseteq c'$ implies that $y \subseteq c'$ for all $y \in P(c)$. That is, if one move x in an information set $P(c)$ can be reached by choosing $c' \in C_i$, then *all* moves y in $P(c)$ can be reached by choosing c'.

Even though the definitions make no reference to finiteness, Kuhn's theorem was originally proved only for the finite case. This was because the general case indeed poses serious technical difficulties. It took over ten years until Aumann (1964), using the random variable approach sketched above, proved a weaker version of Kuhn's theorem for the case of games with perfect recall, an infinite horizon, and action sets that are homeomorphic to the unit interval: given a pure strategy combination of the opponents, for every mixed strategy of player i there is a behavior strategy for that player, which induces the same probability measure on plays as the mixed strategy does.[22] Of course, this states only the sufficiency part of Kuhn's theorem for large games, and it does so only for a given strategy of the opponents. The latter appears acceptable for practical problems. The former, the necessity of perfect recall, was proved another ten years later by Schwarz (1974) under a few mild measurability assumptions.

So, overall, the issue of a general version of Kuhn's theorem is largely but not entirely settled. In any case, it is clear that perfect recall is an indispensable condition for a dynamic analysis of games. In particular, the important technique of backwards induction will not work without it. After all, an "equilibrium in behavior strategies" may not be an equilibrium at all due to profitable deviations in mixed strategies, or there may be equilibria in mixed strategies that cannot be reproduced in behavior strategies (see the remarkable example by Wichardt, 2008).

[22] Technically, Aumann's proof in fact uses the order structure on the unit interval (in paragraph 7), rather than only its measurability structure. Hence, a generalization to separable complete metric action spaces will require the Borel-isomorphism of such spaces with the unit interval.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

3.3 Normal Forms

If for an extensive form every pure strategy combination $s \in S$ induces a unique play $w \in W$, i.e., if there exists a function $\phi : S \to W$, then there is a way to ascertain the outcome of a playing of the game that eliminates the dynamics of the interaction from the picture. Every player picks a strategy, submits it to an umpire, and the umpire executes the strategy combination. This leads to the representation of the game in normal form. Once again, this is now defined without reference to preferences, payoffs, or utilities, on the understanding that preferences need to be added to turn it into a game.

Definition 8 A **normal form** with player set I is a triplet (S, W, ϕ), where $S = \times_{i \in I} S_i$ is the product set of pure strategy combinations, W is a non-empty set of plays or outcomes, and $\phi : S \to W$ is a surjection that associates with every strategy combination $s \in S$ the play $\phi(s) \in W$ that it induces.

Once again, a *normal form game* is obtained from a normal form by adding preferences or utility functions for all players. In a normal form players make their decisions once and for all by choosing a strategy; hence all dynamics is gone. On the other hand, the mathematical structure of a normal form is clearly simpler than that of an extensive form; it is a map on a product set. Most of the basic solution concepts for games are defined in the normal form, e.g., iteratively undominated strategies, rationalizable strategies (Bernheim, 1984; Pearce, 1984), or Nash equilibrium (Nash, 1950, 1951). Only refined notions of Nash equilibrium make use of the dynamics of interaction, e.g., subgame perfect equilibrium (Selten, 1965) or sequential equilibrium (Kreps and Wilson, 1982).

Normal form representations exist for all DEFs by Theorems 4 and 6 of AR2. By Proposition 2 they also exist for every graph-based extensive form, and also for sequence-based extensive forms, as far as those are equivalent to the former.

In the normal form the primitives are now strategies, outcomes, and the relation between these two. Therefore, one is free to specify which strategies are admitted in modeling a game – something that is not true in the extensive form. In particular, it may be desirable to collapse "equivalent" strategies into single representatives.

Two strategies $s_i, s_i' \in S_i$ of player $i \in I$ are said to *agree on* a strategy subset $R_{-i} \subseteq S_{-i} = \times_{j \neq i} S_j$ if $\phi(s_i, s_{-i}) = \phi\left(s_i', s_{-i}\right)$ for all $s_{-i} = (s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n) \in R_{-i}$. They are called *strategically equivalent* if they agree on all of S_{-i}. A normal form in which no two distinct strategies of the same player are strategically equivalent is called a *reduced normal form* (see Dalkey, 1953, p. 222).

If one begins with the strategies derived from an extensive form, the associated reduced normal form is obtained by taking the quotient space with respect to strategic equivalence for each player. This collapses into single representatives all classes of strategies that differ only at contingencies, which cannot arise under these strategies (in the extensive form).

Proposition 4 *For a DEF (T, C) two strategies $s_i, s_i' \in S_i$ of player $i \in I$ agree on the (non-empty) strategy subset $R_{-i} \subseteq S_{-i} = \times_{j \neq i} S_j$ if and only if*

$$s_i(x) \neq s_i'(x) \Rightarrow x \cap \left(\phi(s_i, R_{-i}) \cup \phi\left(s_i', R_{-i}\right)\right) = \emptyset$$

for all $x \in X_i$, , where $\phi(s_i, R_{-i}) = \{\phi(s_i, s_{-i}) \in W \,|s_{-i} \in R_{-i}\}$ for all $s_i \in S_i$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Proof "if:" The hypothesis is equivalent to $x \cap \left(\phi\left(s_i, R_{-i}\right) \cup \phi\left(s_i', R_{-i}\right)\right) \neq \emptyset \Rightarrow s_i(x) = s_i'(x)$ for all $x \in X_i$. Suppose $w \in \phi\left(s_i, R_{-i}\right)$; then $w \in x \in X_i$ implies $s_i(x) = s_i'(x)$ and $w \in s_i(x) = s_i'(x)$ by the construction of ϕ (see the right-hand side of (8.5)). It follows that $w \in \phi\left(s_i', R_{-i}\right)$. Repeating the argument with $w' \in \phi\left(s_i', R_{-i}\right)$ proves that $\phi\left(s_i, R_{-i}\right) = \phi\left(s_i', R_{-i}\right)$.

"Only if": Assume that $\phi\left(s_i, R_{-i}\right) = \phi\left(s_i', R_{-i}\right)$ and let $x \in X_i$ be such that $x \cap \phi\left(s_i, R_{-i}\right) \neq \emptyset$. If $s_i(x) \neq s_i'(x)$, then $s_i(x) \cap s_i'(x) = \emptyset$ by (DEF1) implies $w \notin \phi\left(s_i', R_{-i}\right)$ for any $w \in s_i(x) \cap \phi\left(s_i, R_{-i}\right)$. It follows that $\phi\left(s_i, R_{-i}\right) \neq \phi\left(s_i', R_{-i}\right)$ in contradiction to the assumption. Hence, $s_i(x) = s_i'(x)$ whenever $x \cap \phi\left(s_i, R_{-i}\right) \neq \emptyset$. □

In particular, strategically equivalent strategies take identical choices at all information sets that are reachable with these strategies. Hence, the information suppressed by going to the reduced normal form concerns only events that cannot materialize. And there is a sense in which this appears as inessential information for rational decision making.

4 APPLICATIONS

Even though a well-specified extensive form gives rise to a normal form there is a definite advantage that the former representation has over the latter. This section illustrates this by looking at two applications. The first application concerns the ability to investigate substructures, like subgames or particular information sets. The second concerns the formal virtue of being able to represent simultaneous decisions as truly simultaneous.

4.1 Directed Search

The representation of a game in extensive form allows the analyst to consider the behavior, as implied by an equilibrium, in parts of the tree where certain decisions have already been taken, thereby enabling a better understanding of how equilibrium is supported. A prime example of that is the celebrated equilibrium refinement of subgame perfection (Selten, 1965). Since subgames are games of their own, judging how plausible behavior in a proper subgame is informs the analyst about how sensible the overall equilibrium is.

An example of that point emerges from the literature on directed labor market search (e.g., Montgomery, 1991; Burdett, Shi, and Wright, 2001; Lang, Manove, and Dickens, 2005). The basic model works as follows. There are n risk-neutral firms who wish to fill precisely one vacancy each. The vacancies, if filled, have a gross value of $v_i > 0$ to firm $i = 1, \ldots, n$. On the other side of the market there are m potential workers who seek employment, but who can apply to only one employer. To begin the game, all firms simultaneously post wage offers $y = (y_1, \ldots, y_n) \geq 0$. All these offers are publicly announced, hence become common knowledge among the potential workers. Therefore, the vector y of wage offers corresponds to the root of a subgame (each vector y belongs to a different subgame). In this subgame job seekers have to decide which firm to apply to, that is, each of them chooses precisely one firm.

The simplest case is $n = m = 2$. In the subgame after the two firms have posted their wage offers $y = (y_1, y_2)$ each of the two job seekers may choose either firm 1 or firm 2.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Assuming that facing two applicants a firm randomizes uniformly, the two workers face the following 2×2 game:

$y_1/2$; $y_1/2$	y_1 ; y_2
y_2 ; y_1	$y_2/2$; $y_2/2$

where worker 1 plays rows and worker 2 columns and the upper left entry is 1's payoff and the lower right 2's. If $2y_2 \geq y_1 \geq y_2/2$ and $2y_1 \geq y_2 \geq y_1/2$, the game has two pure strategy equilibria, one where firm 1 and worker 1 and firm 2 and worker 2 are matched, and another where firm 1 and worker 2 and firm 2 and worker 1 are matched. (If $y_2/2 > y_1$ both workers will apply at firm 2, if $y_1/2 < y_2$ both workers apply at firm 1.) There is also a mixed equilibrium where both players mix with probability $(2y_1 - y_2) / (y_1 + y_2)$ for the first row (resp. column).

The latter is in fact the equilibrium on which this literature focuses. The reasons for this focus are not quite clear. An argument sometimes given is that the pure strategy equilibria require coordination on the part of the applicants (e.g., Montgomery, 1991, p. 167) or that it amounts to anonymity (e.g., Lang et al., 2005, p. 1329). But, of course, a mixed equilibrium requires even more coordination and is even more at variance with anonymity than a pure strategy equilibrium. After all, in a mixed strategy equilibrium each player has to compute all possible payoffs of her opponent to keep him precisely indifferent, which requires very precise information about the opponent's preferences. As a consequence, in a mixed strategy equilibrium each player faces no cost for deviating from the mixture that makes the opponent indifferent. Therefore, a mixed strategy equilibrium is a very weak equilibrium that does not appear plausible – in particular, not in the presence of strict equilibria. In the context of a market, mixed equilibria are even more implausible as they are inefficient, meaning that there is a positive probability of the parties remaining unmatched. In real life this would probably cause intervention or intermediation.

Hence, this example illustrates that the consideration of extensive form structures can be informative about how plausible an equilibrium prediction is. If the overall equilibrium is supported by very weak equilibria in subgames, there is reason to doubt what this equilibrium prescribes.

4.2 Job Market Signaling

Beyond enabling scrutiny of substructures, like subgames or information sets, the particular formalization of extensive forms by a DEF (as in Definition 4) has further advantages. The job market model by Spence (1973) serves to illustrate how using a DEF, rather than a graph- or sequence-based definition, may simplify the application of refined equilibrium concepts like perfect Bayesian equilibrium (see also Example 9 of AR3).

In that model chance initially assigns a productivity type $\theta \in \{\theta_L, \theta_H\}$, with $0 < \theta_L < \theta_H$, to a potential worker. The prospective employee then decides on a training level $e \in \mathbb{R}_+$ in order to signal her productivity. This signal, but not the productivity, is observed by a competitive industry, which then offers a wage $y \in [\theta_L, \theta_H]$. This wage equals the expected productivity conditional on the observed signal.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

To complete the specification of a game, assume two firms who compete à la Bertrand by offering wages to the worker, who in turn chooses among wage offers. Since the worker will always choose the higher wage, price competition guarantees that both firms offer a wage equal to expected productivity. This ensures that the productivities expected by the two firms are equal along the equilibrium path. The equilibrium concept of perfect Bayesian equilibrium, however, does not force the two firms to hold the same beliefs *off* the equilibrium path.

Imposing the "common belief property" additionally runs into problems when simultaneous moves are represented by cascading information sets. For this traditional representation, outcomes are ordered four-tuples:

$$(\theta, e, w_1, w_2) \in \{\theta_L, \theta_H\} \times \mathbb{R}_+ \times [\theta_L, \theta_H] \times [\theta_L, \theta_H] = W,$$

assuming that firm 1 moves first and ignoring the trivial decision by the worker. When $x_t(\bar{e}) = \{w \in W \,|\theta = \theta_t, e = \bar{e}\}$ denotes the node where type $t = L, H$ has chosen $\bar{e} \in \mathbb{R}_+$, firm 1's beliefs are $\mu_1(\bar{e}) = (\mu_1(x_L(\bar{e})), \mu_1(x_H(\bar{e})))$. Firm 2 moves second at her infinite information set

$$g_2(\bar{e}) = \{y_t(\bar{e}, \bar{w}_1) \,|t = L, H, \; \bar{w}_1 \in [\theta_L, \theta_H]\}$$

where $y_t(\bar{e}, \bar{w}_1) = \{w \in W \,|\theta = \theta_t, \; e = \bar{e}, \; w_1 = \bar{w}_1\}$ denotes the node reached after type $t = L, H$ has chosen $\bar{e} \in \mathbb{R}_+$ and firm 1 has offered $\bar{w}_1 \in [\theta_L, \theta_H]$. Her beliefs are then given by a probability measure $\mu_2(\cdot\,|\bar{e}) : g_2(\bar{e}) \rightarrow \mathbb{R}_+$. This specification is void of economic content, but complicates the common belief property, because it becomes necessary to specify beliefs of firm 2 about the wage offered by firm 1 (and, out of equilibrium, nothing pins down such beliefs). Common beliefs boil down to the statement that the probability mass $\mu_2(\{y_t(\bar{e}, \bar{w}_1) \,|\bar{w}_1 \in [\theta_L, \theta_H]\} \,|\bar{e})$ numerically equals the probability $\mu_1(x_t \,|\bar{e})$, for $t = L, H$ and all $\bar{e} \in \mathbb{R}_+$. The beliefs μ_1 and μ_2 are radically different formal objects!

Since a DEF allows several players to decide at the same move, the natural representation is much simpler. Outcomes become triplets $(\theta, e, (w_1, w_2)) \in \{\theta_L, \theta_H\} \times \mathbb{R}_+ \times [\theta_L, \theta_H]^2 = W$ and all nodes $x_t(\bar{e})$ are directly followed by terminal nodes; the nodes $y_t(\bar{e}, \bar{w}_1)$ are not needed anymore. The information sets $g(\bar{e}) = \{x_t(\bar{e}) \,|t = L, H\}$, which are now common to both firms, consist of two nodes and beliefs for both firms are two vectors $\mu_1(\bar{e}), \mu_2(\bar{e})$ in the one dimensional unit simplex. The common belief property is captured by the statement $\mu_1(\bar{e}) = \mu_2(\bar{e})$ for all $\bar{e} \in \mathbb{R}_+$.

REFERENCES

Alós-Ferrer, C., and K. Ritzberger (2005a): Some Remarks on Pseudotrees, *Order* **22**, 1–9.

Alós-Ferrer, C., and K. Ritzberger (2005b): Trees and Decisions, *Economic Theory* **25**, 763–798.

Alós-Ferrer, C., and K. Ritzberger (2008): Trees and Extensive Forms, *Journal of Economic Theory* **43**, 216–250.

Alós-Ferrer, C., and K. Ritzberger (2013): Large Extensive Form Games, *Economic Theory* **52**, 75–102.

Alós-Ferrer, C., and K. Ritzberger (2016): Equilibrium Existence for Large Perfect Information Games, *Journal of Mathematical Economics* **62**, 5–18.

Alós-Ferrer, C., and K. Ritzberger (2017a): Characterizing Existence of Equilibrium for Large Extensive Form Games: A Necessity Result, *Economic Theory* **63**, 407–430.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Alós-Ferrer, C., and K. Ritzberger (2017b): Does Backwards Induction Imply Subgame Perfection?, *Games and Economic Behavior* **103**, 19–29.
Alós-Ferrer, C., and K. Ritzberger (2017c): Characterizations of Perfect Recall, *International Journal of Game Theory* **46**, 311–326.
Aumann, R.J. (1961): Borel Structures for Function Spaces, *Illinois Journal of Mathematics* **5**, 614–630.
Aumann, R.J. (1963): On Choosing a Function at Random, in F.B. Wright (ed.), *Ergodic Theory*, New Orleans: Academic Press, 1–20.
Aumann, R.J. (1964): *Mixed and Behavior Strategies in Infinite Extensive Games*, *Advances in Game Theory*, Princeton, NJ: Princeton University Press, 627–650.
Aumann, R.J. (1987): Game Theory, in J. Eatwell, M. Milgate, and J. Newman (eds.), *The New Palgrave, A Dictionary of Economics, Vol. 2*, Basingstoke, UK: Macmillan, 460–482.
Bernheim, D.B. (1984): Rationalizable Strategic Behavior, *Econometrica* **52**, 1007–1028.
Billingsley, P. (1986): *Probability and Measure*, New York: John Wiley & Sons, Inc.
Birkhoff, G. (1961): *Lattice Theory*, Providence, RI: American Mathematical Society.
Burdett, K., S. Shi, and R. Wright (2001): Pricing and Matching with Frictions, *Journal of Political Economy* **109**, 1060–1085.
Dalkey, N. (1953): Equivalence of Information Patterns and Essentially Determinate Games, in H.W. Kuhn and A.W. Tucker (eds.), *Contributions to the Theory of Games, Vol. II, Annals of Mathematics Study No. 28*, Princeton, NJ: Princeton Unversity Press, 217–243.
Feldman, M., and C. Gilles (1985): An Expository Note on Individual Risk Without Aggregate Uncertainty, *Journal of Economic Theory* **35**, 26–32.
Fishburn, P.C. (1988): *Nonlinear Preference and Utility Theory*, Baltimore, MD: Johns Hopkins University Press.
Gilboa, I. (1997): A Comment on the Absent-minded Driver Paradox, *Games and Economic Behavior* **20**, 25–30.
Harris, C. (1985): Existence and Characterization of Perfect Equilibrium in Games of Perfect Information, *Econometrica* **53**, 613–628.
Hewitt, E., and K. Stromberg (1965): *Real and Abstract Analysis*, *Graduate Texts in Mathematics Vol. 25*, New York: Springer Verlag.
Kline, J.J., and S. Luckraz (2016): Equivalence between Graph-based and Sequence-based Extensive Form Games, *Economic Theory Bulletin* **4**, 85–94.
Kohlberg, E., and J.-F. Mertens (1986): On the Strategic Stability of Equilibria, *Econometrica* **54**, 1003–1037.
Kreps, D.M., and R. Wilson (1982): Sequential Equilibrium, *Econometrica* **50**, 863–894.
Kuhn, H.W. (1953): Extensive Games and the Problem of Information, in H.W. Kuhn and A.W. Tucker (eds.), *Contributions to the Theory of Games, Vol. II*, *Annals of Mathematics Study No. 28*, Princeton, NJ: Princeton University Press. 193–216.
Lang, K., M. Manove, and W.T Dickens (2005): Racial Discrimination in Labor Markets with Posted Wage Offers, *American Economic Review* **95**, 1327–1340.
Montgomery, J.D. (1991): Equilibrium Wage Dispersion and Interindustry Wage Differentials, *Quarterly Journal of Economics* **106**, 163–179.
Nash, J.F. (1950): Equilibrium points in *N*-person Games, *Proceedings of the National Academy of Sciences* **36**, 48–49.
Nash, J.F. (1951): Non-cooperative Games, *Annals of Mathematics* **54**, 286–295.
Passell, P. (1994): Game Theory Captures a Nobel, *New York Times Business Day*, October 12.
Osborne, M., and A. Rubinstein (1994): *A Course in Game Theory*, Cambridge, MA: MIT Press.
Pearce, D.G. (1984): Rationalizable Strategic Behavior and the Problem of Perfection, *Econometrica* **52**, 1029–1050.
Piccione, M., and A. Rubinstein (1997): On the Interpretation of Decision Problems with Imperfect Recall, *Games and Economic Behavior* **20**, 3–24.
Ritzberger, K. (1999): Recall in Extensive Form Games, *International Journal of Game Theory* **20**, 69–87.
Ritzberger, K. (2002): *Foundations of Non-Cooperative Game Theory*, Oxford: Oxford University Press.
Savage, L.J. (1954): *The Foundations of Statistics*, New York: John Wiley & Sons.
Schwarz, G. (1974): Ways of Randomizing and the Problem of their Equivalence, *Israel Journal of Mathematics* **17**, 1–10.
Selten, R. (1965): Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit, *Zeitschrift für die gesamte Staatswissenschaft* **121**, 301–324 and 667–689.
Selten, R. (1975): Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games, *International Journal of Game Theory* **4**, 25–55.
Shapley, L.J. (1953): Stochastic Games, *Proceedings of the National Academy of Sciences* **39**, 1095–1100.
Spence, A.M. (1973): Job Market Signaling, *Quarterly Journal of Economics* **83**, 355–377.
von Neumann, J. (1928): Zur Theorie der Gesellschaftsspiele, *Mathematische Annalen* **100**, 295–320.
von Neumann, J., and O. Morgenstern (1944): *The Theory of Games and Economic Behavior*, Princeton, NJ: Princeton University Press.
Wichardt, P.C. (2008): Existence of Nash Equilibria in Finite Extensive Form Games with Imperfect Recall: A Counterexample, *Games and Economic Behavior* **63**, 366–369.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

9. Strategic refinements

*Carlos Pimienta**

1 INTRODUCTION

Nash (1950, 1951) formalized the idea of *strategic equilibrium* already used by Cournot (1838). A Nash equilibrium is a description of behavior for every agent in the strategic interaction with the property that no agent can induce a better outcome for herself by modifying her behavior and keeping the other agents' behavior unchanged. Selten (1965, 1975) noticed that, in many relevant cases, rational agents have reasons to deviate from their prescribed Nash equilibrium behavior, anticipating that other agents will not follow their corresponding Nash equilibrium behavior either. This initiated the main research program in game theory whose objective was to propose a definition of strategic equilibrium that, for each strategic situation, identifies the set of *self-enforcing* norms of behavior. That is, the set of norms of behavior with the property that, if every player in the game knows which norm is going to be followed by every other player in the game, no player has any reason to disobey such a norm. This research program uncovered important implications of strategic rational behavior as well as their connection with decision-theoretic principles.

The evolution of such a research program followed a recognizable pattern that we also employ in many parts of this chapter. Many contributions preceding Kohlberg and Mertens (1986) tried to improve upon an existing equilibrium concept by showing some of its deficiencies in an example, arguing how a selected equilibrium was not reasonable, and proposing a new concept that does not suffer from those same deficiencies. This led to a proliferation of equilibrium concepts that either do not fully capture all the different implications of rationality or, if they do, it is at the cost of not satisfying existence in some subset of games.

In contrast, Kohlberg and Mertens (1986) listed a collection of properties that an ideal equilibrium refinement ought to satisfy. Summarizing, it has to exist for every game, players should not use dominated strategies, it has to satisfy backward and forward induction, it has to be robust to iterated elimination of dominated strategies, and it has to be immune to strategically irrelevant presentation effects. They show that this list of requirements forces us to think of a self-enforcing equilibrium not as a single strategy profile, but as a set of strategy profiles where behavior is not necessarily completely pinned down and can vary because of reasons different from the players' preferences. Nevertheless, the equilibrium concept proposed by Kohlberg and Mertens (1986) is unsatisfactory as it does not satisfy backward induction. Mertens (1989, 1991) offers a reformulation of the concept that does satisfy all those properties including backward induction.

* I am grateful to Francesco De Sinopoli, Hari Govindan, Claudia Meroni, Matías Núñez, and Robert Wilson for very useful comments. I thank financial support from the Australian Research Council's Discovery Projects funding scheme DP140102426.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

We cover some classic equilibrium concepts from Nash (1950) to Kohlberg and Mertens (1986) as well as more recent contributions. However, the structure of the chapter and the selection of topics are organized around what seems to be the most successful notion of equilibrium: *stable sets of equilibria* as defined by Mertens (1989, 1991). Hence, the emphasis is different from surveys such as Hillas and Kohlberg (2002), Van Damme (2002) or Govindan and Wilson (2008b). There is little attempt to justify concepts such as *admissibility*, *backward induction*, or *invariance*. We simply take them as desirable properties and refer the reader to Hillas and Kohlberg (2002) for a critical analysis of such properties and many other aspects related to the notion of equilibrium. Similarly, we do not cover epistemic game theory (see Dekel and Siniscalchi, 2014 for a recent survey) or evolutionary stability (see instead, e.g., Samuelson, 2002).

We, of course, adhere to the classical assumptions in game theory. A given game fully represents the relevant strategic interaction. Players are fully rational and strive to obtain the most desirable outcome subject to the other players' behavior. They make their decision in isolation, there is no possibility of correlation or commitment, and binding contracts are not available. The objective is to identify strategically stable or, in other words, *self-enforcing* equilibria; that is, those rules of behavior such that, if every player in the game knew that every other player is going to abide by them then no player in the game would have an incentive to deviate.

In the next section we introduce basic definitions and notation. In Section 3, we quickly cover some of the most prominent equilibrium concepts in non-cooperative game theory. Section 4 defines several set-valued equilibrium concepts that lead to Mertens' (1989, 1991) definition of stable sets of equilibria. We analyze each part of the definition, its properties, and some of its applications to industrial organization. Section 5 makes some remarks on backward induction and Section 6 covers some recent definitions of forward induction. We mention some contributions to the literature of refinements from an axiomatic approach in Section 7.

2 BASIC CONCEPTS AND NOTATION

A *finite normal-form game* is denoted by G. The set of *players* is $N = \{1, \ldots, n\}$ and each player $i \in N$ has a finite *strategy set of pure strategies* S_i and a *utility function* u_i defined on the set of strategy profiles $S := \prod_{i \in N} S_i$. Player i's set of mixed strategies is Σ_i and the set of mixed strategy profiles is correspondingly denoted $\Sigma := \prod_i \Sigma_i$. We denote also by u_i the linear extension of player i's utility function to the set of mixed strategy profiles. An *extensive-form game* is denoted by Γ. The set of decision nodes is X and the set of final nodes is Z. A typical decision node is denoted by $x \in X$ and a typical final node is denoted by $z \in Z$. We also allow for moves of *nature*. The collection of information sets of player $i \in N$ is represented by H_i. At each information set $h \in H_i$ payer i has a finite set of choices $C_i(h)$ available. A pure strategy for player i is a function that for each information set $h \in H_i$ assigns a choice $c \in C_i(h)$. A mixed strategy for player i is a probability distribution on the set of player i's pure strategies. Additionally, a behavior strategy for player i assigns to each information set $h \in H_i$ a probability distribution b_{ih} on the set of choices $C_i(h)$. The set of player i's behavior strategies is B_i and B is the set of behavior strategy profiles. In an extensive-form game, utility functions are defined over the set of final nodes Z. Each (pure, mixed, or behavior) strategy

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

profile induces an *outcome*, i.e., a probability distribution on Z. Hence, we can easily compute players' expected utilities to each strategy profile and, therefore, construct the normal-form representation of the game. We always assume perfect recall so that Kuhn's theorem (Kuhn, 1953) implies that, for any mixed (behavior) strategy of player i and any mixed or behavior strategy profile of the other players, player i has an equivalent behavior (mixed) strategy that induces the same outcome. Under these conditions, we say that such strategies are equivalent, and we focus on either mixed or behavior strategies whenever it is more convenient.

3 FROM NASH EQUILIBRIA TO STRATEGIC STABILITY

Nash (1950, 1951) defines an *equilibrium* as a strategy profile σ with the property that no player can obtain a strictly higher payoff by unilaterally choosing a different strategy than the one specified for her in the profile σ. Suppose that players hold beliefs about how the game is going to be played, which are represented by the strategy profile σ. That is, we may interpret σ as a common system of beliefs held by the players. Since σ has a product structure, each entry provides beliefs about how player i intends to play. Let $\mathrm{BR}_i(\sigma)$ represent the set of player i's strategies that maximizes player i's utility when every other player plays according to σ. We can argue that, if player i is rational, and every player in the game knows that, a minimal consistency requirement implies that beliefs held by every player about player i (that is, the ith entry in σ) should be an element of $\mathrm{BR}_i(\sigma)$. Repeating the same argument for every player, for any strategy profile σ, we can construct the set $\mathrm{BR}(\sigma) := \prod_i BR_i(\sigma)$ and give the following definition of Nash equilibrium:

Definition 1 (Nash equilibrium) A strategy profile σ is a Nash equilibrium if $\sigma \in \mathrm{BR}(\sigma)$.[1]

Thus, if players expect the game to be played according to σ then no player can unilaterally deviate and thereby increase her payoff. A Nash equilibrium serves as a resting point that divides the multi-person decision problem that the game represents into n individual decision problems. In each of these problems, the corresponding player must maximize her utility. So we can say that a strategy profile is self-enforcing *only if* it is a Nash equilibrium.

Nonetheless, Selten (1965) noted that a Nash equilibrium strategy is not necessarily self-enforcing. Indeed, the strategy profile (R, b) in Figure 9.1 satisfies the Nash equilibrium conditions because player 1 prefers (R, b) to (L, b) and player 2 is indifferent between (R, b)

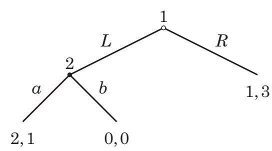


Figure 9.1 Only (a,L) satisfies backward induction

[1] The strategy space Σ is compact and the correspondence $\mathrm{BR} : \Sigma \to \Sigma$ upper-hemicontinuous with non-empty, convex, and compact values, hence Kakutani's fixed-point theorem (Kakutani, 1941) implies that every finite game has a Nash equilibrium.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

and (R, a). However, if player 2's choice matters, that is, if player 1 chooses L then player 2 would certainly play a after observing player 1's choice. To take this into account, we can solve the game using *backward induction* (Zermelo, 1912) whose main idea can be summarized as follows:

> *Players make rational decisions at every juncture of the game, even at those that may have been ruled out by previous behavior.*

Hence, if a strategy profile satisfies backward induction then every player's strategy is still optimal after every contingency, even those that are impossible given the strategy profile. Put differently, the reason why a part of the game is not reached by the strategy profile cannot be based on irrational behavior in such a part of the game. In a game of perfect information like the one in Figure 9.1, the implementation of this principle is easy. First, find maximizing behavior at every decision node where every choice leads to a final node. Then, taking those choices as given, find optimizing behavior at every decision node where every choice leads to either a final node or a decision node whose behavior has been determined in the previous step. Iterating on this process we find a behavior strategy profile that satisfies backward induction. In the example, such a strategy profile is (a, L).

Selten (1965) defined *subgame perfect equilibrium* as a first attempt to apply the idea of backward induction to games with imperfect information. If we accept that rational behavior in a game implies Nash equilibrium then we can request that a candidate equilibrium strategy induce a Nash equilibrium in any *subgame*, i.e., in any part of the game that could be considered as a separate game. For instance, the game on the left-hand side of Figure 9.4 has a subgame starting in the second information set of player 1. Such a subgame has three Nash equilibria (T, L), (B, R), and $(\frac{3}{4}T + \frac{1}{4}B, \frac{1}{4}L + \frac{3}{4}R)$. A subgame perfect equilibrium has to prescribe behavior in the subgame according to one of these three profiles. And player 1 would decide whether or not play *In* depending on whether or not the prescribed Nash equilibrium in the subgame gives him a payoff of at least 2, giving us the subgame perfect equilibria (In, T, L), (Out, B, R), and $(Out, \frac{3}{4}T + \frac{1}{4}B, \frac{1}{4}L + \frac{3}{4}R)$.

By definition, in games without proper subgames, subgame perfection does not impose any restriction beyond Nash equilibrium (e.g., the game on the right-hand side of Figure 9.4). But often the idea of backward induction can still be applied. For instance, consider the game in Figure 9.2 (Selten, 1975, Fig. 1). The strategy profile (A, C, F) is a Nash equilibrium and, since the game does not have proper subgames, also a subgame perfect equilibrium. Under (A, C, F), player 2's decision node is ruled out by player 1 choosing A so player 2 cannot change the outcome of the game. But player 2 is not maximizing her *expected payoff at her information set* taking into account the behavior of player 3. Hence, the principle of backward induction applied here dictates that player 2 should play D, but then player 1 would deviate to B, upsetting the subgame perfect equilibrium (A, C, F).

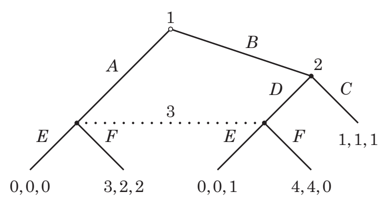


Figure 9.2 Selton's horse

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

To address this issue, Selten (1975) introduced *extensive-form perfect equilibrium*. Given a candidate behavior strategy profile b, the idea is to require that b induce optimal choices at each information set against a slightly modified version of itself where every choice is taken with positive, however small, probability. Under such a modified strategy, no information set is ruled out so that every choice matters and affects the expected payoff.

For our purposes, it is convenient to take a small detour and define first the normal-form version of this concept that we call *normal-form perfect equilibrium* or, simply, *perfect equilibrium*. To that end, take a normal-form G and some (small) number $\varepsilon > 0$.

Definition 2 (Perfect equilibrium) A completely mixed strategy profile σ^{ε} is an ε-*perfect equilibrium* if every player i plays strategies that are non-optimal against $\sigma^{\varepsilon}_{-i}$ with probability less than ε. A strategy σ is a *perfect equilibrium* if and only if there is a vanishing sequence $\{\varepsilon^t\}$ and a sequence $\{\sigma^t\}$ converging to σ such that σ^t is a ε^t-perfect equilibrium for every t.

That is, at an ε-perfect equilibrium σ^{ε}, each player assigns strictly positive probability to each one of her pure strategies, however, always less than ε to those that are non-optimal given σ^{ε}. As ε decreases, the ε-perfect equilibrium is closer and closer to some Nash equilibrium of the game. Such an equilibrium is, by definition, a perfect equilibrium. Every normal-form game has a perfect equilibrium.

We now bring back this concept to extensive-form games. The idea of an extensive-form perfect equilibrium is that, *at every information set* h, the player moving there selects a choice that maximizes her expected utility at h against a completely mixed behavior strategy profile. To implement this idea using Definition 2, we use the *agent normal form* of the extensive-form game. For each player i and each information set $h \in H_i$, define the *agent* i_h of player i and endow her with the strategy set $C_i(h)$. The agent normal form is the normal-form game whose player set is the collection of all agents, each of them with her corresponding strategy set. Every agent of player i has the same utility function as player i. An extensive-form perfect equilibrium is defined as a perfect equilibrium of the agent normal form. In the game of Figure 9.2, the strategy profile (A, C, F) is not an extensive-form perfect equilibrium because if player 1's plays B with some small probability then player 2's choice matters, and she is better off by playing D.

In Figure 9.2, each player moves only once so that the agent normal form and the normal-form representations of that game coincide. In contrast, in Figure 9.3 (Van Damme, 1984, Example 4) player 1 has two agents so that the normal-form representation of that game has two players while the agent normal form is a three-player game. A strategy profile in the agent normal form is a behavior strategy in the extensive form and, by construction, agents of the same player at different information sets behave as different players in the agent normal form. This is not a problem when every player has just one agent (e.g., Figure 9.2) or two agents of the same player never move one after another. However, if two agents of the same player

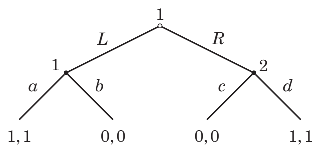


Figure 9.3 Only (La, d) is an admissible strategy profile

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

do move one after another an extensive-form perfect equilibrium may prescribe unreasonable behavior. Going back to Figure 9.3 we see that, in any ε-perfect equilibrium, the second agent of player 1's best response is a and player 2's best response is d. Moreover, the first agent of player 1 prefers R to L in every ε-perfect equilibrium where player 2 plays c with more probability than the second agent of player 1 plays b.

Thus, (R, a, d) is an extensive-form perfect equilibrium even if it seems that the only rational strategy for player 1 is La because it guarantees her a payoff of 1 no matter what player 2 does. In contrast, player 1's utility to playing Ra depends on player 2's behavior.

We say that player i's strategy σ_i is *(weakly) dominated* by strategy σ_i' if for any strategy profile that players other than i may use, player i's payoff if she plays σ_i' is never less than her payoff if she plays σ_i and, for *at least one* strategy profile of the opponents, the difference in utilities is strictly positive. We say that σ_i is *strictly dominated* by strategy σ_i' if such a strictly positive difference holds *for every* strategy profile of the opponents. Following the terminology in decision theory, we say that a strategy is *admissible* if it is not (weakly or strictly) dominated. Luce and Raiffa (1957, p. 287, Axiom 5) propose that the solution to an individual decision problem under uncertainty must be admissible, and point out that such a requirement is equally reasonable in a game-theoretical context. We phrase the admissibility requirement as follows:

Players play admissible strategies in any solution of the game.

The admissibility axiom is typically implemented through normal-form perfect equilibrium.[2] (In fact, Mertens, 2003, p. 397 discusses stronger variants of the admissibility axiom tightly related to perfect equilibrium.) To see this, notice that a weakly dominated strategy is never optimal in an ε-perfect equilibrium because every strategy profile of the opponents is played with positive probability. For example, the strategy profile (Ra, d) is not a perfect equilibrium in Figure 9.3. While d is a best response to every mixed strategy of player 1, the only strategy of player 1 that is a best response to a completely mixed strategy of player 2 is La. Furthermore, Mertens (1995) discusses an example (a symmetrized version of the game in Figure 9.3) where the set of extensive-form perfect equilibria and the set of admissible equilibria are disjoint. Therefore, it seems that we have to discard extensive-form perfect equilibrium or, at least, substitute it for a weaker variant that satisfies backward induction but that never excludes the whole set of admissible equilibria.

An ε-perfect equilibrium (of either the normal form or the agent normal form) assigns positive probability to every pure strategy profile and, consequently, induces well-defined conditional probabilities given any subset of pure strategy profiles. Thus, an agent moving at information set h can compute the conditional probability of a pure strategy profile given that h is reached, that is, given the subset of pure strategy profiles that go through h with positive probability. This allows the agent moving at h to find the set of choices that maximize expected utility at h given such a conditional probability. A sequence of ε-perfect equilibria induces a sequence of conditional probabilities, and the resulting perfect equilibrium is optimal against the limiting conditional probability system. But any arbitrary converging sequence of completely mixed strategies generates a sequence of conditional

[2] Blume, Brandenburger, and Dekel (1991) provide a characterization of perfect equilibrium using a lexicographic belief system representation with full support. They show that a normal-form perfect equilibrium is equivalent to an admissible Nash equilibrium satisfying the common prior and strong independence assumptions.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

probabilities. A weaker version of extensive-form perfect equilibrium can be obtained by ignoring the optimality conditions along the sequence of completely mixed strategies (i.e., the ε-perfect equilibrium conditions) and requiring optimality only against the limiting system of conditional probabilities. This is the concept of *sequential equilibrium* introduced by Kreps and Wilson (1982a). A sequential equilibrium implements backward induction by requiring *sequential rationality* or, in other words, optimality at every information set given a system of conditional probabilities on the set of pure strategy profiles. Such a system of conditional probabilities is called a *system of beliefs*. Those beliefs are not arbitrary; instead they are tied to the strategy profile as follows. Given a behavior strategy profile b (and the equivalent mixed strategy σ), a system of beliefs μ is *consistent* with b if there is a sequence of completely mixed behavior strategy profiles $\{b^t\}$ converging to b (and, thus, an equivalent sequence of completely mixed strategy profiles $\{\sigma^t\}$ converging to σ) such that the induced sequence of beliefs $\{\mu^t\}$ obtained by computing μ^t from σ^t using Bayes' rule converges to μ.[3,4] Given a pure strategy s_i of player i's, define player i's *continuation strategy* at information set $h \in H_i$ to be the restriction of s_i to h and every other information set in H_i that follows h.

Definition 3 (Sequential equilibrium) The behavior strategy profile b is a *sequential equilibrium* if there is a system of beliefs μ consistent with b such that b is sequentially rational, that is, for every player i, every information set $h \in H_i$, and every choice $c \in C_i(h)$ we have $b_{ih}(c) > 0$ only if there exist a continuation strategy that prescribes c at h and that is optimal against the conditional probability induced by μ over the pure strategy profiles of the opponents that enable h.

Every sequential equilibrium is subgame perfect. Looking back at the previous examples, (A, C, F) is not a sequential equilibrium in Figure 9.2 because player 2 is not being sequentially rational in his unique information set. In turn, (R, a, d) is a sequential equilibrium in Figure 9.3 because every player maximizes their expected payoffs at every information set given the behavior of the opponents. For both extensive-form perfect equilibrium and sequential equilibrium we have to construct a sequence of completely mixed behavior strategy profiles converging to the candidate equilibrium. The difference between the two is that while perfect equilibrium requires optimizing behavior against every member of such a sequence, sequential equilibrium only imposes optimality at the limit of the sequence (i.e., at the consistent system of beliefs). Thus, every extensive-form perfect equilibrium is sequential (so every extensive-form game has at least one sequential equilibrium). In fact, Blume and Zame (1994) show that, for *generic payoffs*, the set of sequential and extensive-form perfect equilibria coincide.[5]

---

[3] Kohlberg and Reny (1997) show that a system of beliefs generated in this way properly captures strategic independence, that is, the basic game-theoretical assumption that players take their decision independently. See also Battigalli (1996); Swinkels (1993).

[4] This definition of beliefs is different from Kreps and Wilson's (1982a). They only consider the information needed to implement sequential rationality while here a consistent system of beliefs contains conditional probabilities for any arbitrary subset of pure strategy profiles, even if they do not all go through the same information set. This additional generality will be useful when discussing forward induction in Section 6.

[5] A utility function in an extensive-form game is a point in $\mathbb{R}^{nZ}$, that is, it specifies a utility value for each player at each terminal node. Genericity is used here in the sense that there exists a lower-dimensional set in $\mathbb{R}^{nZ}$ defined by finitely many polynomial inequalities such that, for every point in $\mathbb{R}^{nZ}$ outside that set, every sequential equilibrium is extensive-form perfect (cf. Govindan and Wilson, 2001).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

The set of sequential equilibria always contains equilibria that are admissible. A selection of those admissible sequential equilibria is done by *quasi-perfect equilibrium* (Van Damme, 1984). The concept is also similar to extensive-form perfect equilibrium in that it imposes optimality against a sequence of perturbed mixed strategies as the perturbations vanish. The difference is that when an agent of player i maximizes, he does so ignoring the perturbations in the behavior of future agents of player i. In this way, quasi-perfect equilibrium avoids selecting dominated strategies when two agents of the same player moves one after another.

Definition 4 (Quasi-perfect equilibrium) A completely mixed behavior strategy profile b^ε is an *ε-quasi-perfect equilibrium* if for every player i, every information set $h \in H_i$, and every choice $c \in C_i(h)$ we have $b^\varepsilon_{ih}(c) > \varepsilon$ only if there exists an optimal (given b^ε_{-i}) continuation strategy that prescribes c at h. The profile b is a *quasi-perfect equilibrium* if there is some vanishing sequence $\{\varepsilon^t\}$ and a sequence of profiles $\{b^t\}$ converging to b such that b^t is a ε^t-quasi-perfect equilibrium for every t.

Every extensive-form game has a quasi-perfect equilibrium. In a quasi-perfect equilibrium, each player (as opposed to every agent of each player as in extensive-form perfect equilibrium) is playing a best response against a sequence of completely mixed strategies of the opponents, so every quasi-perfect equilibrium is (normal-form) perfect and, therefore, admissible. Quasi-perfection, like sequential rationality and extensive-form perfect equilibrium, requires player i to optimize at every information set, even at those ruled out by previous behavior. So quasi-perfect equilibrium satisfies backward induction (i.e., it is sequentially rational) and admissibility (i.e., induces a perfect equilibrium). Moreover, for generic assignments of payoffs to ending nodes, the sets of sequential, extensive-form perfect, and quasi-perfect equilibrium coincide (Hillas, Kao, and Schiff, 2002; Pimienta and Shen, 2010).

The definitions of quasi-perfect and sequential equilibrium (as well as backward induction and sequential rationality) make explicit use of the extensive form. They impose some kind of optimality condition at *every* information set so, were the information sets to change, these concepts would impose a different collection of conditions. Namely, it can be that for two extensive-form games that represent the same strategic situation, these concepts give different solutions. This is the case for the two extensive-form games in Figure 9.4. In both games, the set of subgame perfect equilibria and quasi-perfect equilibria coincide. The game on the left-hand side has three quasi-perfect equilibria: (Out, B, R), $(Out, \frac{3}{4}T + \frac{1}{4}B, \frac{1}{4}L + \frac{3}{4}R)$ and (In, T, L). But the game on the right-hand side has (T, L) and a whole continuum from (Out, R) to $(Out, \frac{2}{3}L + \frac{1}{3}R)$.

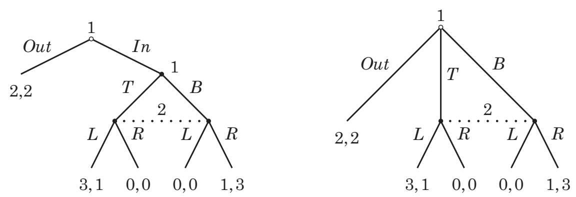


Figure 9.4 Two extensive-form games with the same reduced normal-form representation

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

If we agree that the same strategic interaction can be represented by different extensive-form games then we would ideally want an equilibrium concept whose selection does not depend on which game we choose to model it. Following Kohlberg and Mertens (1986), we call the following principle *invariance*:

> *The solution of a game does not depend on strategically irrelevant aspects of the representation of the game.*

Thus, the analysis of a game should not be affected by presentation effects. In Figure 9.4, we transform the game on the left-hand side into the game on the right-hand side by substituting two consecutive binary choices by a three-way choice. Two extensive forms that can be transformed into each other through a sequence of strategically inessential transformations like this one (Thompson, 1952; Dalkey, 1953; Elmes and Reny, 1994) have the same semi-reduced normal form (which is obtained after eliminating duplicated pure strategies from the normal form, e.g., *OutT* and *OutB* in the normal-form representation of the right-hand-side game in Figure 9.4).

Van Damme (1984) proves that every *proper equilibrium* (Myerson, 1978) of a normal-form game induces a quasi-perfect equilibrium (hence, sequential) in every extensive-form Γ with that semi-reduced normal form (cf. Kohlberg and Mertens, 1986, Proposition 0).

Definition 5 (Proper equilibrium) An *ε-proper equilibrium* is a completely mixed strategy profile σ^ε such that for each player i and any two of his pure strategies s_i and s_i' we have $\sigma_i^\varepsilon(s_i) < \varepsilon\sigma_i^\varepsilon(s_i')$ whenever $u_i(\sigma_{-i}^\varepsilon, s_i) < u_i(\sigma_{-i}^\varepsilon, s_i')$. A strategy profile σ is a *proper equilibrium* if it is the limit point of a sequence $\{\sigma^\varepsilon\}$ of ε-proper equilibria as ε goes to zero.

An ε-perfect equilibrium (Definition 2) does not impose restrictions on non-optimal pure strategies other than they have to be played with probability less than ε. An ε-proper equilibrium, on the other hand, additionally requires that the probability assignment to pure strategies respects the preference ordering of the player at the given profile. Every normal-form game has a proper equilibrium and, as mentioned above, every proper equilibrium induces a quasi-perfect equilibrium of every extensive form with that semi-reduced normal form. Indeed, the set of proper equilibria of the normal-form representation of the extensive forms in Figure 9.4 are (T, L), (Out, R) and $(Out, \frac{1}{4}L + \frac{3}{4}R)$. Notwithstanding this property, proper equilibrium is not a truly invariant concept. The game in Figure 9.5 is equivalent to the (reduced) normal-form representation of the game in Figure 9.4. The extra pure strategy X is a mixed strategy already available before. In this new game, the unique proper equilibrium is (T, L): X strictly dominates B so, in an ε-proper equilibrium player 2 has to consider X to be "infinitely" more likely than B. But then, player 2 must play L as the only reason to play R

	L	R
Out	$2,2$	$2,2$
T	$3,1$	$0,0$
B	$0,0$	$1,3$
$\frac{2}{3}Out + \frac{1}{3}T =: X$	$\frac{7}{3}, \frac{5}{3}$	$\frac{4}{3}, \frac{4}{3}$

Figure 9.5 A game equivalent to the games in Figure 9.4

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

is that B is played with relatively higher probability than T and X. And if player 2 plays L, player 1 must play T.[6]

It follows that a truly invariant solution concept only depends on the reduced normal form of the game, that is, the one obtained after eliminating redundant pure strategies that are already available as convex combination of other existing pure strategies. Proper equilibrium too depends on strategically irrelevant aspects of the game, so it is not surprising that in some examples it selects implausible equilibria. In the well-known example of Figure 9.4, it has been argued that the only reasonable equilibrium is indeed (T, L). Considering either extensive-form representation of this game, the reason is that if player 2 had to make a move while also believing that player 1 is rational, he should attach probability zero to player 1 choosing B because such a strategy is strictly dominated by *Out*. Thus, anticipating this, player 1 should play T and player 2 respond with L. This is an example where a *forward induction* argument selects the only equilibrium that survives invariance and backward induction. Hillas and Kohlberg (2002, p. 1651) conjecture that this relationship may go beyond this simple example (see also Hillas, 1998a). We come back to this point in Section 6.

To capture forward induction in a more encompassing way we can use the following principle suggested by Kohlberg (1990, p. 13) and that we may refer to as *independence of irrelevant alternatives* (see also Van Damme, 2002):

> *A solution is not be vulnerable to the elimination of a strategy that is irrelevant (i.e., certain not to be employed) when players play according to the solution.*

This principle, combined with the admissibility axiom, implies that a solution should not depend on the existence of a dominated strategy. In Figure 9.6, backward induction implies that the solution is the unique proper equilibrium $(T, \frac{1}{2}L + \frac{1}{2}R)$. By admissibility, the weakly dominated strategies M and B are "certain not to be employed". But if we eliminate M, the only admissible equilibrium is (T, R), while if we eliminate B the only admissible equilibrium is (T, L). Thus, independence of irrelevant alternatives tells us that both (T, L) and (T, R) are the solution so, if we want to preserve existence, a satisfactory equilibrium concept must be set-valued and, in this particular case, a solution should contain (T, L), (T, R) and, by backward induction, the unique proper equilibrium $(T, \frac{1}{2}L + \frac{1}{2}R)$. Notice that the principle of independence of irrelevant alternatives is not necessary to arrive at this conclusion in this example. Again, the same can be accomplished with an argument based on invariance and backward induction. Add to the game in Figure 9.6 the pure strategy $X = \frac{1}{2}T + \frac{1}{2}M$ for player 1, the only proper equilibrium in the new game is (T, L). If instead we add the pure strategy $Y = \frac{1}{2}T + \frac{1}{2}B$ then the only proper equilibrium is (T, R).

	L	R
T	1, 1	1, 1
M	1, −1	−1, 1
B	−1, 1	1, −1

Figure 9.6 Iterated elimination of dominated strategies can give different solutions depending on the order of elimination

[6] Importantly, and contrary to proper equilibrium, the set of perfect equilibria is unaffected by adding or deleting pure strategies that are convex combinations of the others.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

To a great extent, the arguments used in this section have been borrowed from Kohlberg and Mertens (1986) who give a more careful description of the properties that an ideal solution concept should satisfy and conclude that such a concept must be set-valued. They proposed three set-valued solution concepts. Here we are only going to define the last one and call it KM-stability.

As mentioned above, a given ε-perfect equilibrium induces a system of conditional probabilities defined on pure strategy profiles. We can make those relative probabilities more explicit by defining a *perturbed game* where probabilities to non-optimal strategies are specified by the description of the game. Thus, consider an arbitrary *completely mixed* strategy profile ς and a strictly positive vector $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$ whose entries are small enough. We call $\eta_i := \varepsilon_i \varsigma_i$ a perturbation for player i and $\eta := (\eta_1, \dots \eta_n)$ a vector of perturbations. We can think of perturbations as changing strategies, so that if player i intends to play σ_i then she actually plays σ_i with probability $(1 - \varepsilon_i)$ and ς_i with probability ε_i. That is, a perturbation η transforms player i's strategy σ_i into $\tau_i(\sigma_i \mid \eta_i) := (1 - \varepsilon_i)\sigma_i + \varepsilon_i \varsigma_i$ and the strategy profile σ into the strategy profile $\tau(\sigma \mid \eta) := (\tau_1(\sigma_1 \mid \eta_i), \dots, \tau_n(\sigma_n \mid \eta_n))$. Equivalently, given a game G and a perturbation $\eta = (\varepsilon_i \varsigma_i)_i$ we can define the (perturbed) game $G(\eta)$ as a game where the strategies are not perturbed but the utility function of player i is given by $u_i^\eta := u_i \circ \tau(\cdot \mid \eta)$. A Nash equilibrium σ^η of $G(\eta)$ is a strategy profile such that no player can get a strictly higher payoff in $G(\eta)$ by changing his or her strategy. And such a strategy profile σ^η satisfies

$$\sigma^\eta \in \mathrm{BR}((1 - \varepsilon_1)\sigma_1^\eta + \varepsilon_1 \varsigma_1, \dots, (1 - \varepsilon_n)\sigma_n^\eta + \varepsilon_n \varsigma_n), \tag{9.1}$$

where BR is the best-response correspondence of the original game G. Of course, a strategy profile σ is a perfect equilibrium of G if and only if there is a vanishing sequence of perturbations η^t and a sequence of strategies $\{\sigma^t\}$ converging to σ such that σ^t is a Nash equilibrium of $G(\eta^t)$ for every t (cf. Definition 2).

A KM-stable set is a minimal set of Nash equilibria that is robust to *every* completely mixed strategy perturbation. We describe such a set of perturbations as follows. For any number $0 < \delta \leq 1$ we construct the set of all completely mixed strategy perturbations of "size" smaller than δ. If we let Σ° be the set of completely mixed strategy profiles then such a set of perturbations is $P_\delta^\circ := \{(\varepsilon_i \varsigma_i)_{i \in N} : 0 < \varepsilon_i \leq \delta,\ \varsigma_i \in \Sigma_i^\circ \text{ for all } i\}$. The equilibrium property of being robust to every completely mixed strategy perturbation is described as follows:

Definition 6 (Property S) A set of equilibria T satisfies Property **S** if it is a closed set of Nash equilibria of G such that for every $\lambda > 0$ there is a δ such that if $\eta \in P_\delta^\circ$ then the game $G(\eta)$ has a Nash equilibrium whose distance with respect to T is less than λ.

This robustness property is similar to the definition of continuity. For any value of λ (which is how far away from T we are allowed to go searching for a Nash equilibrium of the perturbed games) we can restrict the size of the perturbations δ so that *every* perturbed game with a perturbation of size smaller than δ has a Nash equilibrium λ-close to T. The fact that the Nash equilibrium correspondence (which maps each game G into its set of Nash equilibria) is upper-hemicontinuous implies, by definition, that the whole set of Nash equilibria of G satisfies Property **S**. Kohlberg and Mertens (1986) extract a non-trivial refinement by imposing minimality.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Definition 7 (KM-stable set) T is a KM-stable set if it is a minimal set (in terms of set inclusion) satisfying Property **S**.

A KM-stable set always exists. By minimality, every equilibrium in a KM-stable set is perfect. Because of this fact, the collection of KM-stable sets does not change when we add or delete a pure strategy that is a convex combination of other pure strategies. It also satisfies robustness against deletion of irrelevant alternatives in the following two versions. A KM-stable set *contains* a KM-stable set of the game obtained after eliminating a dominated strategy (iterated dominance), and a KM-stable set *contains* a KM-stable set of the game obtained after eliminating a strategy that is never an optimal response against any strategy in the KM-stable set (forward induction).[7]

However, some stable sets may fail to include a proper equilibrium and, therefore, satisfy backward induction. Kohlberg and Mertens (1986) give an example due to Faruk Gul of a KM-stable set made of two equilibrium points, none of which is the unique sequential equilibrium of the game.[8] For completeness, we include Gul's game in Figure 9.7. The subgame starting with player 1's second information set has a unique equilibrium $(\frac{1}{2}M+\frac{2}{2}B, \frac{1}{2}L+\frac{2}{2}R, \frac{1}{2}l+\frac{2}{2}r)$. Therefore, the game has a unique subgame perfect equilibrium, which is also necessarily sequential. This sequential equilibrium is $(T, \frac{1}{2}L + \frac{2}{2}R, \frac{1}{2}l + \frac{2}{2}r)$. By backward induction, a satisfactory solution of this game should include this strategy. But $\{(T, L, l), (T, R, r)\}$ is a KM-stable set of the game: player 1's strategy T is a strict best response against both elements in the set, so it is still a best response against their perturbed versions. Furthermore, for those perturbed games such that $\eta_1(M) \geq \eta_1(B)$ there is a Nash equilibrium where players 2 and 3 play, respectively, L and l; and for those perturbed games such that $\eta_1(B) \geq \eta_1(M)$ there is a Nash equilibrium where players 2 and 3 play, respectively, R and r.

Another desirable property listed by Kohlberg and Mertens (1986) that we have still not mentioned is *connectedness*, that is, every solution must be a connected set of Nash equilibria. Connectedness is a natural condition to identify equilibrium points in the same solution. In this way, we can interpret the different equilibrium points in the same solution as varying continuously with beliefs, presentation effects, or other aspects of the interaction not captured

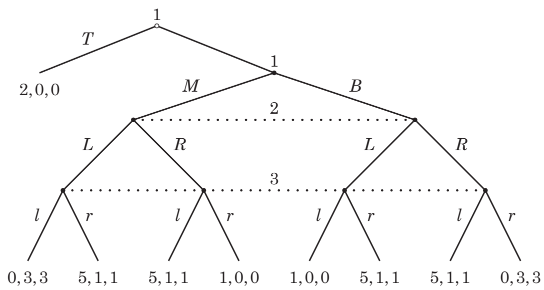


Figure 9.7 Gul's game

[7] Cho and Kreps, (1987) and Banks and Sobel (1987) offer a characterization of KM-stability for a generic class of signaling games. Govindan and Wilson (2006) offer a characterization of Property **S** in terms of lexicographic probability systems.

[8] A strategy profile σ is a *strictly perfect equilibrium* (Okada, 1981) if the set $\{\sigma\}$ satisfies Property **S**. Vermeulen and Jansen (1996) provide a counterexample that shows that a strictly perfect equilibrium may not be proper.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

by preferences. Additionally, for generic extensive-form games, connectedness implies (cf. Kreps and Wilson, 1982a, Theorem 2) that a solution induces a unique outcome so that the departure from a single-valued solution concept is minimal in these games. In any case, as Gul's example demonstrates, KM-stability does not satisfy this requirement either.

4 ESSENTIAL SETS OF FIXED POINTS AND STRATEGIC STABILITY

A game G is defined by a set of exogenous variables (players, strategies, and preferences) from which we can construct the best-response correspondence. A Nash equilibrium is a fixed point of $\mathrm{BR} : \Sigma \to \Sigma$, that is, a point σ that satisfies $\sigma \in \mathrm{BR}(\sigma)$. Some of the desirable properties in the previous section can be implemented by looking at the set of Nash equilibria of games that are either nearby, or that can be constructed from the original game. The set of Nash equilibria of those games can be obtained as fixed points of a perturbed version of the best-reply correspondence. For instance, dominated strategies are eliminated in a perfect equilibrium because they become strictly dominated in any close-by completely mixed strategy perturbed game and, therefore, they cannot be part of a Nash equilibrium in such a perturbed game. As another illustration of this idea, we know that *regular equilibrium* (Harsanyi, 1973; Ritzberger, 1994) satisfies every desirable requirement apart from existence. A property of regular equilibrium is that if σ is a regular equilibrium of G then, locally, there is a bijective function between games that are close to G and their Nash equilibria that are close to σ.

Consider the function from the interval $[0, 1]$ to itself represented in Figure 9.8. Looking at the 45 degree line we see that the function has several fixed points, x_1, the interval $[x_2, x_3]$, x_4, and x_5. Any continuous function "close" to f has a fixed point close to x_1, another one close to the interval $[x_2, x_3]$ and another one close to x_5. But the same is not necessarily true for x_4: some perturbations of f have two fixed points close to x_4 and some other perturbations have no fixed point at all close to x_4. In this sense, x_1 and x_5 are more robust fixed points than x_4. Fort (1951) defines a fixed point of a continuous function to be essential if they satisfy this robustness property. It is not difficult to come up with an example of a continuous function

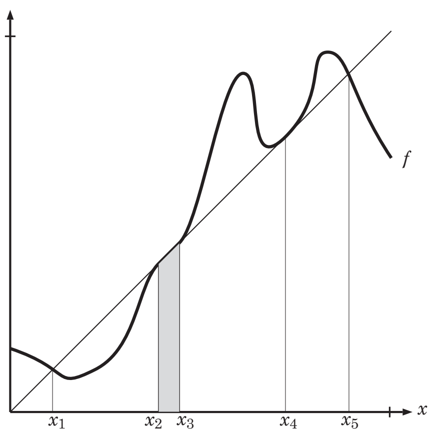


Figure 9.8 Essential and inessential sets of fixed points

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

with no essential fixed point, e.g., the identity function. Hence, Kinoshita (1952) defined a set of fixed points X of a function f to be essential if every function close to f has a fixed point close to X. In Figure 9.8 no fixed point in the interval $[x_2, x_3]$ is essential, but $[x_2, x_3]$ is an essential set of fixed points. But once we accept that we need to consider sets of fixed points if we want to obtain a useful concept, we also have to limit the size of the robust set of fixed points to avoid selecting all of them. In the literature, we normally find conditions such as minimality or connectedness.

Similar ideas can be applied to correspondences in general and to the best-response correspondence in particular. McLennan (2012) provides a unified treatment that we now use to discuss set-valued equilibrium concepts in terms of some robustness property of the best-response correspondence against some set of perturbations. There are at least three parts that we can adjust to obtain a set-valued equilibrium concept in this way: the space of perturbations that we want to consider, the concept of robustness with respect to such perturbations, and how we extract a solution concept from the robustness condition. The way we tune each of these parts determines the game-theoretical properties of the resulting concept.

For the time being, let us fix the robustness property (for every small perturbation there must be a fixed point close to the proposed set) and how we extract a refinement from it (the set should not strictly contain another set that satisfies the robustness property) and look only at the set of perturbations. We first note that the bigger the set of perturbations that we admit, the larger the sets of fixed points that we need to accept. For instance, suppose that the true game in Figure 9.9 is the one obtained when $\varepsilon_1 = \varepsilon_2 = 0$ so that the only admissible equilibrium is (T, L). If we were to require that a selected set of equilibria should be robust against every perturbation of the best-response correspondence generated by payoff perturbations then we have to admit the inadmissible equilibria (T, R) and (B, L) because those are the only equilibria in any close-by game where ε_1 and ε_2 are positive that are close to the set of Nash equilibria of the true game.[9]

KM-stability only considers perturbations of the best-response correspondence generated by strategy perturbations of the game. That is, perturbations of BR that are obtained by composing it with $\tau(\cdot \mid \eta)$ for some $\eta \in P^{\circ}_{\delta}$ (see page 239). Hence, Property **S** can also be stated as follows:

Definition 8 (Property S) A set of Nash equilibria S satisfies property **S** if for every λ there is a δ such that for every $\eta \in P^{\circ}_{\delta}$ we can find a strategy σ that is λ-close to S and, furthermore, $\sigma \in \mathrm{BR}(\tau(\sigma \mid \eta))$.

A KM-stable set is a minimal set that satisfies Property **S**. While the choice of perturbations guarantees admissibility, Figure 9.7 demonstrates that the definition does not guarantee backward induction. Hillas (1990, p. 1372) explains the reason. For those perturbed games

	L	R
T	$1,1$	$1,1+\varepsilon_2$
B	$1+\varepsilon_1,1$	$0,0$

Figure 9.9 A payoff perturbed game

[9] This is the space of perturbation used in the definition of *hyperstability* given by Kohlberg and Mertens (1986).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

such that $\eta_1(M) \geq \eta_1(B)$, even if the strategy profile (T, L, l) is a Nash equilibrium, player 1 strictly prefers B to M but plays the latter with larger probability than the former. That is, (M, L, l) is not a Nash equilibrium of the smallest subgame, which, in turn, implies that (T, L, l) is not subgame perfect. A similar argument applies to (T, R, r). Hillas (1990) defines a new class of perturbations that allows players some "wiggle room" to assign probabilities to non-optimal strategies in a way that respects the players' preference ordering. We call these perturbations γ restrictions.[10] A γ restriction specifies, for each player and each subset of her strategies, the minimal total probability that must be assigned to such a subset. For instance, in Gul's example, one of such γ restrictions specifies the minimum total probability that player 1 must assign to the set $\{M, B\}$, e.g., $\gamma > 0$, as well as the minimum probabilities that M and B must each be played with, e.g., γ^2. Hence, if players 2 and 3 play in a way that player 1 strictly prefers T to B and B to M, then she would play M with probability γ^2 and B with probability $\gamma - \gamma^2$. A minimal set of equilibria such that for every γ restriction there is a Nash equilibrium close to it is called a *quasi-stable set*.[11] Based on the analysis of Gul's example, one can intuitively see how this concept implies backward induction.

Now consider the best-response correspondence of a γ restriction of a game, but defined on the entire set Σ. Let $\varphi : \Sigma \to \Sigma$ be such a correspondence. For each $\sigma \in \Sigma$ it specifies, for each player, the set of optimal responses over those that satisfy the minimal probabilities specified by the γ restriction on the different sets of pure strategies. As such, the set $\varphi(\sigma)$ is as "close" to $\mathrm{BR}(\sigma)$ as those probability restrictions allow. Hillas (1990) defines the space of perturbations of BR to be the collection of those correspondences that are close to BR (and maintains the same notion of robustness as Kohlberg and Mertens, 1986). Note that this collection of perturbations also includes those perturbations of BR considered in Definition 8 so that every set satisfying the following definition also satisfies Property **S**.

Definition 9 (BR-stability) A closed set $T \subset S$ is a BR-set if for any neighborhood V of T there exists a number $\delta > 0$ such that every closed and convex-valued upper-hemicontinuous correspondence φ that is δ-close to the best-response correspondence BR has a fixed point in V.[12] A *BR-stable set* is a BR-set that is connected and only contains perfect equilibria.

In the definition, δ-close means that for every $\sigma \in \Sigma$ the Hausdorff distance between the sets $\mathrm{BR}(\sigma)$ and $\varphi(\sigma)$ is less than or equal to δ (so that $\mathrm{BR}(\sigma)$ is a subset of a neighborhood of $\varphi(\sigma)$ and $\varphi(\sigma)$ is a subset of a neighborhood of $\mathrm{BR}(\sigma)$). Following Hillas et al. (2001), the equilibrium concept is obtained from the robustness condition by directly imposing connectedness and perfection instead of a minimality condition. (A minimality condition can violate the invariance principle: see the ordinality concept discussed in the next section and footnote 15.) BR-stable sets satisfy every desirable property that we have discussed so far. They are connected sets of perfect equilibria and always contain a proper equilibrium. And,

[10] Hillas (1990) calls them δ restrictions. We change the terminology here to avoid re-using the symbol δ.

[11] Hillas (1990) furthermore defines *fully stable sets of equilibria*, modifying the definition of quasi-stability to obtain a concept immune to the addition and deletion of existing mixed strategies as pure strategies.

[12] The results in Hillas et al. (2001) imply that we can alternatively consider those perturbations of the correspondence BR that can be obtained through the collection of continuous functions that map each strategy profile to a strategy perturbation (see also Vermeulen, Potters, and Jansen 1997). Namely, a *CKM-perturbation* of "size" δ is a continuous function $f_\delta : \Sigma \to P_\delta$. A *CKM-set* is a set that such that for any λ there is a δ so that for any CKM-perturbation of size δ the correspondence $\mathrm{BR} \circ \tau(\cdot|f(\cdot)) : \Sigma \to \Sigma$ has a fixed point λ-close to it. A set is a BR-set if and only if it is a CKM-set.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

furthermore, a BR-stable set contains a BR-stable set of a game obtained after deleting a strategy that is either dominated or not a best response against any equilibrium in the BR-stable set.

Nonetheless, BR-stability does not satisfy all the different invariance properties discussed by Mertens (1989, 1991, 1992, 2003). These invariance properties go beyond requiring that the equilibrium concept depends only on the reduced normal form. They correspond to other changes in the description of the game that, by its nature, do not alter the strategic interaction under study. In the next section, we define *Mertens' stable sets* and discuss the invariance properties that they satisfy.

Before that, we mention a recent paper by Grigis de Stefano (2014). He defines a stability concept that changes Kohlberg and Mertens by enlarging the collection of games against which a stable set must be robust. It considers games obtained by introducing new irrelevant players and adding new appropriately chosen strictly dominated strategies. Furthermore, those games are perturbed so that each pure strategy is replaced by several perturbed versions of itself. These auxiliary games are conceived not to alter the relevant strategic interaction of the original game and a stable set of the original game must "produce" an equilibrium for each of these perturbed games in a similar fashion as Property **S**. Even if the final properties obtained improve upon KM-stability, it is also the case that the concept does not satisfy all the invariance properties satisfied by Mertens' stability.

4.1 Mertens' Stable Sets

Mertens' (1989, 1991) definition of stability uses the same set of perturbations as Kohlberg and Mertens (1986). However, he strengthens the robustness property by requiring not only that each sufficiently close perturbed game has a Nash equilibrium close to the stable set, but also imposing a restriction on how those equilibria must "hang together".

We introduce the concept in an incremental way so that we can get some intuition about each of its parts. We begin with the space of perturbations. For any number $0 \leq \delta \leq 1$ we construct the set $P_\delta := \{(\varepsilon_i \varsigma_i)_{i \in N} : 0 \leq \varepsilon_i \leq \delta,\ \varsigma_i \in \Sigma_i \text{ for all } i\}$. The set P_δ contains those completely mixed strategy perturbations P_δ° defined on page 239 and also perturbations that are not completely mixed, i.e., in the boundary ∂P_δ of P_δ. Of course, $P_\delta^\circ := P_\delta \setminus \partial P_\delta$. Recall also that for each perturbation $\eta \in P_\delta$ we can construct the corresponding payoff perturbed game $G(\eta)$ with the same player set and strategy sets as G but where the utility of each player i is given by $u_i^\eta(s) = u_i(\tau(s \mid \eta))$. Thus, we can consider the restriction of the Nash equilibrium correspondence $\mathrm{NE}_u : P_1 \to \Sigma$ that takes each strategy perturbation $\eta \in P_1$ and maps it to the set of Nash equilibria of the game $G(\eta)$. We are only interested in equilibria of perturbed games that are close to some candidate set of equilibria of G. Hence, if the set $T \subset \Sigma$ is such a candidate, we look for a subset E of the graph of NE_u that satisfies $T = \{\sigma \in \Sigma : (0, \sigma) \in E\}$. The subset of E where perturbations are smaller than some δ is denoted $E_\delta = \{(\eta, \sigma) \in E : \eta \in P_\delta\}$. We also construct the subset of E_δ where perturbations are completely mixed $E_\delta^\circ := \{(\eta, \sigma) \in E : \eta \in P_\delta^\circ\}$ and where perturbations are in the boundary $\partial E_\delta := \{(\eta, \sigma) \in E_\delta : \eta \in \partial P_\delta\}$. (Note that ∂E_δ and E_δ° are not necessarily the boundary and interior of E_δ.) Once we have identified a relevant subset of the graph of the Nash equilibrium correspondence of perturbed games, we can use it to impose a robustness condition more strongly than simply requiring that every nearby perturbed game has a nearby equilibrium. For instance, following

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Hillas and Kohlberg (2002, p. 1647), we can provide a slightly strengthened version of KM-stability:

Definition 10 A set of equilibria T is KM*-stable if there is a closed set $E \subset \text{graph}(\text{NE}_u)$ and some δ with $T = \{\sigma : (0, \sigma) \in E_\delta\}$ such that

- for every $0 < \delta' \leq \delta$ the set $E^\circ_{\delta'}$ is connected and satisfies $E_{\delta'} = \text{cl}(E^\circ_{\delta'})$; and
- $p : E_\delta \to P_\delta$ is surjective.

The only extra requirement with respect to KM-stability is the connectedness requirement in the first bullet point (we also dropped the minimality condition). We can think of E°_δ as capturing beliefs supporting the set of equilibria. For any $(\eta, \sigma) \in E^\circ_\delta$ we can compute the completely mixed strategy $\tau(\sigma \mid \eta)$ and, therefore, the corresponding system of conditional probabilities. As δ decreases we can construct sequences of those systems of beliefs which, at the limit, produce a system of beliefs associated with the corresponding perfect equilibrium in T. The connectedness requirement implies that both the collection of beliefs supporting the KM*-stable set and the KM*-stable set itself are connected sets. However, the improvement with respect to KM-stability is not substantial. Hillas (1990, p. 1386) modifies Gul's example to obtain a game with a KM*-stable that also fails to include a proper equilibrium (see also Hillas, 1998b).

Mertens' (1989) reformulation of strategic stability imposes a yet stronger restriction on how the set E_δ should look. An implication of such a restriction is that not only must the projection $p : E_\delta \to P_\delta$ be surjective but also it must not be "equivalent" to a function that is not surjective. If we consider two functions to be equivalent when one can be continuously deformed into the other then the set E_δ cannot have any hole in it. As an illustration, consider Figure 9.10 and the projection from the set consisting of the two short horizontal lines (whose union plays the role of E_δ – the set ∂E_δ is represented by the union of the two dots) to the larger horizontal line (which plays the role of P_δ – the set ∂P_δ is represented by the union of the two short vertical lines). This projection mapping is surjective, however, it can be continuously deformed into a function that is not surjective, namely, one that maps the lower segment of E_δ to the left end of P_δ and the upper segment of E_δ to the right end of P_δ. Of course, in this figure, what would be the set E°_δ is not connected, but a similar example can be constructed in a higher dimensional space where E°_δ is connected and still has a hole in it that is "vertical" to the space of perturbations. The example provided by Hillas (1998b) of a KM*-stable without a proper equilibrium has precisely this feature.

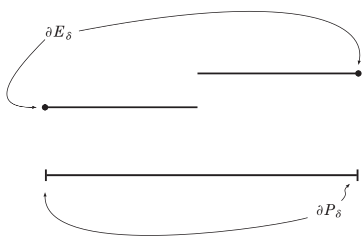


Figure 9.10 The projection mapping can be continuously deformed into a non-surjective function

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Based on this, let us provide a definition of an equilibrium concept that is one step closer to Mertens' stability. This definition appears in Mertens (1991, p. 704), Hillas et al. (2001, p. 615) and Govindan and Wilson (2008a, p. 792). We first formalize what it means to continuously deform a function into another. Consider two topological spaces, X and Y and let $A \subset X$ and $B \subset Y$. We write $f : (X, A) \to (Y, B)$ for a function $f : X \to Y$ such that the image of A is in B, that is $f(A) \subset B$. Two maps between pairs $f, g : (X, A) \to (Y, B)$ are *homotopic relative to A* if there is a continuous function $H : [0, 1] \times X \to Y$ such that $H(0, x) = f(x)$ and $H(1, x) = g(x)$ for all $x \in X$ and, moreover, $H(t, a) = f(a) = g(a)$ for all $t \in [0, 1]$ and all $a \in A$. We say that the map $f : (X, A) \to (Y, B)$ is *essential in homotopy* if it is not homotopic relative to A to a map whose image is in B.

Definition 11 (Homotopy-stable set) A set of equilibria T is homotopy-stable if there is a closed set $E \subset \text{graph}(\text{NE}_u)$ and some δ with $T = \{\sigma : (0, \sigma) \in E_\delta\}$ such that

- for every $0 < \delta' \leq \delta$ the set $E^\circ_{\delta'}$ is connected and satisfies $E_{\delta'} = \text{cl}(E^\circ_{\delta'})$; and
- $p : (E_\delta, \partial E_\delta) \to (P_\delta, \partial P_\delta)$ is essential in homotopy.

Every homotopy-stable set is a BR-set (Hillas et al., 2001). Therefore, they are connected sets of perfect equilibria and always contain a proper equilibrium. They satisfy a stronger version of iterated dominance and forward induction; a homotopy-stable set contains a homotopy-stable set of any game obtained after eliminating a strategy that is at its minimum probability in every ε-perfect equilibrium close to the homotopy-stable set. To get a feeling for the implications of the strengthening of the robustness requirement (the second bullet point in the previous definition) with respect to the definition of KM*-stable sets, we note it implies the following strong fixed-point property: any upper-hemicontinuous, convex-valued correspondence (e.g., any continuous function) $F : (E_\delta, \partial E_\delta) \to (P_\delta, \partial P_\delta)$ has a point of coincidence with the projection mapping p, that is, a point $e \in E$ such that $p(e) \in F(e)$.[13] Then, to see why every homotopy-stable set satisfies backward induction, construct the correspondence that assigns to each perturbation-strategy pair (η, σ) in E_δ the set of strategy perturbations that, for each player i and each pair of strategies, respect the upper bound on probabilities imposed by the $\delta/2$-proper equilibrium conditions on the less preferred pure strategy given σ_{-i}. This correspondence satisfies the necessary conditions for the previous fixed-point result to hold and, therefore, there is a point $(\eta, \sigma) \in E_\delta$ such that σ is a $\delta/2$-proper equilibrium. Taking a sequence as δ vanishes proves that every homotopy-stable set has a proper equilibrium. (See Mertens, 1989, p. 597 for details.)

Homotopy-stable sets are a good approximation to Mertens' stable sets but they still do not satisfy all the implications of the invariance principle. The invariance principle tells us that an equilibrium concept must not depend on strategically irrelevant aspects of the game. So we need to be specific about what we mean by "strategically irrelevant".

We first describe strategically irrelevant aspects of the game that do not modify the set of players. For example, we have already discussed that an invariant equilibrium concept should only depend on the reduced normal form of the game. Mertens (2003) incorporates this requirement as part of the *ordinality* principle. Mertens (1992) describes ordinality as "the unifying idea behind many different desiderata: Normal form invariance, utility invariance,

[13] See, e.g., Govindan and Wilson (2008a, Lemma A.3).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

neutrality as to names of strategies and as to duplication of them, that pure and mixed strategies should be treated identically (the only distinctions stemming from preferences), that only choice sets matter and not the whole preference ordering, that choices are always (in a strong sense) admissible and expected utility maximizing, etc." Shortly, ordinality means the following. Take two games with the same player set. If each player has the same set of admissible best responses in these two games for each belief she might have *over the set of mixed strategy profiles of the opponents* then the two games must have the same solutions. Mertens shows that an equilibrium concept is ordinal if it is both normal-form invariant and admissible best-response invariant. An equilibrium concept is *normal-form invariant* if given two games G and G' with the same reduced normal form (up to relabeling of pure strategies) and a family of surjective linear mappings $f = (f_i)_{i \in N}$ from the strategy sets of one game to the strategy sets of the other game that preserve payoffs, we have that the image of every solution of the first game is a solution in the second game, and that the inverse image of a solution of the second game is the union over all the solutions in the first game whose image is precisely that solution. An equilibrium concept is *admissible best-response invariant* if any any two games with the same set of pure strategies (up to relabeling) and such that their best-response correspondences coincide for every completely mixed strategy profile have the same solutions.[14,15]

A game can be also be presented within an environment with a different player set but where the relevant strategic interaction is not affected. Consider the following variants:

1. *Small worlds axiom* (Mertens, 1992). Suppose that a game G with set of player N is embedded in a larger game G' with a larger set of players N' so that every player in G is in G'. Suppose that in this larger game, the best-response correspondence of any player in N is not affected by the strategic choices of players in N' that are not in N and, moreover, coincides with the best-response correspondence of the first game. The small worlds axiom specifies that the solutions of the game G are precisely the projections of the solutions in G'.
2. *Decomposition*. Consider two different games played by two different sets of players. One can formalize the situation as a big game whose player set is the union of the player sets of the two small games. An equilibrium concept satisfies the decomposition property if the projection of each solution of the big game into one of the smaller games is a solution of the small game and, moreover, the product of two solutions, one for each of the two small games, is a solution of the big game.
3. *Player splitting*. An equilibrium concept satisfies player splitting if it is immune to splitting two agents of the same player that do not move one after the other into two different players.

Mertens' stable sets satisfy ordinality, the small words axiom, decomposition, and player splitting. It is worthwhile pointing out that the set of Nash equilibria satisfy all these invariance

[14] Homotopy-stable sets are admissible best-response invariant but not normal-form invariant.

[15] As mentioned before, a minimality condition imposed on an equilibrium concept may violate the ordinality requirement, in particular, normal-form invariance. Given two games that are ordinally equivalent and a function mapping strategies in the first game to equivalent counterparts in the second game, a minimal set satisfying some robustness condition in the first game can be mapped into a set of the second game that, even if it satisfies the same robustness condition, may not be minimal. Vermeulen and Jansen (2001) give an example of a minimal BR-set that behaves in this way.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

properties as well. The definition of Mertens' stability is obtained after strengthening the robustness property of the definition of homotopy stability so that all these invariance properties can be obtained. For completeness, we include the definition here and note that every stable set is homotopy-stable:

Definition 12 (Mertens' stable set) A set of equilibria T is *stable* if there is a closed set $E \subset \text{graph}(\text{NE}_u)$ and some δ with $T = \{\sigma : (0, \sigma) \in E_\delta\}$ such that

- for every $0 < \delta' \leq \delta$ the set $E^\circ_{\delta'}$ is connected and satisfies $E_{\delta'} = \text{cl}(E^\circ_{\delta'})$; and
- $p : (E_\delta, \partial E_\delta) \to (P_\delta, \partial P_\delta)$ is essential in cohomology.[16,17]

Of course, stable sets are connected sets of perfect equilibria and always contain a proper equilibrium. Every stable set contains a stable set of the game obtained after eliminating a strategy that is played with probability less than ε in every ε-perfect equilibrium close to the stable set. Hence, stable sets also satisfy forward induction and robustness against iterated elimination of dominated strategies.

The robustness condition involves concepts of algebraic topology that can be difficult to verify in applications. However, sometimes, stronger properties are easier to check, e.g., if the projection from E°_δ to P°_δ is a homeomorphism then the projection mapping is necessarily essential in cohomology, which implies that every strict equilibrium, every regular equilibrium, and every *strongly stable equilibrium* (Kojima, Okada, and Shindoh, 1985) are singleton stable sets. Likewise, if we know that the game has finitely many equilibria, like generic normal-form games (Harsanyi, 1973), generic extensive-form games (Kreps and Wilson, 1982a), or generic voting games (De Sinopoli, 2001) then each component of Nash equilibria maps into a unique outcome. Hence, in these cases, it is easier to verify whether a given outcome is stable than whether a particular subset of equilibria is a stable set. For instance, Ritzberger (1994) shows that every Nash equilibrium component with non-zero *index* contains a stable set. Finding a stable outcome is even easier when iterated deletion of dominated strategies reduces the game to a game with a unique perfect equilibrium outcome. It is also the case that every *persistent retract* (Kalai and Samet, 1984) contains a stable set (Mertens, 1992).[18]

4.2 Some Applications of Strategic Stability to Industrial Organization

In applications, it is more tractable to characterize the set of stable outcomes by taking advantage of their properties. For example, Govindan (1995) shows that the Kreps and Wilson (1982b) version of the chain store game has a unique stable outcome. Kreps and Wilson resolve the chain store paradox, introducing incomplete information so that potential entrants

[16] See Govindan and Mertens (2004) for an equivalent definition of stability.

[17] Mertens' original definition is in terms of homological essentiality. A map is essential in (co)homology if the induced map between the corresponding (co)homology groups is different from zero. There are different versions of this definition that depend on the coefficient module used to construct the (co)homology groups, which, therefore, lead to different solution concepts. For the purposes of this chapter, these differences are unimportant. The reader can find an accessible introduction to algebraic topology in Hatcher (2002).

[18] A *retract* is a set $\Lambda := \prod_{i \in N} \Lambda_i$ such that $\Lambda_i \subset \Sigma_i$ is non-empty, closed and convex set of mixed strategies. A persistent retract is a minimal retract Λ such that every strategy profile sufficiently close to Λ has a best response in Λ. Myerson and Weibull (2015) define related concepts called *coarsely tenable block* and *finely tenable block*. These concepts are used to define *settled equilibria*.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

do not know if they are facing a strong monopolist who would not pay any short-term costs to deter entry or a weak monopolist whose short-term incentives dictate that she should acquiesce to entry. The paradox is resolved because, unlike in the perfect information case, the (weak) monopolist does not acquiesce at every stage but fights entry in early stages to build a reputation for being strong only to acquiesce in the last stages. However, there are also other "nonsensical" sequential equilibria that Kreps and Wilson eliminate by assuming that entrants' beliefs are such, at every stage, that they believe that a fight was more likely to have come from the strong monopolist. Even if such an assumption is intuitively appealing it is also rather ad hoc. Govindan (1995) shows that the outcome induced by the Kreps and Wilson (1982b) solution (henceforth *K-W solution*) is induced by the unique stable outcome of the game. We quickly summarize the game and the relevant arguments.

Thus, consider the following T-stage game. In each stage, a different potential entrant decides whether entering a market currently served by a monopolist or staying out. If she enters then the monopolist chooses between fighting (F) or acquiescing (A). The payoffs to the entrant are 0 if she stays out, $0 < b < 1$ if she enters and the monopolist acquiesces, and $b - 1$ if she enters and the monopolist fights. The monopolist can be of two types. With probability p she is strong and with probability $1 - p$ she is weak. Whatever her type, the monopolist receives $a > 1$ if the entrant stays out. If the potential entrant enters the market then the weak monopolist obtains a payoff of -1 if she fights and 0 if she acquiesces. In turn, the strong monopolists receives, respectively, payoffs equal to 0 and -1. The monopolist's payoff is the payoff accumulated in all the stages.

We index time backwards so that the first entrant is Entrant T and the last entrant is Entrant 1. In the K-W solution, Strong always fights entry. Weak fights entry with positive probability unless she acquiesced in the past, in which case her type is revealed, every remaining entrant enters and Weak acquiesces at every stage. At stage $t \neq 1$, if Entrant t's posterior probability p_t that she is facing Strong satisfies $p_t \geq b^{t-1}$ then Weak fights entry with probability 1, otherwise she fights with probability $((1 - b^{t-1})p_t)/((1 - p_t)b^{t-1})$ so that Entrant $t - 1$ is indifferent between entering or not. At stage $t = 1$ Weak acquiesces. Entrant t's posterior probability p_t is either p if $t = T$ or if Weak always fought with probability 1 in the past, 0 if Weak acquiesced sometime in the past, or b^t if Weak mixed in the previous stage and the current history terminates with a fight. Therefore, Entrant t stays out if $p_t > b^t$ and enters if $p_t < b^t$. If $p_t = b^t$ then Entrant t stays out with probability $1/a$ (so that Weak is indifferent between fighting and acquiescing at $t + 1$).

Govindan (1995) shows that the outcome generated by the K-W solution is the unique stable outcome. The reader is referred to the original paper for details but, vaguely speaking, the argument is as follows:

1. In a perfect equilibrium, given a history h, if Strong acquiesces with positive probability then Weak acquiesces with probability one after some history h' of fights and outs that follows h and before the last stage. Otherwise, Weak would be better off by acquiescing after h and then mimicking strong.
2. If σ is a perfect equilibrium such that Strong always fights with positive probability then σ induces the K-W outcome. This is shown by induction assuming that if from t onwards σ induces the K-W outcome then from $t + 1$ onwards σ also induces the K-W outcome.
3. Consider the perturbed game where Strong always fights and Weak always acquiesces with at least probability η, while each entrant enters and stays out with at least probability

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

δ. Since Strong always fights with small positive probability, in an equilibrium of the perturbed game, Weak also fights with positive probability not to reveal her type. But then, by (1), Strong never acquiesces with positive probability. By (2), an equilibrium of this perturbed game must be close to a perfect equilibrium that induces the K-W outcome.
4. The subset of perfect equilibria that generates the K-W outcome is disconnected from the subset of perfect equilibria that do not.
5. Since stable sets are connected sets of perfect equilibria the unique stable outcome of the game is the K-W outcome.

Govindan (1995) also shows that KM-stability does not suffice to eliminate other sequential equilibrium outcomes. But sometimes it is enough to show that a candidate equilibrium set does not even satisfy KM-stability. Indeed, Glazer and Weiss (1990) study a two-player two-period model where players (firms) play the prisoner's dilemma in the first stage (firms choose either high or low prices and the payoffs capture price competition where the dominant strategy is to undercut the competition) followed by a coordination game (where firms choose either complementary advertising, giving them both a high payoff, or negative advertising, giving them both a lower payoff). They show how no stable set contains the collusive outcome where firms choose high prices and then engage in cooperative advertising, nor the outcome where we observe two different prices.

Even if the details of the payoffs matter, we can informally argue why the collusive outcome is not part of a stable set. The collusive outcome relies on playing the Pareto-dominant equilibrium in the second stage if both players collaborate in the first stage, and playing the Pareto-dominated equilibrium in the second stage if a player deviated in the first stage. However, charging a lower price in the first stage and then participating in one's own punishment in the second stage is a strategy that, by the choice of payoffs, is dominated by the equilibrium outcome and, correspondingly, is played with minimum probability at any ε-perfect equilibrium close to a perfect equilibrium generating the collusive outcome. Thus, the collusive outcome is not a stable outcome.

Analogous forward-induction arguments are explored by Kohlberg (1990), Van Damme (1989) and Osborne (1990). This method of identifying stable outcomes is also used in some political economy models (De Sinopoli, 2000; De Sinopoli, Dutta, and Laslier, 2006). We analyze yet another example at the end of Section 6.

5 BACKWARD INDUCTION

In a sequential equilibrium, behavior *off the equilibrium path* disciplines behavior *on the equilibrium path*. Reny (1992) argues that if a player happens to deviate from the equilibrium path then this is evidence that his behavior was indeed not disciplined by the strategy profile and that he either holds different beliefs about how the others are going to play after a deviation or that, maybe, she is not rational at all. Hence, Reny argues that we may not insist on a player's optimizing behavior after she has deviated and weakens sequential equilibrium as follows:

Definition 13 (Weakly sequential equilibrium) The behavior strategy profile b is a *weakly sequential equilibrium* if there is a system of beliefs μ consistent with b such that, for each

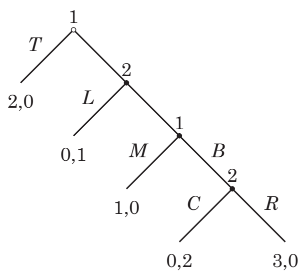
Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

player i, every information set h that b_i does not exclude, and every choice $c \in C_i(h)$ we have $b_{ih}(c) > 0$ only if there exists an optimal (given μ) continuation strategy that prescribes c at h.

Reny proves that every perfect equilibrium induces a weakly sequential equilibrium. The opposite is not true. However, if in Definition 13 instead of requiring optimality at the limit against the system of beliefs μ we require optimality along the sequence of completely mixed strategies supporting μ then we obtain a concept that we may call *weakly quasi-perfect equilibrium* (cf. Definition 4). And every quasi-perfect equilibrium induces a perfect equilibrium.

To illustrate Definition 13 and how it fits within the context of stable sets consider the game of Figure 9.11 (this is Figure 3 in Reny, 1992). The unique stable set coincides with the set of perfect equilibria, that is, the continuum from the unique backward induction solution (T, L) to (T, C). Of course, every point in this continuum corresponds to a weakly quasi-perfect equilibrium of the extensive from. We can interpret them as the set of equilibria such that the equilibrium that is actually in place changes continuously with player 2's beliefs about player 1's intentions in her second information set, which could conform with backward induction (in which case player 2's strategy should be L) or not (so that C starts looking like a viable option). In other words, our analysis of the game implies that player 1, if rational, would play T. But then, this means that a self-enforcing norm of behavior has to be somewhat imprecise about how players would behave if player 1 did not play T as, if that happened, there is no argument based purely on rationality in the game that can delineate behavior. (Hillas and Kohlberg, 2002, p. 1624, make a similar point.) Moreover, a version of this argument can be generalized to generic extensive-form games given that they have finitely many sequential equilibrium outcomes. This, in turn, implies that different equilibria in the same stable set only differ off the equilibrium path.

To conclude this section, we note that Hillas (1990) lists different variants of backward induction. One of them coincides with the concept that we have used (a solution must include a proper equilibrium). Another one is a *subgame consistency* property of a solution concept, meaning that a solution of the game, when projected onto a subgame, contains a solution of the subgame. This property could be conceived as a basis of an algorithm procedure to compute a solution of a game from the solutions of its subgames. Govindan (1996) shows that stable sets satisfy this property, i.e., for any stable set of (the reduced normal form of) an extensive-from game contains a stable set of (the reduced normal form of) any of its subgames.[19]



	L	C	R
T	2,0	2,0	2,0
M	0,1	1,0	1,0
B	0,1	0,2	3,0

Figure 9.11 The unique stable set is the set of perfect equilibria

[19] Brandenburger and Friedenberg (2011) propose a related concept of backward induction. Fix a solution concept. A given solution defines a collection of choices that are "disallowed" by the solution because, at every

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

6 FORWARD INDUCTION

The principle of independence of irrelevant alternatives and the derived concept of forward induction that were introduced in Section 3 are self-referential, e.g., a stable set contains a stable set of the game obtained by deleting a strategy that is played with probability less than ε in every ε-perfect equilibrium close to the stable set. In this section, we revise some definitions of forward induction that have been proposed to determine whether or not a strategy or a set of strategies conform with forward induction.

Consider again Figure 9.4 and the quasi-perfect equilibrium strategy (Out, B, R). In this strategy profile, player 1 selects *Out* because, if she chooses otherwise, the subgame is played according to (B, R). This behavior strategy corresponds to (Out, R) in the normal-form representation on the right-hand side. We can ask what would player 2 do if she was given the opportunity to move (in the extensive-form representation) or if her strategy choice mattered (in the normal-form representation). The answer depends on how likely player 2 thinks player 1's choice of B is relative to T.

Forward induction is the principle that prescribes players to maintain the assumption that their opponents chose their strategies rationally as long as that assumption is justifiable. In an extensive-form game this implies that, given an equilibrium strategy profile, players moving at information sets that are not supposed to be reached should, whenever possible, make sense of those deviations as being part of a rational pattern of behavior. In practical terms, this should be reflected in the beliefs held at those zero probability information sets and on its consequences on optimizing behavior. Under such a principle, if (Out, R) is proposed as a way of playing the game in Figure 9.4 and if player 2 is confronted with the unexpected event that she has to move then she would have to assume that player 1 chose rationally. That translates into discarding the possibility that player 1 chose B as it is a strictly dominated strategy. Strategy T, on the other hand, is a perfectly rational choice if player 1 expects player 2 to understand that she deviated to T and not to B. Hence, (Out, R) violates forward induction because player 2's strategy R is only justified by assuming that player 1 took an irrational deviation while a rational one was available. The only quasi-perfect equilibrium that survives this argument in this game is (T, L).

While it is clear what forward induction entails in a game like the one in Figure 9.4, capturing the idea through an operational definition is more complex. Cho and Kreps (1987) focus on signaling games and propose the *intuitive criterion*. A sequential equilibrium satisfies the intuitive criterion if the receiver's belief after a non-equilibrium message assigns probability zero to those types who prefer the equilibrium outcome to any outcome that follows after her sending that message. (If no sender can benefit from sending that message then they can take any value.) Banks and Sobel (1987) propose the D1 (and the D2) criterion. Suppose that a type t_1 of sender would obtain a utility above her equilibrium utility if after sending a non-equilibrium message m the receiver responds with some action in some set C. If there is another type t_2 of sender (a set of types T not including t_1) and a strict superset C' of C such that for any choice in $c \in C'$ type t_2 (some type in T) prefers the outcome where

point of the solution, they are either taken with probability zero or belong to an information set that is never reached. A solution concept satisfies Brandenburger and Friedenberg's definition of backward induction if, for any solution of an extensive-form game and any subgame, the induced set of outcomes *coincides* with the set of outcomes induced by a solution of the reduced game obtained after eliminating the choices disallowed by some solution of the subgame.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

she sends message m and the receiver responds with c to the equilibrium outcome then the D1 (the D2) criterion dictates that the receiver's beliefs assign probability zero to type t_1 after message m. An iterative application of these criteria leads to divinity and universal divinity (Banks and Sobel, 1987).

Govindan and Wilson (2009) provide a general definition of forward induction for extensive-form games with perfect recall consistent with the idea of assigning probability zero to those strategy profiles that do not satisfy a rationality requirement. In their definition, forward induction is a property of an equilibrium outcome and not of a strategy profile. A pure strategy s_i of player i is *relevant for outcome* $\zeta \in \Delta(Z)$ if there is a weakly sequential equilibrium b (see Definition 13) inducing ζ such that s_i induces an optimal continuation strategy against the beliefs consistent with b at every information set that s_i does not exclude. Therefore, to find player i's pure strategies that are relevant for ζ one has to find every weakly sequential equilibrium inducing ζ and, for each of them, the strategies s_i that prescribe an optimal continuation at every information set not ruled out by s_i. Because s_i is an optimal response against some weakly sequential equilibrium, player i must be indifferent between ζ and the induced outcome if she deviated to s_i. A player's set of strategies that are relevant for an outcome are to be understood as those strategies that satisfy the rationality requirement mentioned above. Now, an information set is *relevant for an outcome* if it is not ruled out by *every* strategy profile that is relevant for that outcome. Govindan and Wilson's forward induction imposes restrictions on beliefs at those information sets that are relevant for an outcome:

Definition 14 (GW-forward induction) An outcome satisfies forward induction if it is induced by a weakly sequential equilibrium in which, at every information set that is relevant for that outcome, the support of the belief of the corresponding player assigns probability zero to those strategy profiles that are not relevant for that outcome.

Govindan and Wilson (2009) show that this concept implies the intuitive criterion, D1, and D2. But a more a remarkable result is that, for a generic assignment of payoffs to ending nodes of a two-player extensive form, an invariant sequential equilibrium outcome satisfies forward induction. This result was suggested by Hillas, (1998a) and Hillas and Kohlberg, (2002). For an illustration, see the analysis of Figures 9.5 and 9.6. To describe what an invariant sequential equilibrium outcome is, note that given an outcome $\zeta \in \Delta(Z)$ there is a set of strategy profiles that induce ζ. Each of those strategy profiles can be replicated by a strategy profile of the reduced normal form. Hence, we can find the set of strategies of the reduced normal form that induce ζ. Say that two extensive-form games Γ_1 and Γ_2 are *equivalent* if they have the same reduced normal form. Furthermore, if Γ_1 and Γ_2 are equivalent games with normal-form representation G, we say that outcome ζ_1 of game Γ_1 is *equivalent* to outcome ζ_2 of game Γ_2 if ζ_1 and ζ_2 are induced by the same set of strategies of G. Govindan and Wilson (2009) show that, for a given outcome of a generic extensive-form games, if every equivalent game has an equivalent sequential equilibrium outcome then such an outcome satisfies Definition 14.

Being a property of an element of $\Delta(Z)$, GW-forward induction is likely to depend on the details of the extensive-form representation. That is, two extensive-form games representing the same strategic interaction may have different sets of final nodes. The game in Figure 9.12 is another representation of the game in Figure 9.4. And, in this representation, the outcome induced by the profile (Out, R) satisfies forward induction. Man (2012) argues that forward

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

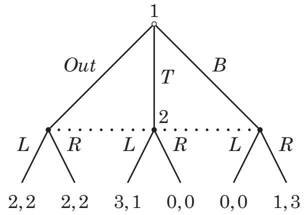


Figure 9.12 (Out, R) satisfies GW-forward induction

induction should not depend on the extensive-form representation. Govindan and Wilson (2009, Appendix B) offer such a definition of forward induction that only depend on the reduced normal form. However, we follow the variant of the definition suggested by Man (2012). That is, we strengthen weakly sequential equilibria to weakly quasi-perfect equilibria and provide a definition of forward induction based on perfect equilibria (see page 251).

Consider a reduced normal form and a function mapping pure strategy profiles to a finite set of alternatives Ω where players have their preferences defined. For instance, if it is the reduced normal form of an extensive-form game where each ending node leads to a different alternative then $\Omega = Z$. Consider a perfect equilibrium σ inducing a probability distribution ζ on Ω. The set of all probability distributions on Ω induced by perfect equilibria has finitely many connected components. Let $P(\zeta)$ be the component that contains ζ and let $\mathrm{PE}(\zeta)$ be the set of perfect equilibria that induces a probability distribution in $P(\zeta)$.

A pure strategy $s_i \in S_i$ of player i is *(first-order) relevant* for $\mathrm{PE}(\zeta)$ if there exists a sequence of ε-perfect equilibria $\{\sigma^t\}$ converging to a perfect equilibrium in $\mathrm{PE}(\zeta)$ such that for every player $i \in N$ and every $s_i' \in S_i$

$$u_i(\sigma_{-i}^t, s_i) \geq u_i(\sigma_{-i}^t, s_i').$$

Thus, this is analogous to the previous definition of relevant strategy, but adapted for perfect equilibrium. The next step is to consider relevant strategies as infinitely more likely than strategies that are not relevant. If, under such a condition, a perfect equilibrium inducing ζ can be sustained then it satisfies (first-order) forward induction:

Definition 15 A perfect equilibrium σ inducing outcome ζ satisfies (first-order) forward induction if there exists a sequence of ε-perfect equilibria $\{\sigma^t\}$ converging to σ such that for all players $i \in N$ and any two $s_i, s_i' \in S_i$, if s_i is (first-order) relevant for $\mathrm{PE}(\zeta)$ and s_i' is not then

$$\lim_{t\to\infty} \frac{\sigma_i(s_i')}{\sigma_i(s_i)} = 0. \tag{9.2}$$

Man (2012) provides an example where a first-order relevant strategy is never a best response against a forward induction equilibrium in $\mathrm{PE}(\zeta)$. Hence, there is room for an iterative definition. Call a strategy second-order relevant for $\mathrm{PE}(\zeta)$ if it is a best response to a sequence of ε-perfect equilibria satisfying (9.2) and converging to a, consequently, first-order forward induction equilibrium in $\mathrm{PE}(\zeta)$. Then a perfect equilibrium inducing outcome ζ satisfies second-order forward induction if it is the limit point of a sequence of ε-perfect equilibria

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

where second-order relevant strategies for PE(ζ) are infinitely more likely than first-order relevant strategies for PE(ζ), and first-order relevant strategies for PE(ζ) are infinitely more likely than strategies that are not first-order relevant for PE(ζ). Continuing in this fashion we may say that a perfect equilibrium σ inducing outcome ζ is a forward induction equilibrium if it survives this iterative process.

In signaling games, this version of forward induction implies the intuitive criterion, D1, and D2. Furthermore, Man (2012) also shows that every KM-stable set contained in a single connected component of equilibria contains a forward induction equilibrium. Thus, it is also the case that every Mertens stable set has a forward induction equilibrium as defined by Man (2012).

6.1 An Example

We present an example with an equilibrium that (vaguely) satisfies the definitions of forward induction of the previous section. However, the forward induction property of stable sets eliminates such an equilibrium. As we can see, we need some form of *collective forward induction* argument to reject the equilibrium.

Consider the extensive-form game Γ depicted in Figure 9.13. The strategy profile $\sigma = (Out, Ns, Lr)$ is a perfect and a proper equilibrium. The outcome $(3,3,3)$ is induced, e.g., by a sequential equilibrium in which player 3's beliefs are such that (W, N) is infinitely more likely than (E, n) and (E, s) is infinitely more likely than (W, S). Player 1's strategies W and E are irrelevant for that outcome, indeed, they are never a best response against any sequential equilibrium generating a payoff vector close to $(3,3,3)$. If, say, W was optimal against a sequential equilibrium generating a playoff for player 1 close to 3, then player 3 must be playing l with probability close to $3/4$. But if that is the case, player 2's unique sequentially rational best response is to play S after W and n after E. Additionally, for player 3 to be mixing and playing l with probability close to $3/4$ it must be the case that she believes that (W, S) and (E, s) are equally likely. These two last facts together with belief consistency imply that player 3 must put probability 1 on the profile (E, n) at her top information set and play R there. Hence, player 1 would be playing E and obtaining a payoff of 4.

This argument implies that the two information sets of player 3 are not relevant for the outcome $(3,3,3)$ and that forward induction as defined by either Govindan and Wilson (2009)

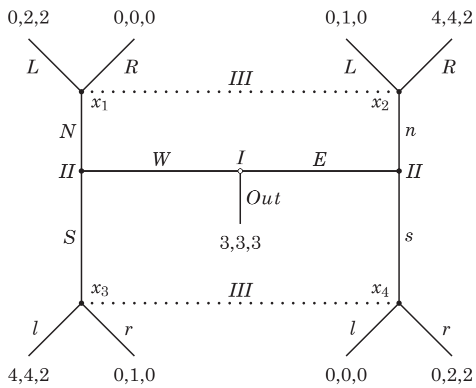


Figure 9.13 The unique stable outcome is '4,4,2'

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

or Man (2012) does not have a bite in this game. However, it seems that a typical forward induction argument is possible and that, under an equilibrium generating a payoff vector $(3,3,3)$ player 3 can only make sense of having to move, say, at her top information set if by concluding that players 1 and 2 played according to (E,n). As a matter of fact, the unique Mertens stable set of this game generates outcome $(4,4,2)$ and one sees that there is no stable set with payoffs $(3,3,3)$ by noticing that such a stable set would not contain a stable set of the game obtained by eliminating either W or E, i.e., it would violate the forward induction property of stable sets.

7 THE AXIOMATIC APPROACH

Stable sets (Mertens, 1989) satisfy desirable rationality requirements (admissibility, backward induction, forward induction, and iterated deletion of dominated strategies) as well as the most exhaustive list of invariance properties. The definition imposes conditions on how the graph of the Nash equilibrium correspondence of perturbed games must look close to the stable set of the game. We can ask whether this definition imposes more than what is warranted by the properties. One way to address this question is treating some desirable equilibrium properties as axioms and then looking on how those axioms shape the collection of equilibrium concepts that satisfy them. A complete axiomatic approach to strategic stability would identify what is implied by game-theoretical axioms, and whether there are any unintended consequences in the definition of stability.[20] We do not currently have such a complete theory available but only some partial contributions.

Even if the approach is not truly axiomatic, Govindan and Wilson (2008a) define *metastability* motivated by the following two axioms:

- *Invariance to embedding*: It is a combination of immunity to the addition of redundant strategies and the small worlds axiom. A game G can be trivially embedded in a larger game $\tilde{G}$ with a larger player set and larger strategy space in a way that the extra features in $\tilde{G}$ do not affect the strategic interaction as represented by G. Embedding requires that each selected set of equilibria of G *is* the projection of a selected set of equilibria of any larger game $\tilde{G}$ in which G is embedded.
- *Continuity*: A game whose best-response correspondence is nearby (as in the robustness condition of BR-sets; see Definition 9) must have a Nash equilibrium close to the selected set.

Taking into account these two axioms we can define an equilibrium concept as follows:

Definition 16 A connected set of the equilibria T of a game G is *metastable* if and only if every neighborhood of T contains the projection (into the set of strategy profiles of G) of a

[20] Balkenborg and Vermeulen (2015) study "diversity games" where players have to avoid choosing the same strategy. If they do they all get a payoff equal to 1, otherwise they all get a payoff equal to 0. All pure strategy equilibria of this game belong to the same component. Balkenborg and Vermeulen (2015) show that if every player has two strategies then such a component is a stable set if and only if the number of players is even. (If the number of players is odd the component does not even contain a BR-stable set.) While this alternating behavior seems counter intuitive, it may be related to backward induction and the small worlds axiom (see Balkenborg and Vermeulen, 2015, Footnote 21).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

fixed point of every sufficiently small perturbation of the best-reply correspondence of any larger game $\tilde{G}$ in which G is embedded.

Govindan and Wilson (2008a) show that this definition is equivalent to a definition of the same type as those definitions in Section 4. They have to satisfy a connectedness condition and an essentiality condition. The latter condition is somewhere in between homotopic essentiality and cohomologic essentiality: not only must the projection mapping be essential in homotopy, but must remain essential when the domain and range are expanded to take into account the larger games in which the game of interest can be embedded. A metastable set satisfies all the properties of stable sets but with only two differences in the small worlds axiom and composition property (cf. properties 1 and 2 in page 247):

1. The projection of a metastable set of the game obtained by trivially increasing the set of players *contains* a metastable set.
2. Given two independent games G and G' and a metastable set T of G, there exists a metastable set T' of G' such that $T \times T'$ is a metastable set of the game $G \times G'$.

A truly axiomatic approach is taken by Govindan and Wilson (2012). They strengthen the meaning of invariance to embedding and combine it with the admissibility and backward induction axioms. The meaning of admissibility is the usual: backward induction is taken to mean that a selected set of equilibria must contain a quasi-perfect equilibrium. The strengthened version of *invariance to embedding* has the small worlds axiom and immunity to the addition of redundant strategies as special cases. In particular, embedding a game in a larger game may introduce new players whose strategy choice may change the redundant pure strategies that the original players have available. Invariance to embedding requires that each solution of a game *is* the image of a solution of a bigger game in which the game is embedded. Formally, let G have player set N and each $i \in N$ a set of mixed strategies Σ_i. We embed G into a larger game $\tilde{G}$ with player set $N \cup N^o$, where each player $i \in N$ has mixed strategy set $\tilde{\Sigma}_i$, and where the set of mixed strategy profiles of players in N^o is $\tilde{\Sigma}^o$. Such an embedding is represented by means of a family of $f_i : \tilde{\Sigma}_i \times \tilde{\Sigma}_o \to \Sigma_i$ satisfying some conditions that guarantee that the extra features in $\tilde{G}$ do not affect the strategic interaction of players in N. Then, invariance to embedding implies that the solutions of G are the images under $f = (f_i)_{i \in N}$ of the solutions in $\tilde{G}$.

Govindan and Wilson (2012) show that robustness to strategy perturbations can be substituted by invariance to embedding in larger games. Based on this, they prove that, in generic two-player extensive-form games, any equilibrium concept that satisfies admissibility, backward induction and invariance to embedding must select Mertens' stable sets. Hence, in this class of games, stability is implied purely by decision-theoretic criteria rather than by properties inherited from perturbed games.

REFERENCES

Balkenborg, D. and Vermeulen, D. Where strategic and evolutionary stability depart – a study of minimal diversity games. *Mathematics of Operations Research*, 41(1):278–292, 2015.

Banks, J.S. and Sobel, J. Equilibrium selection in signaling games. *Econometrica*, 55(3): 647–661, 1987.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Battigalli, P. Strategic independence and perfect Bayesian equilibria. *Journal of Economic Theory*, 70(1): 201–234, 1996.
Blume, L. and Zame, W. The algebraic geometry of perfect and sequential equilibrium. *Econometrica*, 62(4):783–794, 1994.
Blume, L., Brandenburger, A., and Dekel, E. Lexicographic probabilities and equilibrium refinements. *Econometrica*, 59(1):81–98, 1991.
Brandenburger, A. and Friedenberg, A. Are admissibility and backward induction consistent? New York University, 2011.
Cho, I. and Kreps, D. Signaling games and stable equilibria. *Quarterly Journal of Economics*, 102(2): 179–221, 1987.
Cournot, A.-A. *Recherches sur les principes mathématiques de la théorie des richesses par Augustin Cournot*. Paris: Chez L. Hachette, 1838.
Dalkey, N. Equivalence of information patterns and essentially determinate games. In Kuhn, H. and Tucker, A. (eds), *Contributions to the Theory of Games, Vol. 2*, pp. 217–244. Princeton, NJ: Princeton, University Press, 1953.
De Sinopoli, F. Sophisticated voting and equilibrium refinements under plurality rule. *Social Choice and Welfare*, 17(4):655–672, 2000.
De Sinopoli, F. On the generic finiteness of equilibrium outcomes in plurality games. *Games and Economic Behavior*, 34(2):270–286, 2001.
De Sinopoli, F., Dutta, B., and Laslier, J.-F. Approval voting: Three examples. *International Journal of Game Theory*, 35(1):27–38, 2006.
Dekel, E. and Siniscalchi, M. *Epistemic game theory*. Amsterdam: North-Holland, 2014.
Elmes, S. and Reny, P.J. On the strategic equivalence of extensive-form games. *Journal of Economic Theory*, 62(1):1–23, 1994.
Fort, M.K. Points of continuity of semicontinuous functions. *Publicationes Mathematicae Debrecen*, 2 (1951):100–102, 1951.
Glazer, J. and Weiss, A. Pricing and coordination: Strategically stable equilibria. *Games and Economic Behavior*, 2(2):118–128, 1990.
Govindan, S. Stability and the chain store paradox. *Journal of Economic Theory*, 66(2):536–547, 1995.
Govindan, S. A subgame property of stable equilibria. *Mathematics of Operations Research*, 21(4): 991–999, 1996.
Govindan, S. and Mertens, J.-F. An equivalent definition of stable equilibria. *International Journal of Game Theory*, 32(3):339–357, 2004.
Govindan, S. and Wilson, R. Direct proofs of generic finiteness of Nash equilibrium outcomes. *Econometrica*, 69(3):765–769, 2001.
Govindan, S. and Wilson, R. Sufficient conditions for stable equilibria. *Theoretical Economics*, 1(2): 167–206, 2006.
Govindan, S. and Wilson, R. Metastable equilibria. *Mathematics of Operations Research*, 33(4): 787–820, 2008a.
Govindan, S. and Wilson, R. Refinements of Nash equilibrium. In Durlauf, S.N. and Blume, L.E. (eds), *The New Palgrave Dictionary of Economics, Vol. 6*, 2nd edition. Basingstoke, UK: Palgrave Macmillan, 2008b.
Govindan, S. and Wilson, R. On forward induction. *Econometrica*, 77(1):1–28, 2009.
Govindan, S. and Wilson, R. Axiomatic equilibrium selection for generic two-player games. *Econometrica*, 80(4):1639–1699, 2012.
Grigis de Stefano, F. Strategic stability of equilibria: The missing paragraph. *CORE Discussion Papers* 2014015, Université Catholique de Louvain, Center for Operations Research and Econometrics (CORE), 2014.
Harsanyi, J.C. Oddness of the number of equilibrium points: A new proof. *International Journal of Game Theory*, 2(1):235–250, 1973.
Hatcher, A. *Algebraic Topology*. Cambridge, UK: Cambridge University Press, 2002.
Hillas, J. On the definition of the strategic stability of equilibria. *Econometrica*, 58(6):1365–1390, 1990.
Hillas, J. How much of forward induction is implied by backward induction and ordinality? Mimeo. University of Auckland, New Zealand, 1998a.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Hillas, J. A game illustrating some features of the definition of strategic stability, unpublished, 1998b.

Hillas, J. and Kohlberg, E. Foundations of strategic equilibria. In Aumann, R.J. and Hart, S. (eds), *Handbook of Game Theory with Economic Applications*, Vol. 3, pp. 1597–1663. Amsterdam: North-Holland, 2002.

Hillas, J., Jansen, M., Potters, J., and Vermeulen, D. On the relation among some definitions of strategic stability. *Mathematics of Operations Research*, 26(3):611–635, 2001.

Hillas, J., Kao, T., and Schiff, A. A semi-algebraic proof of the generic equivalence of quasi-perfect and sequential equilibria. Mimeo, University of Auckland, 2002.

Kakutani, S. *A Generalization of Brouwer's Fixed Point Theorem*. Durham, NC: Duke University Press, 1941.

Kalai, E. and Samet, D. Persistent equilibria in strategic games. *International Journal of Game Theory*, 13(3):129–144, 1984.

Kinoshita, S. On essential components of the set of fixed points. *Osaka Mathematical Journal*, 4(1): 19–22, 1952.

Kohlberg, E. Refinement of Nash equilibrium: The main ideas. In Ichiishi, T., Neyman, A., and Tauman, Y. (eds), *Game Theory and Applications*, pp. 3–45. San Diego, CA: Academic Press, 1990.

Kohlberg, E. and Mertens, J.-F. On the strategic stability of equilibria. *Econometrica*, 54(5):1003–1037, 1986.

Kohlberg, E. and Reny, P. Independence on relative probability spaces and consistent assessments in game trees. *Journal of Economic Theory*, 75(2):280–313, 1997.

Kojima, M., Okada, A., and Shindoh, S. Strongly stable equilibrium points of n-person noncooperative games. *Mathematics of Operations Research*, 10(4):650–663, 1985.

Kreps, D. and Wilson, R. Sequential equilibria. *Econometrica*, 50(4):863–894, 1982a.

Kreps, D. and Wilson, R. Reputation and imperfect information. *Journal of Economic Theory*, 27(2):253–279, 1982b.

Kuhn, H. Extensive games and the problem of information. *Annals of Mathematical Studies*, 28: 193–216, 1953.

Luce, R.D. and Raiffa, H. *Games and Decisions: Introduction and Critical Survey*. New York: Courier Dover Publications, 1957.

Man, P.T. Forward induction equilibrium. *Games and Economic Behavior*, 75(1):265–276, 2012.

McLennan, A. *Advanced Fixed Point Theory for Economics*, Mimeo, 2012.

Mertens, J.-F. Stable equilibria – a reformulation, part I: Definition and basic properties. *Mathematics of Operations Research*, 14(4):575–625, 1989.

Mertens, J.-F. Stable equilibria – a reformulation, part II: Discussion of the definition and further results. *Mathematics of Operations Research*, 16(4):694–753, 1991.

Mertens, J.-F. The small worlds axiom for stable equilibria. *Games and Economic Behavior*, 4(4): 553–564, 1992.

Mertens, J.-F. Two examples of strategic equilibrium. *Games and Economic Behavior*, 8(2):378–388, 1995.

Mertens, J.-F. Ordinality in non cooperative games. *International Journal of Game Theory*, 32(3): 387–430, 2003.

Myerson, R. Refinements of the Nash equilibrium concept. *International Journal of Game Theory*, 7(2): 73–80, 1978.

Myerson, R. and Weibull, J. Tenable strategy blocks and settled equilibria. *Econometrica*, 83(3): 943–976, 2015.

Nash, J.F. Equilibrium points in n-person games. *Proceedings of the National Academy of Science*, 36(1):48–49, 1950.

Nash, J.F. Non-cooperative games. *The Annals of Mathematics*, 54(2):286–295, 1951.

Okada, A. On stability of perfect equilibrium points. *International Journal of Game Theory*, 10(2): 67–73, 1981.

Osborne, M.J. Signaling, forward induction, and stability in finitely repeated games. *Journal of Economic Theory*, 50(1):22–36, 1990.

Pimienta, C. and Shen, J. On the equivalence between (quasi-) perfect and sequential equilibria. *International Journal of Game Theory*, 1–8, 2010.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Reny, P.J. Backward induction, normal form perfection and explicable equilibria. *Econometrica*, 60(3): 627–649, 1992.
Ritzberger, K. The theory of normal-form games from the differentiable viewpoint. *International Journal of Game Theory*, 23(3):207–236, 1994.
Samuelson, L. Evolution and game theory. *The Journal of Economic Perspectives*, 16(2):47–66, 2002.
Selten, R. Spieltheoretische behandlung eines oligopolmodells mit nachfragetraheit. *Zeitschrift für die gesamte Staatswissenschaft (JITE)*, 12:301–324, 1965.
Selten, R. Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, 4(1):24–55, 1975.
Swinkels, J.M. Independence for conditional probability systems. *Discussion Papers* 1076, Northwestern University, Center for Mathematical Studies in Economics and Management Science, 1993.
Thompson, F.B. Equivalence of games in extensive form. *RAND Research Memorandum*, 759, 1952.
Van Damme, E. A relation between perfect equilibria in extensive-form games and proper equilibria in normal-form games. *International Journal of Game Theory*, 13(1):1–13, 1984.
Van Damme, E. Stable equilibria and forward induction. *Journal of Economic Theory*, 48(2):476–496, 1989.
Van Damme, E. Strategic equilibrium. In Aumann, R.J. and Hart, S., *Handbook of Game Theory with Economic Applications, Vol. 3*, pp. 1521–1596. Amsterdam: North-Holland, 2002.
Vermeulen, A. and Jansen, M. Are strictly perfect equilibria proper? A counterexample. *Journal of Optimization Theory and Applications*, 90(1):225–230, 1996.
Vermeulen, A. and Jansen, M.J. An ordinal selection of stable sets in the sense of Hillas. *Journal of Mathematical Economics*, 36(2):161–167, 2001.
Vermeulen, A.J., Potters, J.A.M., and Jansen, M.J.M. On stable sets of equilibria. In Parthasarathy, T., Dutta, B., Potters, J.A.M. Raghaven, T.E.S. Ray D., and Sen, A., *Game Theoretical Applications to Economics and Operations Research*, pp. 133–148. New York: Springer, 1997.
Zermelo, E. Über eine anwendung der mengenlehre aug die theorie des schachspiels. *Proceedings of the 5th Congress of Mathematicians*, 2:501–504, 1912.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

10. Stackelberg games*

Ludovic A. Julien

1 INTRODUCTION

The Stackelberg market structure portrays a landscape where strategic interactions fall within hierarchical competition. Stackelberg's book *Marktform und Gleichgewicht* (1934) is probably alongside Cournot's monograph (1838) one of the most creative and innovative books devoted to the functioning of markets under strategic interactions. The noncooperative duopoly model was developed widely, leading to many game-theoretic models. As an extension of the simultaneous move games, it was notably at the origin of some refinement for the Nash equilibrium concept, namely the subgame perfect Nash equilibrium (Selten, 1975). It also contributed to operation research insofar as some algorithms were developed to determine strategic sequential equilibria (Murphy, Sherali, and Soyster, 1983). Finally, it applied in many areas of economics like industrial organization (Carlton and Perloff, 2004), international economics (Brander and Spencer, 1983), and macroeconomics (Kydland and Prescott, 1977). In this chapter devoted to Stackelberg games, the motivations are threefold. First, we aim at deepening the logic of Stackelberg competition to capture some salient features of strategic interactions in hierarchical markets. Second, we study some extensions and generalizations of the basic Stackelberg duopoly game. Third, we compare the performance of the Stackelberg market structure(s) with the performance of the Cournot and the competitive market structures.

The standard Stackelberg model depicts a market in which two strategic firms move sequentially. By introducing some hierarchy among firms in the Cournot duopoly model, Stackelberg competition puts forward a second kind of asymmetry. As in the Cournot duopoly market, there is an asymmetry between the demand side, which embodies a large number of price-taking consumers, and the supply side, which includes strategic firms whose behavior is noncooperative. The second asymmetry consists of assuming that strategic firms do not behave in the same way. Thus, the strategic interactions are cast in a sequential model of decisions. Indeed, in the basic duopoly model one firm (the leader) moves first and makes its (her) decision by taking into consideration the reaction of the second-mover firm (the follower) it (she) perfectly knows. The difference between these two kinds of strategic behavior does not stem primarily from the fundamentals, but is merely linked to the noncooperative nature of the game itself. A Stackelberg equilibrium is a noncooperative equilibrium of a two-stage game with complete and perfect information in which the players are the firms, the strategies are their supply decisions, and the payoffs are their profit functions. Therefore, a Stackelberg equilibrium constitutes a pure strategy subgame perfect

* Some parts of this chapter have been presented in seminars held at the University of Udine and at the University of Paris West-Nanterre between the years 2013 and 2015. I am grateful to Giulio Codognato, Stefano Comino, Alex Dickson, Clara Grazziano, Marco Marini, Louis de Mesnard, Fabrice Tricou, and Simone Tonin for their helpful comments, remarks and suggestions.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Nash equilibrium of the two-stage game (Fudenberg and Tirole, 1990). In this chapter, we explore the logic of Stackelberg competition, and we consider a framework, which enables us to widen and generalize the basic game.

This chapter devoted to Stackelberg games does not aspire to exhaustiveness. It rather aims at focusing on the main properties of deterministic and static quantity-setting games, which feature capacity competition in time.[1] It also aims at exploring some of their possible (new) extensions that would be of interest for industrial organization. To this end, we focus on Stackelberg games with (one) homogeneous product(s). In addition, we consider games with complete and perfect information. The reasons are twofold. First, we start from a basic model general enough to allow simple extensions and generalizations. Second, the equilibrium properties of these games do not depend upon special functional forms assumed for costs and market demand. Our starting point is the duopoly model we progressively enrich. Then, we consider strategic interactions between several leaders and followers. Finally, we study a two-commodity multiple leader–follower bilateral market game. This transposition notably tests the limits of some results that hold in industry games. Games with differentiated products are dealt with in Appendix A for comparison. For each game, we discuss the main assumptions, define the corresponding market equilibrium concept, and characterize the optimality conditions. Then, we focus on market power and on the welfare implications of imperfectly competitive behavior. More specifically, we provide welfare comparisons between the various Stackelberg equilibria; and, for each Stackelberg equilibrium concept, we study its relationship with the Cournot and the competitive equilibria. The endogenization of the order of moves, mergers and free entry are also dealt with.

The remainder of this chapter is organized as follows. Section 2 studies the basic Stackelberg duopoly game. Section 3 deals with the multiple leader–follower game. Section 4 is devoted to the bilateral oligopoly multiple leader–follower market game. Section 5 concludes. Appendices A and B complete the story.

2 THE BASIC LEADER–FOLLOWER GAME

The basic static and deterministic leader–follower quantity game is a natural starting point from which to study some critical features of Stackelberg competition. First, we describe the duopoly model. Second, we study the Stackelberg duopoly equilibrium (SDE thereafter). Third, we consider the market power and the welfare properties of the SDE. To this end, we compare the SDE with the Cournot duopoly equilibrium (CDE) and with the competitive equilibrium (CE). Fourth, as the timing of moves is given we consider the endogenization of the order of moves. Fifth, we study free entry, and notably the convergence of the extended SDE to the CE.

[1] Indeed, the price-setting models may generate dual results. For instance, in the basic duopoly price leadership game with linear demand and identical quadratic costs, the follower behaves as a price taker and achieves higher payoffs than the leader. In addition, in the symmetric duopoly price leadership game with differentiated products, the leader sets a higher price than the follower whether the goods are substitutes or complements; and, the leader achieves lower (higher) payoffs than the follower when the goods are substitutes (complements). See Appendix A.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

2.1 The Model

Consider a market with one single divisible homogeneous product. There are two risk-neutral firms, namely L (the leader) and F (the follower) who compete strategically on quantities. Firm L (resp. F) produces x_L (resp. x_F) units of the good. Each firm bears some costs. Let $C_L : \mathbb{R}_+ \to \mathbb{R}_+$, with $x_L \to C_L(x_L)$ be the total cost function of the leader. Similarly, we define $C_F : \mathbb{R}_+ \to \mathbb{R}_+$, with $x_F \to C_F(x_F)$. We assume the costs functions are twice-continuously differentiable, increasing and convex. Both firms face a large number of consumers who behave as price takers. Thus, there is a continuous and decreasing market demand function, namely $D(p)$, where p is the unit price of the good expressed in a numéraire. Indeed, let $X \longmapsto p(X) = D^{-1}(X)$ be the market inverse demand function. This inverse demand represents the maximum price the consumers are willing to pay to buy the quantity X as well as the minimum price the firms are willing to charge to sell the corresponding production X. There is a market-clearing condition that stipulates that the demand balances the aggregate supply X, with $X \equiv x_L + x_F$. We assume the price function is twice-continuously differentiable and decreasing. In addition, it satisfies $\frac{dp(X)}{dX} + kx\frac{d^2p(X)}{(dX)^2} \leqslant 0$, where $k > 0$. This assumption says that marginal revenue for any single firm is a decreasing function of total industry output.[2]

The profit functions $\pi_L(.)$ and $\pi_F(.)$ may be written:

$$\pi_L(x_L, x_F) = p(X)x_L - C_L(x_L) \tag{10.1}$$

$$\pi_F(x_L, x_F) = p(X)x_F - C_F(x_F). \tag{10.2}$$

Let us notice that, under the assumptions made, (10.2) is strictly concave in x_F. The strict concavity of (10.1) is studied in subsection 2.2.

We associate now a noncooperative game $\mathbf{\Gamma}$. The players are the firms, the strategies are their supply decisions, and the payoffs are their profits. Let $\mathcal{S}_L = [0, \infty)$ be the strategy set of the leader, where the supply x_L represents the pure strategy of the leader. Correspondingly, let $\mathcal{S}_F = [0, \infty)$,where x_F is the pure strategy of the follower. A strategy profile will be represented by the vector (x_L, x_F). This sequential game displays two stages of decision and no discounting. We also assume that the timing of positions is given. First, the leader chooses a quantity to sell, and then the follower determines his supply on the residual demand. The leader cannot revise her decision: her supply is irreversible. In addition, firms meet once and cannot make binding agreements. Finally, information is assumed to be complete and perfect. It notably implies that the leader has perfect knowledge of the follower's reactions. It also implies that, for the follower, each information set is a single decision node. In addition, in each decision node, the follower makes an optimal choice, so sequential rationality

[2] This formulation deserves two comments. First, we do not impose $\frac{d^2p(X)}{(dX)^2} \leqslant 0$, so our assumption does not preclude (strictly) convex market demand functions. But it precludes (strong) strategic complementarities. Second, our formulation of the decreasing marginal revenue hypothesis embodies some positive parameter, namely k. For any follower firm, we have $k = 1$, as in the Cournot model (Hahn, 1962, Okugushi, 1976). Let us notice, however, that for any leader firm, the parameter k may be different from unity unless the leader behaves as a follower. This formulation puts forward an important feature of Stackelberg competition, which explicitly takes into account the leader's beliefs. Indeed, the leader has perfect information regarding the optimal reaction of the follower to a change in the leader's strategy (see notably Julien, 2011).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

prevails. As sequential rationality is common knowledge, the game $\mathbf{\Gamma}$ is solved by backward induction, considering first the optimal strategic reactive decision of the follower, and then the equilibrium strategic choice of the leader.

2.2 Stackelberg Duopoly Equilibrium

First, given any feasible strategy profile x_L, firm F determines its optimal decision as a mapping that depends upon the leader's strategy. Let $\varphi : \mathcal{S}_L \to \mathcal{S}_F$, with $x_L \mapsto \varphi(x_L)$ the follower's best response, which is the solution to:

$$\varphi(x_L) \in \arg\max \pi_F(x_L, x_F), \forall x_L \in \mathcal{S}_L. \tag{10.3}$$

The first-order sufficient condition, which is given by $\frac{\partial \pi_F}{\partial x_F} = 0$, may be written:

$$p\,(X) + x_F \frac{dp\,(X)}{dX} - \frac{dC_F(x_F)}{dx_F} = 0. \tag{10.4}$$

Expression (10.4) is the standard optimality condition for a firm that behaves á la Cournot. By the strict concavity of the profit function (10.1), the reaction function $\varphi(x_L)$, solution to (10.3), is unique. We can deduce:

$$\frac{\partial \varphi(x_L)}{\partial x_L} = -\frac{\frac{\partial^2 \pi_F}{\partial x_F \partial x_L}}{\frac{\partial^2 \pi_F}{\partial (x_F)^2}} = -\frac{\frac{dp\,(X)}{dX} + x_F \frac{d^2 p\,(X)}{(dX)^2}}{2\frac{dp\,(X)}{dX} + x_F \frac{d^2 p\,(X)}{(dX)^2} - \frac{d^2 C_F(x_F)}{(dx_F)^2}}. \tag{10.5}$$

Then $\frac{\partial \varphi}{\partial x_L} \in [-1, 0]$ as $\frac{dp\,(X)}{dX} + x_F \frac{d^2 p\,(X)}{(dX)^2} \leqslant 0$, so the best response is nonincreasing. Thus, the game displays strategic substitutabilities: the strategies of the follower and the leader are substitutes. But, if $\frac{dp\,(X)}{dX} + x_F \frac{d^2 p\,(X)}{(dX)^2} > 0$, then the game exhibits strategic complementarities: the best response is upward sloping. It may happen when the demand function is not linear and/or with nonconvex costs functions (see Amir and Grilo, 1999, Vives, 1999).

In the first stage, the leader knows, via the effective demand that is addressed to her, i.e., $p\,(x_L + \varphi(x_L))$, how the market price is affected by the follower's reaction. Therefore, the leader solves:

$$x_L \in \arg\max\ \pi_L(x_L, \varphi(x_L)). \tag{10.6}$$

The first-order condition $\frac{\partial \pi_L}{\partial x_L} = 0$ may be written:

$$p\,(X) + (1+\nu) x_L \frac{dp\,(X)}{dX} - \frac{dC_L(x_L)}{dx_L} = 0. \tag{10.7}$$

This expression is rather different from (10.4) as it includes the term ν, where $\nu = \frac{\partial \varphi(x_L)}{\partial x_L}$, with $\nu \in [-1, 0]$. It represents the reaction of the follower to the *leader*'s strategy, that is, the slope of the *best response* evaluated in equilibrium. Indeed, if $\nu = -1$, then $p\,(X) = \frac{dC(x_L)}{dx_L}$ (which can also be deduced from (10.5)): the leader and his follower behave as

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

price-takers. The leader expects its rival to compensate exactly any expansion or reduction to its own supply, leaving the market price unchanged (see Dixit, 1986, Jeffrey, Karp, and Perloff, 2007). If $\nu = 0$, i.e., the leader makes no expectation regarding the slope of the *best response*, which means that firm L behaves as a Cournot competitor, i.e., $p(X) + x_L \frac{dp(X)}{dX} - \frac{dC_L(x_L)}{dx_L} = 0$.[3] Otherwise, if $\nu \in (-1, 0)$, then firm L behaves as a leader. The second-order sufficient condition, i.e., $\frac{\partial^2 \pi_L}{(\partial x_L)^2} < 0$, with $k = 1 + \nu$, may be written:

$$k\left(kx_L \frac{d^2 p(X)}{(dX)^2} + 2\frac{dp(X)}{dX}\right) - \frac{d^2 C_L(x_L)}{(dx_L)^2} < 0. \tag{10.8}$$

This inequality is true as $\frac{dp(X)}{dX} + kx_L \frac{d^2 p(X)}{(dX)^2} \leqslant 0$. Equation (10.7) determines the leader's equilibrium strategy, namely $\tilde{x}_L$, which by using $\varphi(x_L)$, yields the follower's equilibrium strategy $\tilde{x}_F = \varphi(\tilde{x}_L)$. Then, given the strategy profile $(\tilde{x}_F, \tilde{x}_L)$, we deduce the market price $\tilde{p}$. The payoffs $(\tilde{\pi}_L, \tilde{\pi}_F)$ follow from (10.1)–(10.2).

We now provide a formal definition of a Stackelberg duopoly equilibrium (SDE):

Definition 1 (SDE) A Stackelberg duopoly equilibrium of $\mathbf{\Gamma}$ is given by a strategy profile $(\tilde{x}_L, \tilde{x}_F) \in \mathcal{S}_L \times \mathcal{S}_F$ such that:

a. $\pi_F(\tilde{x}_L, \tilde{x}_F) \geqslant \pi_F(\tilde{x}_L, x_F)$, $\forall x_F \in \mathcal{S}_F$, $x_F = \varphi(x_L)$;
b. $\pi_L(\tilde{x}_L, \varphi(\tilde{x}_L)) \geqslant \pi_L(x_L, \varphi(x_L))$, $\varphi(x_L) \in \mathcal{S}_F$ and $\forall x_L \in \mathcal{S}_L$.

At an SDE, each firm behaves optimally given its beliefs about its rival, and the firms' beliefs are fulfilled for these strategies. An SDE is thus a noncooperative equilibrium of a two-stage game with complete and perfect information; i.e., a pure strategy subgame perfect Nash equilibrium of $\mathbf{\Gamma}$. It requires the strategies of the leader and of the follower to constitute a Nash equilibrium of any subgame. Indeed, neither the leader firm nor the follower firm want to unilaterally deviate from their choice based on their conjectures, and the choices they make are consistent with these conjectures. In addition, it is a subgame perfect Nash equilibrium without empty threats: it rules out noncredible threats by the follower. The reason is that the strategy of the follower is optimal for any supply of the leader. The follower can set their own supply according to any possible function of the quantity set by the leader, with the belief that the leader will not counter-react. Similarly, the leader expects the follower to conform to the decisions given by his best-response function. For the leader, the only requirement it imposes on her strategy is that it does not generate losses, i.e., any supply x_L within $[0, p(0) - c_L(x_L)]$ is sustainable as a Nash equilibrium of the two-stage game with complete and perfect information. Let us notice that the leader's payoff in a SDE is always higher than her payoff in the CDE as she knows that her rival behaves as a Cournot firm.

[3] Let us notice that there is no such best response under Cournot competition. No firm can observe the choice of any other firm since both firms play a simultaneous-move game. For instance, the term "optimal decision mapping" is more appropriate in this case (see notably Johansen, 1982 and Daughety, 2009). Therefore, the fact that there is a zero slope conjecture is not inconsistent with the fact that the slope of the optimal decision mapping of any firm differs from zero since there is *no* reaction of any firm to possible deviations of the other firm.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

What can be said about existence and uniqueness of an SDE? We know that (10.3) has a unique solution: the mapping $\varphi(x_L)$ is a well-defined function from the strict concavity of (10.1). Indeed, the objective function (10.1) is a continuous function, and it is defined over a compact(ified) strategy set. So, the mapping $\varphi(x_L)$ exists and is point-valued. In addition, as the function (10.2) is also strictly concave in x_L, the solution to (10.6) is unique and given by the solution to equation (10.7), that is $\tilde{x}_L$. Then, we deduce $\tilde{x}_F = \varphi(\tilde{x}_L)$. Existence and uniqueness of SDE are notably studied by Leitmann (1978) in a nonzero-sum two-player game, and in Alj, Breton, and Haurie (1988), and by Freiling, Jank, and Lee (2001) within differential games.

2.3 Market Power and Welfare

We now consider market power, and the welfare implications of strategic interactions. More specifically, we compare the SDE market outcome with the CDE and the CE ones. Then, we study the relationships between market concentration and social surplus, and between market power and payoffs.

2.3.1 The SDE, CDE and CE aggregate market outcomes

Between the two polar sides of market structures, namely monopoly and perfect competition, there are several oligopoly market structures, among which are the Cournot and the Stackelberg duopoly competitions. The SDE aggregate market outcome is located somewhere between the CDE and the CE market outcomes. Proposition 1 compares market outcomes, and Proposition 2 compares SDE and CDE strategies.

Proposition 1 *The SDE aggregate supply (market price) is higher (lower) than the CDE supply (market price). But the SDE aggregate supply (market price) is lower (higher) than the CE aggregate supply (market price).*

The difference between the SDE and the CDE market outcomes can be explained as follows. The Cournot market structure features a simultaneous-move game. Therefore, firm L makes its decision by (correctly) expecting that the supply of firm F is independent of its own decision (and conversely). On the contrary, the sequential noncooperative game implies that the leader, who moves first, makes its decision by taking into consideration the reactions of the follower that the leader perfectly knows: the leader knows the reactions of the follower to a change of the leader's strategy, i.e., that the follower behaves as a Cournot firm. Then, the leader can set a higher supply than the supply corresponding to a Cournot behavior. In addition, the increment in the leader's supply more than compensates the decrease of the follower's supply when the best response is negatively sloped (above -1), whilst it goes in the same direction when the best response increases, i.e., when strategies are complements. Therefore, the aggregate output (market price) is higher (lower) in the SDE than in the CDE either when strategies are substitutes or when they are complements.

Proposition 2 *If* $\nu = 0$*, then the SDE coincides with the CDE.*

Proof Immediate, by comparing (10.7) with (10.4) when $\nu = 0$. ■

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

So, when the follower's best response has a zero slope, the leader rationally expects that each strategic decision she should undertake would entail no reaction from the follower. Each firm believes the others behave à la Cournot; thus all firms behave as if they played a simultaneous move game. This result holds when the market demand is isoelastic (Colombo and Labrecciosa, 2008).

The preceding propositions have welfare implications, which we now explore.

2.3.2 Market concentration and social surplus

Welfare is measured in numéraire units with the concept of surplus. Let S be the social surplus, which may be defined as a function of X:

$$\mathbf{S}(X) := \int_0^X p(z)dz - [C_L(s_L X) + C_F(s_F X)], \text{ with } X \leqslant X^*, \tag{10.9}$$

where $s_L \equiv \frac{x_L}{X}$ and $s_F \equiv \frac{x_F}{X}$ are the market shares. Differentiating partially with respect to X and decomposing $p(X)$ lead to $\frac{\partial \mathbf{S}(X)}{\partial X} = s_L[p(X) - \frac{dC_L(x_L)}{dX}] + s_F[p(X) - \frac{dC_F(x_F)}{dX}] \geqslant 0$ for fixed s_L and s_F, with $\frac{\partial \mathbf{S}(X)}{\partial X}_{|X=X^*} = 0$. Therefore, the social surplus increases with the aggregate supply: it is higher at the SDE than at the CDE, and highest at the CE. The consumer's surplus is $S_C(X) := \int_0^X p(z)dz - p(X)X$, with $\frac{\partial S_C(X)}{\partial X} = -X\frac{dp(X)}{dX} > 0$. In addition, if we let $S_\Pi(X) := p(X)(s_L + s_F)X - C_L(s_L X) - C_F(s_F X)$, then $\frac{dS_\Pi(X)}{dX} = p(X) + X\frac{dp(X)}{dX} - [s_L\frac{dC_L(s_L X)}{dX} + s_F\frac{dC_F(s_F X)}{dX}] < 0$ (costs are convex). Let us notice that the surplus of firm L is higher in the SDE than in the CDE.

The fact that the firm's aggregate surplus decreases with the aggregate output may be linked to market concentration. One interesting feature of Stackelberg competition is that the concentration is higher than under Cournot competition. Consider the Herfindahl index, which we denote by H, and that is defined as:

$$H(x_L, x_F) := (s_L)^2 + (s_F)^2 = \left(\frac{x_L}{x_L + x_F}\right)^2 + \left(\frac{x_F}{x_L + x_F}\right)^2. \tag{10.10}$$

In the linear model with identical maginal costs, market concentration is highest at the SDE. In addition, social surplus (resp. firm's average profit) and concentration are negatively (resp. positively) correlated. We now investigate the link between market power and individual welfare.

2.3.3 Market power and payoffs

We compare the SDE and CDE payoffs. To this end, we write (10.4) and (10.7) as:

$$p(X) = (1 + \kappa_F)\frac{dC_F(x_F)}{dx_F}, \text{ with } \kappa_F = \frac{1}{1 + \frac{1}{\epsilon}s_F} - 1 \tag{10.11}$$

$$p(X) = (1 + \kappa_L)\frac{dC_L(x_L)}{dx_L}, \text{ with } \kappa_L = \frac{1}{1 + \frac{1+\nu}{\epsilon}s_L} - 1, \tag{10.12}$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

where κ_F and κ_L represent the markups of the follower and the leader, and $\epsilon = \frac{dp(X)}{dX}\frac{p(X)}{X}$ is the price elasticity of market demand. We have $\kappa_F \in [0,1)$ and $\kappa_L \in [0,1)$, where $\kappa_F = \kappa_L = 0$ (resp. $0 < \kappa_F, \kappa_L < 1$) corresponds to CE (resp. SDE) behavior. Thus, $\kappa_L \gtreqless \kappa_F$ when $(1+\nu)s_L \gtreqless s_F$. To analyze the relation between market power and payoffs, it is instructive to rewrite (10.11) and (10.12) as:

$$L_F = \frac{\kappa_F}{\kappa_F + 1} = -\frac{1}{\epsilon}s_F \tag{10.13}$$

$$L_L = \frac{\kappa_L}{\kappa_L + 1} = -\frac{1+\nu}{\epsilon}s_L, \tag{10.14}$$

where $L_F \equiv \frac{p(X) - \frac{dC_F(x_F)}{dx_F}}{p(X)}$ and $L_L \equiv \frac{p(X) - \frac{dC_L(x_L)}{dx_L}}{p(X)}$ are the Lerner indexes.

Remark 1 The Lerner index of the leader may be written as $L_L = -\frac{1}{\xi}s_L$, where ξ is the price elasticity of the effective demand on the basis of which the leader might behave as if she were a monopolist.[4] Likewise, for the follower, we have $L_F = -\frac{1}{\zeta}s_L$, where ζ is the price elasticity of the residual demand.

Proposition 3 *If $L_L > L_F$, then $\tilde{\pi}_L > \tilde{\pi}_F$. In addition, assume $L_L = L_F$. Then $\tilde{\pi}_L \gtreqless \tilde{\pi}_F$ if and only if $\nu \lesseqgtr 0$.*

Proof Immediate from (10.13) and (10.14). ■

First, when the leader's cost is lower than the follower's cost, then the leader has more market power and achieves higher payoffs. Note the first part of the property does not preclude the case for which $\tilde{\pi}_L > \tilde{\pi}_F$ with $L_L < L_F$, when $\frac{dC_L(x_L)}{dx_L} > \frac{dC_F(x_F)}{dx_F}$. But it excludes the case for which $\tilde{\pi}_L < \tilde{\pi}_F$ with $L_L > L_F$. The following example illustrates this:

Example 1 Let $p(X) = a - bX$, $a, b > 0$, and $C_L(x_L) = (c+\theta)x_L$, $\theta > 0$, and $C_F(x_F) = cx_F$. Then, the SDE is given by $(\tilde{x}_L, \tilde{x}_F) = \left(\frac{a-2-2\theta}{2b}, \frac{a-c+2\theta}{4b}\right)$, $\tilde{p} = \frac{a+3c-2\theta}{4}$, and $(\tilde{\pi}_L, \tilde{\pi}_F) = \left(\frac{a-c-2\theta}{4b}\right)\left(\frac{a-c-6\theta}{2}, \frac{a-c-2\theta}{4}\right)$. Then, we have $\tilde{\pi}_L > \tilde{\pi}_F$ whenever $\theta < \frac{a-c}{10}$ but $L_1 < L_2$, where $(L_1, L_2) = \left(\frac{a-c-2\theta}{a+3c-2\theta}, \frac{a-c+2\theta}{a+3c-2\theta}\right)$ as $\nu = -\frac{1}{2}$ and $\epsilon = -\frac{a+3c-2\theta}{3(a-c)-2\theta}$.

Second, when $L_1 = L_2$, we have $C_L(x_L) = C_F(x_F)$. The leader achieves higher payoffs only if the best response is decreasing (a result that holds in the case of constant marginal costs). Indeed, the next example illustrates that the leader may have higher market power than the follower even if the two Lerner indexes are equal:

[4] A leader behaves as a monopolist (on the effective demand) when market demand is linear and firms bear the same constant marginal costs.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Example 2 Let $p(X) = a - bX$, $a, b > 0$, $C_L(x_L) = cx_L$ and $C_F(x_F) = cx_F$, with $c < a$. The SDE market shares and market price are given by $(\tilde{s}_L, \tilde{s}_F) = \left(\frac{2}{3}, \frac{1}{3}\right)$ and $\tilde{p} = \frac{a+3c}{4}$. Then, $\tilde{L}_L = \tilde{L}_F = \frac{a-c}{a+3c}$ as $\nu = -\frac{1}{2}$ and $\epsilon = -\frac{1}{3}\frac{a+3c}{a-c}$. In addition, the inverse of the effective demand is given by $p(X) = \frac{a+c}{2} - \frac{b}{2}x_1$. Then, $\tilde{L}_L = \frac{a-c}{a+3c}$ as $\tilde{s}_L = 1$ and $\xi = -\frac{a+3c}{a-c}$. Finally, the inverse of the residual demand is given by $p(X) - x_1 = \frac{a+c}{2} - bx_2$. Then, $\tilde{L}_F = \frac{a-c}{a+3c}$ as $\tilde{s}_F = 1$ and $\zeta = -\frac{a+3c}{a-c}$.

More generally, in case of strategic substituabilities, i.e., $\nu < 0$, the leader can decrease the supply of the follower by increasing his supply. The leader exploits the fact that the second-mover behaves with Cournot beliefs (Gal-Or, 1985, Vives, 1999, and Julien, 2011). When the best response has a zero slope in equilibrium, firm L behaves as a follower and both firms reach the Cournot payoffs (Colombo and Labrecciosa, 2008). The leader has lower payoffs than the follower when the best response increases in equilibrium, reflecting strategic complementarities (Gal-Or, 1985, Dowrick, 1986, Vives, 1999, and Julien, 2011). Recall an SDE is a subgame perfect equilibrium where empty threats are not allowed. So, when the follower's best reply has a positive slope in equilibrium, the leader will be deterred from raising output by the threat of retaliation from the follower, who can increase his supply. The next example, based on Amir and Grilo (1999), illustrates a second-mover advantage:

Example 3 Let $p(X) = \frac{1}{(X+1)^{\gamma}}$, $\gamma > 2$ and $C_L(x_L) = C_F(x_F) = 0$. The best response is given by $\varphi(x_L) = \frac{1}{\gamma-1} + \frac{1}{\gamma-1}x_L$, with $\nu = \frac{1}{\gamma-1} > 0$. We get $(\tilde{x}_L, \tilde{x}_F) = \frac{1}{\gamma-1}\left(1, \frac{\gamma}{\gamma-1}\right)$, $\tilde{p} = \left(\frac{\gamma-1}{\gamma}\right)^{2\gamma}$, and $(\tilde{\pi}_L, \tilde{\pi}_F) = \left(\frac{\gamma-1}{\gamma}\right)^{2\gamma}\left(\frac{1}{\gamma-1}, \frac{\gamma}{(\gamma-1)^2}\right)$, so $\tilde{\pi}_L < \tilde{\pi}_F$. Finally, we have $\left(\tilde{L}_L, \tilde{L}_F\right) = \frac{1}{2\gamma-1}\left(\frac{\gamma}{\gamma-1}\right)^{2(\gamma+1)}\left(\frac{1}{\gamma}, 1\right)$. Then $\tilde{L}_L < \tilde{L}_F$.

Therefore, the leader always achieves higher payoff than the follower for a certain class of nonlinear market demand functions intersecting the price axis. This echoes the results of Kaplan and Wettstein (2000) and of Tasnadi (1999).[5]

2.4 Strategic Behavior and Endogenous Timing

The preceding discussion devoted to the comparison between individual payoffs, notably Propositions 1 and 3, suggests that there may be situations in which the leader has lower payoffs than the follower. In addition, the timing of moves is purely arbitrary: the positions of moves is given even if firms are interchangeable players. Therefore, it questions the leadership position in quantity-setting games. Making Stackelberg leadership endogenous also enables us to test the prediction of the duopoly model. We briefly review some results in some quantity-setting theoretical models and in experimental games literature.[6]

[5] The former authors show that the Bertrand paradox can only be avoided if the demand curve does not intersect the vertical axis, whereas the latter establishes that the avoidance of a mixed-strategy equilibrium in the standard Bertrand-Edgeworth game requires the same condition.

[6] There is a literature devoted to Bertrand games with price leadership in oligopoly markets. Van Damme and Hurkens (2004) consider a linear price-setting duopoly game with differentiated products and endogenously determine which of the players will lead and which one will follow. While the follower role is better for each firm,

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

2.4.1 Some results

The Stackelberg duopoly game provides some insights into the idea of commitment. Schelling (1960) develops the idea of commitment in strategic environments, and shows through examples that committing is beneficial only if the rival does not commit.[7] The theoretical literature devoted to the endogenization of the order of moves in oligopoly starts from the articles by Gal-Or (1985), Reinganum (1985), Dowrick (1986), Boyer and Moreaux (1987a), and Saloner (1987). More recently, several game-theoretic contributions establish conditions under which firms are likely to play either a simultaneous-move game or a sequential-move game in quantity-setting competition with complete information (see Hamilton and Slutsky, 1990, Robson, 1990a, Pal, 1991, Anderson and Engers, 1992, Amir, 1995, Amir and Grilo, 1999, Van Damme and Hurkens, 1999, Matsumura, 1999, among others). Other contributions consider either incomplete information (Mailath, 1993, Normann, 2002), or uncertainty (Albaek, 1990, Spencer and Brander, 1992, Sadanand and Sadanand, 1996). In these papers, the order of supply decisions is not exogenously specified. Rather, it is derived from the firms' decisions about timing. We propose to focus on Hamilton and Slutsky's seminal paper (1990).

Hamilton and Slutsky (1990) consider a duopoly game with an initial stage in which both firms simultaneously decide whether to move first or to move second. They propose two ways of endogenizing the order of moves in a duopoly market with complete information: either with an extended game with observable delay or with an extended game with action commitment. In the former, both firms announce the period in which they decide to move before choosing an action. After the announcement, they choose their actions in the sequence that results from the timing decisions by knowing when its rival will play. In the latter, a firm can move first only by selecting an action to which it is committed. But a firm that plays first does not know whether its opponent plays first or delays. In the extended game with action commitment, the two Stackelberg equilibria are the only pure strategy equilibria in undominated strategies; the Cournot-Nash equilibrium is an outcome in dominated strategies. In the extended game with observable delays, there are three sets of pure strategy subgame perfect equilibria. In case 1, there are two Stackelberg duopoly equilibria if each firm prefers its SDE follower payoffs to its CDE payoffs. In case 2, there is one SDE with one leader if the other firm prefers its SDE follower payoffs than its CDE payoff. In case 3, there is one CDE if each firm prefers its CDE payoffs to its SDE follower payoffs.

These results provide some theoretical foundation for endogenous timing in Stackelberg duopoly.[8] But as emphasized by Hamilton and Slutsky (1990), without further assumptions

---

they show that waiting is more risky for the low-cost firm. Consequently, risk dominance leads to the result that only the high-cost firm will choose to wait. Hence, the low-cost firm will emerge as the endogenous price leader. Amir and Stepanova (2006) generalize the preceding model by considering the issue of first-mover versus second-mover advantage in differentiated products Bertrand duopoly with general demand and asymmetric linear costs. They use the supermodularity framework to generalize existing results for all possible combinations where prices are either strategic substitutes or complements. They show that a firm with a sufficiently large cost can have a first-mover advantage. For the linear version of the model, they invoke a natural endogenous timing scheme coupled with equilibrium selection according to risk dominance. The sequential game produces a unique equilibrium outcome with the low-cost firm as leader. See also the seminal paper by Robson (1990a).

7 This situation displays a coordination problem. Thus, some participants could decide not to commit to avoid a conflict that would result from antagonist leadership positions.

8 Anderson and Engers (1992) consider a game in which firms choose whether to reveal their supplies. This game includes Stackelberg and Cournot settings as possible outcomes. They show that the equilibrium is the Stackelberg setting; all firms decide to reveal in the sequence.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

the existence of multiple perfect equilibria makes crystal clear prediction impossible. Nevertheless, in the extended game with observable delay, the Pareto preferred outcome depends on characteristics of the firms' isoprofits and of the best responses. Indeed, given that the two firms have an optimal mapping to the strategy of the rival, the leader achieves higher payment at the SDE than at the SCE. In addition, it is possible to figure out three configurations. First, if the optimal decision mappings have slopes of the same sign, then the market outcome corresponds to cases 1 or 3. When both optimal mappings are negatively sloped, neither functions intersect the Pareto preferred set to the CDE (the set of strategies that increase both payoffs), then there is a first-mover advantage; whilst, when both optimal mappings are positively sloped, both functions intersect the Pareto preferred set to the CDE, then there may be a second-mover advantage. Second, if these functions have slopes of opposite signs, then the market outcome corresponds to case 2. Thus, the optimal mapping that is negatively sloped intersects the Pareto preferred set to the CDE.[9] By contrast, in the game with action commitment, both sequential games with both order of moves are the only two pure strategy Stackelberg subgame perfect Nash equilibria, regardless of slopes of best responses and Pareto dominance.

2.4.2 Experimental games

Some recent papers consider experimental games to study which kind of duopoly equilibria would emerge. Indeed, results from this literature may indicate whether models of simultaneous output or price decisions (Cournot, Bertrand) or sequential decisions (Stackelberg, price leadership) are plausible.

The experimental evidence on Hamilton and Slutsky's (1990) commitment game shows that: first, simultaneous-move Cournot market outcomes are more frequent than the Stackelberg market outcomes; second, simultaneous-move outcomes are often played in the second period; and, third, in some cases the follower punishes the leader. For instance, Huck, Konrad, and Müller (2002) use the extended duopoly game with action commitment of Hamilton and Slutsky (1990), which predicts the emergence of endogenous Stackelberg leadership. They search for the equilibrium predictions of the extended game with action commitment. Their data do not confirm the prediction. Indeed, while Stackelberg equilibria are rare, they often observe endogenous Cournot outcomes and sometimes collusive play. Hence, any endogenous Stackelberg follower learns over time to reward cooperation and to punish exploitation (see Huck et al., 2001a who suggest that the Stackelberg leader–follower structure is also beneficial for welfare). Given the *empirical response function* of the follower, the Stackelberg leader would have a better payoff by producing less than the amount that corresponds to the subgame perfect equilibrium.

Santos-Pinto (2008) studies endogenous timing by assuming that players are averse to inequality in payoffs. He explores the implications of inequity aversion and compares them to the empirical evidence. He shows that inequity aversion is able to explain most of the experimental evidence (among which collusive outcome) on the action commitment game of Hamilton and Slutsky (1990). He finds that when inequity aversion is high, the game

[9] The case of decreasing optimal responses correspond to Cournot competition. The case of increasing optimal mappings may correspond to Bertrand price competition with differentiated substitute products. The case of opposite slopes in sign may correspond to a quantity-setting firm for the increasing one and a price-setting firm for the decreasing one (see Singh and Vives, 1984). For a discussion of this classification, see Hamilton and Slutsky (1990) and Vives (1999).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

displays only simultaneous-move symmetric equilibria where both firms produce in the first production period, whilst when the inequity aversion is low there is a continuum of simultaneous-move symmetric equilibria, but there are also two Stackelberg equilibria with sequential play.[10]

Fonseca, Müller, and Normann (2006) test the prediction of the game with observable delay of Hamilton and Slutsky (1990), which has a unique subgame perfect equilibrium in which both players choose to behave as Cournot producers in the first period. To this end, they carry out the game both with a random- and a fixed-matching scheme. With random matching, they find that players choose the predicted production period more frequently over time but choices do not converge to the predicted level as nearly one-third of all players still choose to delay toward the end of the experiment. With a fixed-matching scheme, the subgame perfect equilibrium has no predictive power with regard to timing choices as throughout the experiment only half of the timing observations are period 1 choices. These differences in timing choices are explained by the deviations from the prediction observed in the sequential move subgame: with random matching, more competitive behavior in the Stackelberg subgame provides an incentive to avoid it by choosing to produce early.

2.5 Free Entry

Free entry and market performance are linked under Cournot competition: the Cournot oligopoly market outcome coincides with the CE outcome when the number of firms increases without limit. We would like to know whether Stackelberg quantity competition has the same property. To this end, we first consider strategic free entry.[11] In particular, does quantity commitment still give strategic advantage to incumbent firms under free entry? Under which condition(s) does a leader firm always have higher payoffs in the Stackelberg game? Then, we study market power in the extended Stackelberg duopoly game.

2.5.1 SDE and free entry

The SDE may be analyzed as the market outcome resulting from a rivalry between an incumbent and one potential entrant (Spence, 1977 and Dixit, 1980).[12] Indeed, the SDE constitutes a subgame perfect Nash equilibrium with free entry. For instance, in Salop (1979) the entrant is a leader, whilst in Basu and Singh (1990) the leader is the incumbent firm.[13] Other contributions merely focus on the welfare implications of market power (Etro, 2006, 2008, Mukherjee and Zhao, 2009). We here consider the market performance of the Stackelberg structure to emphasize the differences, if any, with Cournot competition. It is well known that in a Cournot game with symmetric firms, free entry may increase or reduce the incumbents' supplies, but it always reduces their profits (Seade, 1980). The following

[10] Rassenti et al. (2000) examine results from laboratory experiments in which five persons participate as sellers in a Cournot oligopoly game. They wonder whether repeated play will lead to convergence to a unique noncooperative Nash equilibrium. The results provide observed intertemporal variation in total output and *heterogeneity* in individual choices that are inconsistent with convergence to the static Nash equilibrium.

[11] See also Appendix D in Anderson, Erkal, and Piccinin (2016).

[12] The starting point of these models is a criticism of the limit price approach according to which the potential entrant firm is assumed to believe that the incumbent firm will maintain the same action after entry as before the entry (Dixit, 1979, Nti and Shubik, 1981, Eaton and Ware, 1987).

[13] Entry may also be linked to endogenous timing (Sadanand and Sadanand, 1996).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

illustration given by Mukherjee and Zhao (2009) emphasizes that Stackelberg competition does not always lead to higher supply when we compare it with the CDE without entry:

Example 4 Let $p(X) = a - X$, $a > 0$, $C_L^1(x_L^1) = 0$, $C_L^2(x_L^2) = c_L x_L^2$, and $C_F(x_F) = c_F x_F$. Assume first there is no entry and both firms behave as Cournot competitors. The CDE is given by $\left(\hat{x}_L^1, \hat{x}_L^2\right) = \left(\frac{a+c_L}{3}, \frac{a-2c_L}{3}\right)$, $\hat{p} = \frac{a+c_L}{3}$, and $(\hat{\pi}_L^1, \hat{\pi}_L^2) = \left(\left(\frac{a+c_L}{3}\right)^2, \left(\frac{a-2c_L}{3}\right)^2\right)$. Consider now the case of entry. The follower's best response is $\varphi(x_L^1, x_L^2) = \frac{a-c_F}{2} - \frac{1}{2}(x_L^1 + x_L^2)$. The SDE is given by $\left(\tilde{x}_L^1, \tilde{x}_L^2, \tilde{x}_F\right) = \left(\frac{a+2c_L+c_F}{3}, \frac{a-4c_L+c_F}{3}, \frac{a+2c_L-5c_F}{6}\right)$, $\tilde{p} = \frac{a+2c_L+c_F}{6}$, and $\tilde{\pi}_L^1 = \frac{1}{2}\left(\frac{a+2c_L+c_F}{3}\right)^2$, $\tilde{\pi}_L^2 = \frac{1}{2}\left(\frac{a-4c_L+c_F}{3}\right)^2$ and $\tilde{\pi}_F = \left(\frac{a+2c_L-5c_F}{6}\right)^2$. Let us notice that $\tilde{x}_F > 0$ whenever $c_F < \frac{a+c_L}{5} \equiv \bar{c}_F$ and that $\bar{c}_F \geqslant c_F$ for $c_L \leqslant \frac{a}{3}$. Therefore, consider $c_L \leqslant \frac{a}{3} \equiv \bar{c}_L$ as a condition for entry. Let us now compare the SDE supplies with the CDE supplies. We have $\tilde{x}_L^1 > \hat{x}_L^1$ but $\tilde{x}_L^2 > \hat{x}_L^2$ if $c_L \in [0, \frac{a}{8}]$ and $c_F \in (2c_L, \bar{c}_F)$. The reason is that under Stackelberg competition, both firms face a higher (effective) demand and the second leader faces higher marginal cost, so its supply increases provided the cost of the follower entrant is relatively high with respect to the second leader's cost. In addition, $\tilde{\pi}_L^1 > \hat{\pi}_L^1$ if $c_F \in (\underline{c}_F(c_L), \bar{c}_F)$, where $\underline{c}_F(c_L) \equiv (\sqrt{2}-1)(a - c_L\sqrt{2})$ and $c_L \in (\underline{c}_L, \bar{c}_L)$, where $\underline{c}_L \equiv a\frac{5\sqrt{2}-6}{12-5\sqrt{2}}$, whilst $\tilde{\pi}_L^2 < \hat{\pi}_L^2$. First, the supply of the first leader increases with the higher (effective) demand, and with the cost of the follower, and second, as supplies are strategic substitutes, the decrease of the other leader's supply is favorable for the first leader. The total effect is positive for the first firm when costs are relatively high.

This example shows that, when firms have different marginal costs and the incumbents behave as Stackelberg leaders, while the entrant behaves as a follower, then entry (does not necessarily) increases the supply (the profit) of the higher efficient cost leader but (not necessarily) decreases the supply (the profit) of the less efficient leader. The next example shows that entry of a second follower in a Stackelberg market may be damaging for the incumbent firms:

Example 5 (Example 4 with firm 2 as a follower) Let now $C_L^1(x_L^1) = 0$, $C_F^1(x_F^1) = c_F^1 x_F^1$, and $C_F^2(x_F^2) = c_F^2 x_F^2$. Assume first, there is no entry and both firms behave as in the SDE game, with firm L as the leader. The best response is $\varphi(x_L^1) = \frac{a+c_F^1}{2} - \frac{1}{2}x_L^1$. The SDE obtains as $\left(\tilde{x}_L^1, \tilde{x}_F^1\right) = \left(\frac{a+c_F^1}{2}, \frac{a-3c_F^1}{4}\right)$, $\hat{p} = \frac{a+c_F^1}{4}$, and $(\hat{\pi}_L^1, \hat{\pi}_F^1) = \left(\frac{1}{2}\left(\frac{a+c_F^1}{2}\right)^2, \frac{(a+c_F^1)(a-3c_F^1)}{16}\right)$. Consider now the case of entry with a new entrant follower. The best responses are $\varphi(x_L^1) = \frac{a-2c^1+c^2}{3} - \frac{1}{3}x_L^1$ and $\varphi^2(x_L^1) = \frac{a+c_F^1-2c_F^2}{3} - \frac{1}{3}x_L^1$. The Stackelberg equilibrium with free entry is given by $\left(\check{x}_L^1, \check{x}_F^1, \check{x}_F^2\right) = \left(\frac{a+c_F^1+c_F^2}{2}, \frac{a-5c^1+c^2}{6}, \frac{a+c_F^1-5c_F^2}{6}\right)$, $\check{p} = \frac{a+c_F^1+c_F^2}{6}$, and $\check{\pi}_L^1 = \frac{1}{3}\left(\frac{a+c_F^1+c_F^2}{2}\right)^2$, $\check{\pi}_F^1 = \left(\frac{a-5c_F^1+c_F^2}{6}\right)^2$ and $\check{\pi}_F^2 = \left(\frac{a+c_F^1-5c_F^2}{6}\right)^2$. We have $\check{x}_F^2 > 0$ whenever $c^2 < \frac{a+c_F^1}{5} \equiv \bar{c}_F^2$ and that $\bar{c}_F^2 \geqslant c_F^1$ for $c_F^1 \leqslant \frac{a}{4}$. Therefore, consider $c_F^1 \leqslant \frac{a}{4} \equiv \bar{c}_F^1$ as a

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

condition for entry ($\bar{c}_F^1 < \frac{a}{3}$ for the no entry case). We now compare the SDE supplies with entry with the SDE supplies without entry. We have $\check{x}_L^1 > \tilde{x}_L^1$ but $\check{x}_F^1 < \tilde{x}_F^1$ as $c_F^2 < \frac{a+c_F^1}{2}$. The incumbent follower faces a lower (effective) demand and the conditions on costs are more stringent with the presence of a new follower. In addition, $\check{\pi}_L^1 < \tilde{\pi}_L^1$ and $\check{\pi}_F^1 < \tilde{\pi}_F^1$. But $\check{X} - \widetilde{X} = \frac{a+c_F^1-2c_F^2}{3} > 0$, and $\check{p} - \tilde{p} = -\frac{a+c_F^1-2c_F^2}{12} < 0$.

This example outlines the consequences of altering the basic duopoly game with new intrants. But may large entry lead to the competitive market outcome?

2.5.2 The extended Stackelberg duopoly game (*T*-stage game *I*)

The first extension to the SDE is performed in Boyer and Moreaux (1986) who consider a linear hierarchical market game as a T-stage oligopoly model with one firm per stage. They show that, by enlarging the game, the Stackelberg equilibrium may coincide at the limit with the CE. In addition, the profit of the first leader is strictly higher than the profit it reaches in the corresponding Cournot game when there are at most two stages of decision. Anderson and Engers (1992) assume $X(p) = a - bp^{\alpha}$, $a, b, \alpha > 0$, and show that each firm behaves as if it were a monopolist facing the residual demand inherited from the preceding movers in the hierarchy. Indeed, a T-stage Stackelberg game, with one firm per stage, is a succession of monopoly choices on residual demand.[14] In addition, the first firm in the hierarchy may reach lower payoffs than the payoffs reached in the Cournot game if and only if $n > \alpha + 2$, all other firms having less payoffs than their Cournot counterparts for all $\alpha, n > 0$. Then, the advantage of being a leader may vanish when the number of stages increases.[15]

Robson (1990b) considers the existence and the welfare property of the Stackelberg equilibrium with free entry. There are T firms, each being indexed by i, $i = 1, \ldots, T$. Each firm has average cost given by the function $a(x^i)$, with $\frac{da(x^i)}{dx^i} \leqslant 0$, $\forall x^i \in (0, x^*]$, $\frac{da(x^i)}{dx^i} > 0$, $\forall x^i > x^*$ for some $x^* > 0$. The minimum level of average cost is given by $p^* = a(x^*)$, and total cost is given by $c(x^i) = x^i a(x^i)$, with $c(x^i) \geqslant 0$, for $x^i \geqslant 0$. The price function $p\,(X)$ is continuous and continuously differentiable. In addition, $\lim_{X\to\infty} p\,(X) < p^* < p\,(0) < \lim_{x^i\to 0} a(x^i)$: the demand and the limiting competitive aggregate supply intersect nontrivially at p^*, and supply per firm is bounded below and the number of active firms is bounded above. The strategy set of firm i is $\mathcal{S}^i = [\alpha\bar{x}, X^*]$, where X^* is the unique solution to $p\,(X) = p^*$, $\bar{x}$ is the unique solution to $a(x^i) = p\,(0)$, $x^i < x^*$, and $\alpha \in (0, 1]$ is a scale factor. Consider firm k in the hierarchy, where $1 < k < T$. Its best-response function is given by $x^k = \varphi^k(x^1, \ldots, x^{k-1})$, $1 < k < T$. A set of pure strategies is uniquely determined recursively as x^1, $x^2 = \varphi^2(x^1)$, $x^3 = \varphi^3(x^1, x^2), \ldots, x^T = \varphi^T(x^1, x^2, \ldots, x^{T-1})$. Then, there exists a subgame perfect Nash equilibrium (SPNE) in pure strategies. In addition, the sequence of Stackelberg equilibria with free entry converges to the CE.

What can be said about the sequential entry and its relationship with welfare in the hierarchical game? There are two kinds of models that study sequential entry. The first kind allows for entry deterrence, whilst the second kind does not. Economides (1993) considers a

---

[14] This is no longer true when the costs are quadratic and/or with a linear demand system for symmetrically differentiated products of the type $p^i = \alpha - \beta x^i - \gamma \sum_{j\neq i} x^j$, where $\beta > \gamma > 0$, and $x^i > 0$, $i = 1, \ldots, n$ (see Vives, 1988).

[15] But it does not imply that firms prefer to play à la Cournot. See also Matsumura (1999).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

separation of the entry and supply decisions in a generalized version of Boyer and Moreaux (1986). The decision of entry is taken at an earlier stage, and entry is simultaneous. Therefore, entry deterrence cannot occur. The inverse demand function is concave and the cost function includes a fixed cost of entry. The aggregate supply and market price of the extended SE with free entry are the same as those of the CE with free entry, with the supply of the last firm equal to its Cournot supply, and with fewer active firms in the SE than in the CE. Second, any firm's supply varies inversely with its order in the sequence of decisions, and equilibrium payoffs decrease, with the last firm making zero payoff. Third, *the extended SE with free entry Pareto dominates the CE with free entry* since the higher social surplus is due to higher payoffs for active firms. The increase in social welfare is caused by more efficient utilization of the technology: there are less active firms in the SE and more firms bear the fixed costs in the CE. Therefore, earlier active firms exploit their strategic advantage by supplying more (efficiently) than any later-acting firm. But entry deterrence is not dealt with.

Vives (1988) considers a set of entrants ordered in a sequence either with one or with several incumbents. All firms have the same linear technology and have to bear an entry cost in case they decide to produce. First, if there are a few (many) potential entrants, the incumbent lets them entry (prefers to prevent them from entering). More potential entrants do not decrease welfare. Second, the potential entrants may face several incumbents. Equilibria with and without entry may coexist. In both cases, the incumbents must supply high levels to prevent entry. Thus, the efficiency of a public policy that consists of lowering the cost of entry depends on the number of entrants. In the same spirit, Anderson and Engers (1994) consider competition over entry time to show how the differences in profits can be dissipated. The order of entry is endogenous and depends on costs. Lower fixed costs can lower the social surplus. But here the incumbent leader firm will no longer choose only between either letting all firms enter or preventing them from entering. Finally, Pal and Sarkar (2001) show that additional entry of a lower firm cost in the hierarchy can increase the supply and profits of existing firms.

3 THE MULTIPLE LEADER–FOLLOWER GAME

The multiple leader–follower two-stage noncooperative game was introduced by Sherali (1984), and explored by Daughety (1990), DeMiguel and Xu (2009), and Julien (2017).[16] It constitutes an interesting and nontrivial extension to the basic Stackelberg duopoly game, and it provides a richer set of strategic interactions between several (heterogeneous) firms than the preceding games. First, strategic interactions are more complex since the entire game consists of two Cournot games embedded in a hierarchical competition game. Therefore, the resulting market outcome, namely existence and uniqueness, is more difficult to handle. Second, some features regarding the working and the consequences of market power are specific to this model and are not captured in the basic duopoly game. We describe the multiple leader–follower oligopoly game. Then, we characterize the optimal behavior. To understand the nature of strategic interactions at work, we study existence and uniqueness of a Stackelberg oligopoly equilibrium (SOE thereafter). Finally, we consider the market power

[16] To the best of our current knowledge, the first model with one leader and several followers was introduced by Leitmann (1978), and developed by Murphy et al. (1983) and Sherali (1984). Note that Stackelberg (1934) already envisaged the possibility of several market participants (see his Chapter 3).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

and the welfare properties of the SOE. To this end, we compare the SOE with the Cournot oligopoly equilibrium (COE), and with the CE. We also outline the main differences with the SDE game.

3.1 The Model

There are now several risk-neutral firms of type L and of type F who compete on quantities to sell the homogeneous divisible product. Thus, the set of firms partitions into two subsets $\mathcal{F}_L = \{1, \ldots, n_L\}$ and $\mathcal{F}_F = \{1, \ldots, n_F\}$, with $\mathcal{F}_L \cup \mathcal{F}_F = \mathcal{F}$ and $\mathcal{F}_L \cap \mathcal{F}_F = \varnothing$. We consider $|\mathcal{F}_L| \geqslant 1$ and $|\mathcal{F}_F| \geqslant 1$, where $|A|$ denotes the cardinality of the set A. Firms of type L are leaders, while firms of type F are followers. Leaders are indexed by i and followers are indexed by j.[17] The cost function of leader i is denoted by $C_L^i(x_L^i)$, $i \in \mathcal{F}_L$. Likewise, for each $j \in \mathcal{F}_F$, we let $C_F^j(x_F^j)$. The costs functions are twice-continuously differentiable, increasing and convex. The price function $p(X)$ is twice-continuously differentiable and decreasing. In addition, it still satisfies $\frac{dp(X)}{dX} + kx\frac{d^2p(X)}{(dX)^2} \leqslant 0$, where $k > 0$.

The profit functions $\pi_L^i(.)$ of firm i and $\pi_F^j(.)$ of firm j may be written:

$$\pi_L^i(x_L^i, \mathbf{x}_L^{-i}, \mathbf{x}_F) = p(X)x_L^i - C_L^i(x_L^i),\ i \in \mathcal{F}_L \tag{10.15}$$

$$\pi_F^j(x_F^j, \mathbf{x}_F^{-j}, \mathbf{x}_L) = p(X)x_F^j - C_F^j(x_F^j),\ j \in \mathcal{F}_F. \tag{10.16}$$

Let us notice that (10.16) is strictly concave with respect to x_F^j given $\mathbf{x}_F^{-j}$ and $\mathbf{x}_L$. In addition, like in Section 2, the concavity of (10.15) is more difficult to state.

Consider now the noncooperative game $\mathbf{\Gamma}$ associated with this economy. The players are the $(n_L + n_F)$ firms, the strategies are their production decisions, and the payoffs are their profits. Let $\mathcal{S}_L^i = [0, \infty)$ be the strategy set of leader $i \in \mathcal{F}_L$, where the supply x_L^i represents the pure strategy of the leader. Similarly, let $\mathcal{S}_F^j = [0, \infty)$, where x_F^j is the pure strategy of the follower $j \in \mathcal{F}_F$. A strategy profile will be represented by the vector $(\mathbf{x}_L, \mathbf{x}_F)$, with $(\mathbf{x}_L, \mathbf{x}_F) \in \prod_{i \in \mathcal{F}_L} \mathcal{S}_L^i \times \prod_{j \in \mathcal{F}_F} \mathcal{S}_F^j$. The corresponding payoffs functions are given by $\pi_L^i(.)$, $i \in \mathcal{F}_L$, and $\pi_F^j(.)$, $j \in \mathcal{F}_F$. This sequential game displays two stages of decisions and no discounting. We also assume the timing of positions is given. Each leader chooses first a quantity to sell, and each follower determines their supply on the residual demand. Information is again assumed to be complete and perfect. The multiple leader–follower model is thus described by a two-stage game that embodies two simultaneous-move partial games. Indeed, the leaders play a two-stage game with the followers, but the leaders (the followers) play a simultaneous-move game together.

[17] We adopt the following notational conventions. Let $\mathbf{x} \in \mathbb{R}_+^n$. Then, $\mathbf{x} \geq \mathbf{0}$ means $x_i \geqslant 0$, $i = 1, \ldots, n$; $\mathbf{x} > \mathbf{0}$ means there is some i such that $x_i > 0$, with $\mathbf{x} \neq \mathbf{0}$, and $\mathbf{x} >> \mathbf{0}$ means $x_i > 0$ for all i, $i = 1, \ldots, n$. Let $\mathbf{x}_L = \left(x_L^1, \ldots, x_L^i, \ldots, x_L^{n_L}\right)$ be a strategy profile of leaders, and $\mathbf{x}_F = \left(x_F^1, \ldots, x_F^j, \ldots, x_F^{n_F}\right)$ be a strategy profile of followers, where x_L^i and x_F^j represent respectively the supply of leader $i \in \mathcal{F}_L$, and of follower $j \in \mathcal{F}_F$. In addition, let $\mathbf{x}_L^{-i} = \left(x_L^1, \ldots, x_L^{i-1}, x_L^{i+1}, \ldots, x_L^{n_L}\right)$ and $\mathbf{x}_F^{-j} = \left(x_F^1, \ldots, x_F^{j-1}, x_F^{j+1}, \ldots, x_F^{n_F}\right)$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

3.2 Stackelberg Oligopoly Equilibrium

In this framework strategic interactions occur within each partial game but also between both partial games through sequential decisions. It is worth noticing that the critical difference with the previous duopoly games stem from the fact that the optimal decision of any follower does not necessarily coincide with their best response (see Julien, 2017).[18] Each follower interacts strategically and simultaneously with all other followers. We determine the optimal behavior. Then, we define the SOE, and we state some existence and uniqueness results.

3.2.1 Optimal behavior

The extended game $\mathbf{\Gamma}$ is solved by backward induction. Let $\phi^j : \prod_{-j\in\mathcal{F}_F} \mathcal{S}_F^{-j} \times \prod_{i\in\mathcal{F}_L} \mathcal{S}_L^i \to \mathcal{S}_F^j$, be follower j's optimal decision, with $x_F^j = \phi^j(\mathbf{x}_F^{-j}, \mathbf{x}_L)$, $j \in \mathcal{F}_F$, which is the solution to:

$$\phi^j\left(\mathbf{x}_F^{-j}, \mathbf{x}_L\right) \in \arg\max \pi_F^j\left(x_F^j, \mathbf{x}_F^{-j}, \mathbf{x}_L\right). \tag{10.17}$$

The first-order sufficient condition may be written:

$$p\,(X) + x_F^j \frac{dp\,(X)}{dX} - \frac{dC_F^j\left(x_F^j\right)}{dx_F^j} = 0, j \in \mathcal{F}_F. \tag{10.18}$$

By the strict concavity of the profit function (10.16), the optimal decision $\phi^j(\mathbf{x}_F^{-j}, \mathbf{x}_L)$ is unique. Let us notice this function is not a best-response function since it depends also on the decisions of the other followers. Let us notice that:

$$\frac{\partial\phi^j}{\partial x_F^{-j}} = -\frac{\frac{\partial^2\pi_F^j}{\partial x_F^j\partial x_F^{-j}}}{\frac{\partial^2\pi_F^j}{\partial\left(x_F^j\right)^2}} = -\frac{\frac{dp\,(X)}{dX} + x_F^j\frac{d^2p\,(X)}{(dX)^2}}{2\frac{dp\,(X)}{dX} + x_F^j\frac{d^2p\,(X)}{(dX)^2} - \frac{d^2C_F^j\left(x_F^j\right)}{\left(dx_F^j\right)^2}}, \tag{10.19}$$

where $\frac{\partial\phi^j}{\partial x_F^{-j}} \in (-1, 0)$, when $\phi^j > 0$, and $\frac{\partial\phi^j}{\partial x_F^{-j}} = 0$ when $\phi^j = 0$. Then, $\frac{\partial\phi^j}{\partial x_F^{-j}} \in (-1, 0]$, $-j, j \in \mathcal{F}_F$. In addition, we can show that $\frac{\partial\phi^j}{\partial x_L^i} \in (-1, 0]$, $i \in \mathcal{F}_L, j \in \mathcal{F}_F$.

To determine the effective demand that is addressed to any leader, i.e., $p\,(x_L^i + X_L^{-i} + \sum_j \varphi^j(\mathbf{x}_L))$, we must show there exist best responses $\varphi^j : \prod_{i\in\mathcal{F}_L} \mathcal{S}_L^i \to \mathcal{S}_F^j$, with $x_F^j = \varphi^j(\mathbf{x}_L)$, $j \in \mathcal{F}_F$. This is the critical difference with the previous duopoly game in which the optimal decision of the follower coincides with their best response. Julien (2017) provides a consistency criterion to determine each optimal decision as a function of the strategy profile of the leaders. In addition, it is possible to show that the best responses are not increasing, so the game displays actions that are strategic substitutes. Note the condition is sufficient, so strategic complementarities could exist provided they are not too strong.

[18] One difficulty stems from the fact that the followers' optimal decision mappings might be mutually inconsistent.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Therefore, leader i's optimal decision, which is defined by $\psi^{i} : \prod_{-i\in\mathcal{F}_L} \mathcal{S}_L^{-i} \to \mathcal{S}_L^{i}$, with $x_L^i = \psi^{i}\left(\mathbf{x}_L^{-i}\right)$ and $\boldsymbol{\varphi}(X_L) = (\varphi^1(X_L), \ldots, \varphi^n(x_L))$, is the solution to:

$$\psi^{i}\left(\mathbf{x}_L^{-i}\right) \in \arg\max \pi_L^i\left(x_L^i, \mathbf{x}_L^{-i}, \boldsymbol{\varphi}\left(\mathbf{x}_L\right)\right), i \in \mathcal{F}_L. \tag{10.20}$$

As $p\left(X\right) = p\left(x_L^i + X_L^{-i} + \sum_j \varphi^j\left(x_L^i + X_L^{-i}\right)\right)$, the condition $\frac{\partial \pi^{i}}{\partial x_L^i} = 0$ leads to:

$$p\left(X\right) + \left(1 + \nu^{i}\right) x_L^i \frac{dp\left(X\right)}{dX} - \frac{dC_L^i\left(x_L^i\right)}{dx_L^i} = 0, i \in \mathcal{F}_L. \tag{10.21}$$

The term $\nu^{i} = \frac{\partial \sum_j \varphi^j(\mathbf{x}_L)}{\partial x_L^i}$, with $\nu^{i} \in [-1, 0]$, represents the reaction of all followers to leader i's strategy, i.e., the slope of the aggregate best response to i, $i \in \mathcal{F}_L$. By construction $\nu^{i} = \nu^{-i} = \nu$ for all $i, -i \in \mathcal{F}_L$. This term has the same interpretation as in subsection 2.2, with $k = (1 + \nu)$. It is possible to check that the second-order sufficient condition holds:

$$\frac{\partial^2 \pi_L^i}{(\partial x_L^i)^2} = k\left(k x_L^i \frac{d^2 p\left(X\right)}{(dX)^2} + 2\frac{dp\left(X\right)}{dX}\right) - \frac{d^2 C_L^i\left(x_L^i\right)}{\left(dx_L^i\right)^2} < 0. \tag{10.22}$$

Finally, let us notice that $\frac{\partial^2 \pi^{i}}{\partial x_L^i \partial x_L^{-i}} = k\left(k\frac{dp\left(X\right)}{dX} + x_L^i \frac{d^2 p\left(X\right)}{(dX)^2}\right) \leqslant 0$, $i \in \mathcal{F}_L$; and $\frac{\partial \psi^{i}}{\partial x_L^{-i}} = -\frac{\frac{\partial^2 \pi_L^i}{\partial x_L^i \partial x_L^{-i}}}{\frac{\partial^2 \pi_L^i}{\partial (x_L^i)^2}} = -\frac{k\frac{dp\left(X\right)}{dX} + k^2 x_L^i \frac{d^2 p\left(X\right)}{(dX)^2}}{2k\frac{dp\left(X\right)}{dX} + k^2 x_L^i \frac{d^2 p\left(X\right)}{(dX)^2} - \frac{d^2 C_L^i(x_L^i)}{(dx_L^i)^2}}$, so $\frac{\partial \phi^j}{\partial x_F^{-j}} \in (-1, 0]$, for all $-i \neq i, -i, i \in \mathcal{F}_L$.

3.2.2 SOE: definition, existence and uniqueness

We first provide a definition of an SOE.

Definition 2 (SOE) A Stackelberg oligopoly equilibrium of $\boldsymbol{\Gamma}$ is given by a strategy profile $\left(\tilde{\mathbf{x}}_L, \tilde{\mathbf{x}}_F\right) \in \prod_{i\in\mathcal{F}_L} \mathcal{S}_L^i \times \prod_{j\in\mathcal{F}_F} \mathcal{S}_F^j$, with $\mathbf{x}_F = \boldsymbol{\varphi}(\mathbf{x}_L)$, where $\boldsymbol{\varphi} : \prod_{i\in\mathcal{F}_L} \mathcal{S}_L^i \to \prod_{j\in\mathcal{F}_F} \mathcal{S}_F^j$, such that:

a. $\pi_F^j\left(\tilde{x}_F^j, \tilde{\mathbf{x}}_F^{-j}, \tilde{\mathbf{x}}_L\right) \geqslant \pi_F^j\left(x_F^i, \tilde{\mathbf{x}}_F^{-j}, \tilde{\mathbf{x}}_L\right)$, $\forall x_F^j \in \mathcal{S}^j$, $j \in \mathcal{F}_F$

b. $\pi_L^i\left(\tilde{x}_L^i, \tilde{\mathbf{x}}_L^{-i}, \boldsymbol{\varphi}\left(\tilde{x}_L^i, \tilde{\mathbf{x}}_L^{-i}\right)\right) \geqslant \pi_L^i\left(x_L^i, \tilde{\mathbf{x}}_L^{-i}, \boldsymbol{\varphi}(x_L^i, \tilde{\mathbf{x}}_L^{-i})\right)$, $\forall \boldsymbol{\varphi}(\mathbf{x}_L) \in \prod_{j\in\mathcal{F}_F} \mathcal{S}_F^j$, $\forall \mathbf{x}_L^{-i} \in \prod_{-i\in\mathcal{F}_L} \mathcal{S}_F^{-i}$ *and* $\forall x_L^i \in \mathcal{S}_L^i$, $i \in \mathcal{F}_L$.

Existence and uniqueness problems are more difficult to handle in this framework, insofar as strategic interactions occur within each partial game but also between both partial games through sequential decisions. Indeed, the n_L leaders play a two-stage game with the n_F followers, but the leaders (the followers) play a simultaneous-move game together. But

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

the presence of several leaders and followers displays a richer set of strategic interactions. Strategic interactions occur between both stages of the game but also within each stage. Indeed, the best responses could not be well defined in the presence of several heterogeneous followers. As a consequence, existence and uniqueness are more difficult to handle with several heterogeneous firms in each stage. By taking into account this problem, Julien (2017) shows existence and uniqueness of an SOE under general assumptions on costs and demand functions.[19] Therefore, is it possible to state the following two theorems:

Theorem 1 (Existence) *Consider the game* $\mathbf{\Gamma}$*, and let Assumptions 1 and 2 be satisfied. Then, there exists a Stackelberg oligopoly equilibrium.*

Proof For more details, see Julien (2017). ■

Theorem 2 (Uniqueness) *Let Assumptions 1 and 2 be satisfied. Then, if there exists a Stackelberg oligopoly equilibrium, then it is unique.*

Proof To this end, consider $\mathbf{\Upsilon}_L = \left(\frac{\partial \pi_L^1}{\partial x_L^1}, \ldots, \frac{\partial \pi_L^i}{\partial x_L^i}, \ldots, \frac{\partial \pi_L^{n_L}}{\partial x_L^{n_L}}\right)$ (see Julien, 2017 for details). Let $\left|J_{-\mathbf{\Upsilon}_L}(\tilde{\mathbf{x}}_L, \tilde{\mathbf{x}}_F)\right|$, with $J_{-\mathbf{\Upsilon}_L} = -\left(\frac{\partial^2 \pi_L^i}{\partial x_L^i \partial x_L^{-i}}\right)$, where $\frac{\partial \pi_L^i}{\partial x_L^i} = p(X) + k x_L^i \frac{dp(X)}{dX} - \frac{dC_L^i(x_L^i)}{dx_L^i}$. From Corollary 2.1 in Kolstad and Mathiesen (1987), as leaders in the partial game $\mathbf{\Gamma}_L$ behave as Cournot firms, we show this criterion is satisfied, so the SPNE in $\mathbf{\Gamma}_L$ is unique. It is possible to show that:

$$\left|J_{-\mathbf{\Upsilon}_L}\right| = \left(1 - k \sum_{i \in \mathcal{F}_L} \frac{\frac{dp(X)}{dX} + k x_L^i \frac{d^2 p(X)}{(dX)^2}}{\frac{d^2 C_L^i(x_L^i)}{(dx_L^i)^2} - k \frac{dp(X)}{dX}}\right) \prod_{i \in \mathcal{F}_L} \left(\frac{d^2 C_L^i(x_L^i)}{(dx_L^i)^2} - k \frac{dp(X)}{dX}\right). \tag{10.23}$$

Then, using the assumptions on costs and demand, we deduce:

$$sign\left|J_{-\mathbf{\Upsilon}_L}\right| = sign\left(1 - k \sum_{i \in \mathcal{F}_L} \frac{\frac{dp(X)}{dX} + k x_L^i \frac{d^2 p(X)}{(dX)^2}}{\frac{d^2 C_L^i(x_L^i)}{(dx_L^i)^2} - k \frac{dp(X)}{dX}}\right) > 0. \tag{10.24}$$

[19] Sherali (1984) shows existence and uniqueness with identical convex costs for leaders, and states some results under the assumptions of linear demand either with linear or quadratic costs (see Ehrenmann, 2004). Sherali's model constitutes an extension of the seminal paper of Murphy et al. (1983) that covers the case of many followers who interact with one leader. But they do not study the conditions under which the followers' optimal decisions are mutually consistent. In the same vein, Tobin (1992) provides an efficient algorithm to find a unique SE by parameterizing the price function by the leader's strategy. Some strong assumptions on the thrice-differentiability of the price function and cost function of the leader profit function are made. Following De Wolf and Smeers (1997) who extend Murphy et al. (1983), De Miguel and Xu (2009) extend Sherali (1984) to uncertainty with stochastic market demand. Unlike Sherali (1984) they allow costs to differ across leaders. Nevertheless, to show the concavity of the expected profit of any leader, they assume that the follower aggregate best response is convex. But as this assumption does not always hold, they must resort to a linear demand. Fukushima and Pang (2005), Wang and Yu (2008), and Jia, He, and Xiang (2015) prove existence of an equilibrium point of a finite game with two leaders and several followers without specifying the assumptions made on demand and costs.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

As $\left|J_{-\Upsilon_L}((\tilde{\mathbf{x}}_L, \tilde{\mathbf{x}}_F)\right| > 0$ there exists a unique Nash equilibrium in $\mathbf{\Gamma}_L$, and then, a unique pure strategy Nash equilibrium of $\mathbf{\Gamma}_F$. Then, the SPNE of $\mathbf{\Gamma}$ is unique, which proves uniqueness of the SOE. ■

Assuming symmetry, the condition for the sign of $\left|J_{-\mathbf{\Upsilon}_L}((\tilde{\mathbf{x}}_L, \tilde{\mathbf{x}}_F)\right|$ might be rewritten as $\frac{dp\,(X)}{dX} + kx_L^i \frac{d^2p\,(X)}{(dX)^2} < \frac{1}{kn_L}\left(\frac{d^2C_L^i(x_L^i)}{(dx_L^i)^2} - k\frac{dp\,(X)}{dX}\right)$, which would say that "on average" leaders' marginal revenues could be increased but not too much. In addition, $\frac{d^2C_L^i(x_L^i)}{(dx_L^i)^2} - k\frac{dp\,(X)}{dX} + n_L\left(\frac{dp\,(X)}{dX} + kx_L^i\frac{d^2p\,(X)}{(dX)^2}\right) = \frac{\partial^2\pi_L^1}{(\partial x_L^1)^2} + (n_L-1)\frac{\partial^2\pi_L^i}{\partial x_L^i\partial x_L^{-i}} < 0$: the effect on i's marginal profit of a change in x_L^i dominates the sum of the cross-effects of similar changes of other leaders' supply.

We illustrate the SOE concept with the linear model of Daughety (1990).

3.2.3 The linear game

Consider a market that embodies n firms: there are $n_L \geqslant 1$ leader(s) and $n_F \geqslant 1$ follower(s), with $n_L + n_F = n$. Let $p\,(X) = a - bX$, $a, b > 0$, where $X \equiv X_L + X_F$, with $X_L \equiv \sum_{i=1}^{n_L} x_L^i$ and $X_F \equiv \sum_{j=1}^{n_F} x_F^j$. The costs functions are given by $C_L^i(x_L^i) = cx_L^i$, $i = 1, \ldots, n_L$, and by $C_F^j\left(x_F^j\right) = cx_F^j$, $j = 1, \ldots, n_F$, with $c < a$. The strategy sets are given by $\mathcal{S}_L^i = \mathcal{S}_F^j = \left[0, \frac{a}{b} - c\right]$, $i \in \mathcal{F}_L$, $j \in \mathcal{F}_F$. As a reference the CE aggregate supply and market price are given respectively by $X^* = \frac{a-c}{b}$ and $p^* = c$. In addition, the COE is given by $\hat{x}_L^i = \hat{x}_F^j = X^*\frac{1}{n_L+n_F+1}$, $\widehat{X} = X^*\frac{n_L+n_F}{n_L+n_F+1}$, $\tilde{p} = \frac{a+c(n_L+n_F)}{n_L+n_F+1}$, $\hat{\pi}_L^i = \hat{\pi}_F^j = \frac{(a-c)^2}{b(n_L+n_F+1)^2}$, $i \in \mathcal{F}_L, j \in \mathcal{F}_F$.

Follower j's program is $\max \pi_F^j\left(x_F^j, \mathbf{x}_F^{-j}, \mathbf{x}_L\right) = \left[a - b(x_F^j + X_F^{-j} + X_L) - c\right]x_F^j$. The solution to equation (10.18) is given by $\phi^j\left(\mathbf{x}_F^{-j}, \mathbf{x}_L\right) = \frac{a-c}{2b} - \frac{1}{2}\left(X_F^{-j} + X_L\right)$, where we let $X_F^{-j} \equiv \sum_{-j\neq j} x_F^{-j}$. Then, the best response of follower j is:

$$\varphi^j(\mathbf{x}_L) = \frac{a-c}{b(n_F+1)} - \frac{1}{n_F+1}X_L, j \in \mathcal{F}_F \tag{10.25}$$

The effective demand is $p\,(X_L) = \frac{a+cn_F}{n_F+1} - \frac{bn_F}{n_F+1}X_L$. Then, leader i solves:

$$\max\ \pi_L^i(x_L^i, \mathbf{x}_L^{-i}) = \left[\frac{a+cn_F}{n_F+1} - \frac{bn_F}{n_F+1}(x_L^i + X_L^{-i}) - c\right]x_L^i. \tag{10.26}$$

Assuming $x_L^i = x_L^{-i}$, $-i \neq i$, we deduce the equilibrium strategy of leader i:

$$\tilde{x}_L^i = X^*\frac{1}{n_L+1}, i \in \mathcal{F}_L. \tag{10.27}$$

Then, we deduce the equilibrium strategy of follower j:

$$\tilde{x}_F^j = X^*\frac{1}{(n_L+1)(n_F+1)}, j \in \mathcal{F}_F. \tag{10.28}$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Therefore $\widetilde{X} = X^* \frac{n_L(n_F+1)+n_F}{(n_L+1)(n_F+1)}$, so $\tilde{p} = \frac{a+c[n_L(n_F+1)+n_F]}{(n_L+1)(n_F+1)}$. Then, the payoffs are given by:

$$\tilde{\pi}_L^i = \frac{(a-c)^2}{b(n_L+1)^2(n_F+1)}, i \in \mathcal{F}_L \tag{10.29}$$

$$\tilde{\pi}_F^j = \frac{(a-c)^2}{b[(n_L+1)(n_F+1)]^2}, j \in \mathcal{F}_F. \tag{10.30}$$

We remark that for each $i \in \mathcal{F}_L$, $\tilde{\pi}_L^i \geqslant \hat{\pi}_L^i$ whenever $n_L \leqslant \sqrt{n_F+1}$ any leader achives higher payoff provided the number of leaders is not too high.

We now investigate the welfare implications of the SOE game.

3.3 Market Power and Welfare

We now compare the SOE market outcome with the COE, and with the CE. Then, we study the relation between market concentration and surplus, and also the relation between individual market power and payoffs. Do the welfare properties of the SOE differ from those of the SDE? Why or why not?

3.3.1 The SOE, COE and CE aggregate market outcomes

We can state the following proposition, which relies on Proposition 1:

Proposition 4 *The SOE aggregate supply (market price) is higher (lower) than the COE supply (market price). But the SOE aggregate supply (market price) is lower (higher) than the CE aggregate supply (market price).*

The leaders can set higher supply since they know the followers behave as Cournot competitors. In addition, the increment in the total supply of leaders more than compensates for the decrease of the total supply of followers when the aggregate best response decreases, whilst it goes in the same direction when the aggregate best response increases, i.e., when strategies are complements. Therefore, like in the SDE, the aggregate supply (market price) is higher (lower) in the SOE than in the COE either when strategies are substitutes or when they are complements. The next example illustrates this (Daughety, 1990).

Example 6 (linear game continued) Consider the market outcome given by (10.27)–(10.28). We deduce $\widetilde{X}_L = X^* \frac{n_L}{n_L+1}$ and $\widetilde{X}_F = X^* \frac{n_F}{(n_L+1)(n_F+1)}$. Then, $\widetilde{X} = X^* \frac{n_L n_F + n_L + n_F}{(n_L+1)(n_F+1)}$, which may be written as $\widetilde{X}(n_L, n) = X^* \frac{n + n n_L - n_L^2}{(n_L+1)(n-n_L+1)}$. We see that $\widetilde{X} < X^*$, and then, $p^* < \tilde{p} = \frac{a+c[n_L(n_F+1)+n_F]}{(n_L+1)(n_F+1)}$. We remark that $\widetilde{X}(0,n) = \widetilde{X}(n,n) = X^* \frac{n}{n+1}$, which corresponds to the two Cournot oligopoly equilibria, and $\widetilde{X}(2,n) = \frac{4n-9}{4(n-2)} < \widetilde{X}(2,n) = \frac{3n-4}{3(n-1)} > X(1,n) = \frac{2n-1}{2n} > X(0,n)$. Then, for fixed n, the aggregate supply is concave in n_L, i.e., $\frac{\partial^2 X(n_L,n)}{(\partial n_L)^2} = -\frac{2X^*}{(n_L+1)^3(n-n_L+1)} < 0$. Indeed, the Cournot aggregate supply is given by $\widehat{X}(n_L, n_F) = X^* \frac{n_L+n_F}{n_L+n_F+1}$. Then, we have $\widehat{X}(n_L, n_F) < \widetilde{X}(n_L, n_F)$.

The next proposition echoes Proposition 2, but it refers to the slope of the *aggregate* best response (see Julien, 2011):

Proposition 5 *If* $\nu = 0$, *then the SOE coincides with the COE.*

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Proof Immediate from (10.18) and (10.21). ■

3.3.2 Market concentration and welfare

Welfare is defined as in subsection 2.3 but since there are now several leaders and followers, we must take into consideration, when considering the variation in aggregate supply, the shares of aggregate supply of leaders and of followers. Let $\vartheta_L \equiv \frac{X_L}{X}$, with $0 \leqslant \vartheta_L \leqslant 1$, and $\vartheta_F \equiv \frac{X_F}{X}$, with $0 \leqslant \vartheta_F \leqslant 1$, and where $\vartheta_L + \vartheta_F = 1$. Therefore, the social surplus may now be defined as:

$$\mathbf{S}(X) := \int_0^X p(z)dz - \left[\sum_{i=1}^{n_L} C_L^i(s_L^i \vartheta_L X) + \sum_{j=1}^{n_F} C_F^j(s_F^j \vartheta_F X)\right], \text{ with } X \leqslant X^*, \tag{10.31}$$

where $s_L^i \equiv \frac{x_L^i}{X_L}$ is the leader i's market share, and $s_F^j \equiv \frac{x_F^j}{X_F}$ is follower j's market share. Differentiating partially with respect to X and decomposing $p(X)$ lead to $\frac{\partial \mathbf{S}(X)}{\partial X} = \sum_{i=1}^{n_L} s_L^i \left[p(X) - \vartheta_L \frac{dC_L^i(x_L^i)}{dX}\right] + \sum_{j=1}^{n_F} s_F^j \left[p(X) - \vartheta_F \frac{dC_F^j(x_F^j)}{dX}\right] \geqslant 0$ for fixed s_L^i, s_F^j, ϑ_L and ϑ_F, with $\frac{\partial \mathbf{S}(X)}{\partial X}_{|X=X^*} = 0$. Then, the social surplus is higher at the SOE than at the COE, and reaches its maximum value at the CE. Indeed, $\frac{\partial S_C(X)}{\partial X} = -X\frac{dp(X)}{dX} > 0$, with $S_C(X) := \int_0^X p(z)dz - p(X)X$. In addition, if we let $S_P(X) := p(X)(\vartheta_L \sum_{i=1}^{n_L} s_L^i + \vartheta_F \sum_{j=1}^{n_F} s_F^j)X - \sum_{i=1}^{n_L} C_L^i(s_L^i \vartheta_L X) - \sum_{j=1}^{n_F} C_F^j(s_F^j \vartheta_F X)$, then $\frac{dS_P(X)}{dX} = p(X) + X\frac{dp(X)}{dX} - [\vartheta_L \sum_{i=1}^{n_L} s_L^i \frac{dC_L^i(s_L^i \vartheta_L X)}{dX} + \vartheta_F \sum_{j=1}^{n_F} s_F^j \frac{dC_F^j(s_F^j \vartheta_F X)}{dX}] < 0$ (costs are convex). Therefore, one essential feature of the SOE game is that the strategic interactions between leaders and followers may be welfare enhancing.

Daughety (1990) shows that, by taking the aggregate supply as a measure of welfare, welfare may be maximized when there is considerable asymmetry in the market, whilst symmetric (Cournot) equilibria for which $n_L = 0$ and $n_L = n$ minimize welfare (see Daughety, 1990). Example 7 illustrates this:

Example 7 (linear game continued) Let $\widetilde{X}(n_L, n) = X^* \frac{n + nn_L - n_L^2}{(n_L+1)(n - n_L + 1)}$ as a measure of welfare. We know that, for fixed n, the aggregate supply is concave in n_L for fixed n. Therefore, let $\max_{n_L} \widetilde{X}(n_L, n)$. The solution to this program, namely the optimal number of leaders $\tilde{n}_L$, exists and is unique; it is given by $\tilde{n}_L = \frac{n}{2}$ when n is even, and by either $\tilde{n}_L = \frac{n-1}{2}$ or $\tilde{n}_L = \frac{n+1}{2}$ when n is odd. Then, for a fixed number of firms, welfare is maximized when there is the same number of leaders and followers (see Daughety, 1990).

What can be said in this context regarding concentration and welfare? Consider the Herfindahl index, which is now denoted by $H(n_L, n_F)$, and that is defined as:

$$H(n_L, n_F) := \sum_{i=1}^{n_L} (\vartheta_L s_L^i)^2 + \sum_{j=1}^{n_F} (\vartheta_F s_F^j)^2. \tag{10.32}$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Unlike the SDE game, one critical feature of the multiple leader–follower game is the concentration index, which might be no longer appropriate for measuring welfare, as the next example illustrates (see Daughety, 1990)!

Example 8 (linear game continued) Equation (10.23) leads to $\tilde{H}(n_L, n) = \frac{\vartheta_L^2}{n_L} + \frac{(1-\vartheta_L)^2}{n-n_L}$ for $1 \leqslant n_L < n-1$ and to $\tilde{H}(n_L, n) = \frac{1}{n}$ for $n_L = 0$ or $n_L = n$. Some algebra yields $\frac{\partial \tilde{H}(n_L,n)}{\partial n_L} \geqslant 0$ when $n_L \leqslant 1$, and $\frac{\partial \tilde{H}(n_L,n)}{\partial n_L} < 0$ when $n_L > 1$. Recall that $\widetilde{X}(n_L, n) = X^* \frac{n+nn_L-n_L^2}{(n_L+1)(n-n_L+1)}$ and $\tilde{n}_L = \frac{n}{2}$ when n is even, and either $\tilde{n}_L = \frac{n-1}{2}$ or $\tilde{n}_L = \frac{n+1}{2}$ when n is odd. With one leader there is considerable asymmetry in the market. When the number of leaders increases, the aggregate shares of leaders ϑ_L and the individual shares s_L^i increase. Therefore, the market outcome becomes less asymmetric (the Herfindahl index decreases). But welfare first increases until the optimal number of leaders $\tilde{n}_L$ is reached and it decreases after. Then, welfare is not necessarily positively correlated with concentration!

Therefore, unlike the SDE game, the SOE game shows that market concentration may be of little concern with welfare.

3.3.3 Market power and payoffs

The next example shows that there is no positive correlation between market power and concentration:

Example 9 (linear game continued) For fixed n, the average profit defined as $\bar{\pi}(n_L, n) := \frac{\sum_i \pi_L^i + \sum_j \pi_F^j}{n}$, and given by $\bar{\pi}(n_L, n) = \frac{(a-c)^2}{bn} \frac{n+nn_L-n_L^2}{(n_L+1)^2(n-n_L+1)^2}$, is convex in n_L. Moreover, we have $\frac{\partial \bar{\pi}(n_L,n)}{\partial n_L} = \frac{(a-c)^2}{bn} \frac{(n+nn_L-n_L^2-1)(2n_L-n)}{(n_L+1)^3(n-n_L+1)^2}$. Then, $\frac{\partial \bar{\pi}(n_L,n)}{\partial n_L} \gtreqqless 0$ when $n_L \gtreqqless \frac{n}{2}$. Recall that $\frac{\partial \tilde{H}(n_L,n)}{\partial n_L} < 0$ when $n_L > 1$. So average firm profit and concentration are positively correlated when the number of leaders is less than the socially optimal number of leaders (Daughety, 1990). Thus, unlike the SDE, there is no positive correlation between $\bar{\pi}$ and $\tilde{H}$ no longer holds in the SOE unless there are few leaders.

We now compare the SDE payoffs with the CDE payoffs. The optimal conditions (10.18) and (10.21) may be expressed respectively as $p(X) = (1+m_F^j)\frac{dC_F^j(x_F^j)}{dx_F^j}$, with $\kappa_F^j = \frac{1}{1+\frac{1}{\epsilon}\vartheta_F s_F^j} - 1$, $j \in \mathcal{F}_F$, and as $p(X) = (1+\kappa_L^i)\frac{dC_L^i(x_L^i)}{dx_L^i}$, with $\kappa_L^i = \frac{1}{1+\frac{1+\nu}{\epsilon}\vartheta_L s_L^i} - 1$, $i \in \mathcal{F}_L$, where κ_F^j and κ_L^i are the markups of follower j and leader i. To analyze the relation between market power and individual payoffs, consider:

$$L_F^j = -\frac{1}{\epsilon}\vartheta_F s_F^j, j \in \mathcal{F}_F \text{ and } L_L^i = -\frac{1+\nu}{\epsilon}\vartheta_L s_L^i, i \in \mathcal{F}_L, \tag{10.33}$$

where L_F^j and L_L^i are the Lerner indexes of follower j and of leader i respectively.

The following corollary is the counterpart of Proposition 3:

Corollary 1 *If $L_L^i > L_F^j$, then $\tilde{\pi}_L^i > \tilde{\pi}_F^j$, $i \in \mathcal{F}_F$, $j \in \mathcal{F}_F$. In addition, assume for all $i \in \mathcal{F}_F$ and $j \in \mathcal{F}_F$, $L_L^i = L_F^j$. Then, $\tilde{\pi}_L^i \gtreqqless \tilde{\pi}_F^j$ if and only if $\nu \lesseqqgtr 0$, $i \in \mathcal{F}_F$, $j \in \mathcal{F}_F$.*

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Proof Immediate by comparing expressions in (10.33). ■

Corollary 1 extends Proposition 3. However, differences in leaders' (resp. followers') payoffs are caused by asymmetries in costs. As the extended game embodies strategic interactions among several leaders and followers, we now explore the possibility of merging.

3.3.4 Mergers, payoffs and welfare

The welfare effects and the profitability of horizontal merger in oligopoly quantity-setting games depend on firms' costs functions. For instance, in Cournot competition bilateral mergers between two firms can be profitable if costs are sufficiently convex (Reynolds, Salant, and Switer, 1983, Perry and Porter, 1985, Farell and Shapiro, 1990). But the (strategic) effects of merging on welfare and on the profitability also depend on the market structure. Indeed, the advantages and disadvantages of merging are due to noncooperative strategic behavior that prevails in the SOE (Daughtety, 1990, Heywood and McGinty, 2007, 2008, and Huck et al., 2001b, Huck, Muller and Normann, 2002). The following two examples illustrate the welfare effects of merging and the conditions under which bilateral merger is profitable in two distinct experiments (see Daughety, 1990 and Huck et al., 2001b). These examples show that the multiple leader–follower game may provide counterintuitive results:

Example 10 (linear model continued) *Effects of mergers on welfare:* Let $\widetilde{X}(n_L, n) = X^* \frac{n+nn_L-n_L^2}{(n_L+1)(n-n_L+1)}$. First, a merger means that one firm disappears from the market. Consider the following three cases: (1) merger of two leaders so the post-merger market has $n_L - 1$ leaders but still $n - n_L$ followers; (2) merger of two followers, so there are n_L leaders but $n - n_L - 1$ followers; and (3) merger of one leader and one follower, so there are n_L leaders but $n - n_L - 1$ followers. Therefore, in case 1, some calculations yield $\tilde{X}(n_L - 1, n - 1) - \widetilde{X}(n_L, n) = -X^* \frac{(n-1)n_L(n_L+3)-2n_L+1}{n_L(n_L+1)(n-n_L+1)} < 0$. In cases 2 and 3, we get $\widetilde{X}(n_L, n - 1) - \widetilde{X}(n_L, n) = -X^* \frac{1}{(n_L+1)(n-n_L)(n-n_L+1)} < 0$. Thus, welfare is always reduced. Second, consider now the number of leaders increases. The comparative statics yields: $\frac{\partial \widetilde{X}(n_L,n)}{\partial n_L} = X^* \frac{n-2n_L}{(n_L+1)^2(n-n_L+1)^2} \lesseqqgtr 0$ for $n \lesseqqgtr 2n_L$, $\frac{\partial \widetilde{X}(n_L,n)}{\partial n} = X^* \frac{1}{(n_L+1)(n-n_L+1)^2} > 0$, and $\frac{\partial^2 \widetilde{X}(n_L,n)}{\partial n_L \partial n} = X^* \frac{n-2n_L}{(n_L+1)^2(n-n_L+1)^2} \lesseqqgtr$ for $3n_L + 1 \lesseqqgtr n$. The last effect captures the effect on welfare of changes in industry structure. Consider that two followers merge and behave as a leader firm. There are now $n - 1$ firms with $n_L + 1$ leaders. Some algebra leads to $\widetilde{X}(n_L + 1, n - 1) - \widetilde{X}(n_L, n) = X^* \frac{n-3(n_L+1)}{(n_L+1)(n_L+2)(n-n_L-1)(n-n_L+1)} > 0$ whenever $n_L < \frac{n}{3} - 1$: so, when there are few leaders, merging can increase aggregate supply. Asymmetry is socially desirable as it enhances welfare. But when $n_L > \frac{n}{2}$, less leaders and more followers could increase welfare.

In both experiments, the advantages or disavantages in terms of welfare do not stem from costs but merely from the noncooperative SOE itself: the mergers alter the strategic behavior of market participants (Daughety, 1990). The difference between the two experiments explains by the fact that, in the second case, the reduction of the number of followers is associated with an increase in the number of leaders. But which are the effects of bilateral merging on payoffs?

Example 11 (linear model continued) *Effects of mergers on payoffs in the first experiment:* Let us write (10.29) and (10.30) as $\tilde{\pi}_L^i(n_L, n)$ for all $i \in \mathcal{F}_F$, and as $\tilde{\pi}_F^j(n_L, n)$ for all $j \in \mathcal{F}_F$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Select one i and one j. Consider the first experiment with the three merging cases. Therefore, in the case of a merger of two leaders, some calculations yield $\tilde{\pi}_L^i(n_L-1, n-1) - 2\tilde{\pi}_L^i(n_L, n) = -\frac{(a-c)^2(n_L^2-2n_L-1)}{bn_L^2(n_L+1)^2(n-n_L+1)} > 0$ if $n_L = 2$. In case of a merger of two followers, we get $\tilde{\pi}_F^j(n_L, n-1)$ $-2\tilde{\pi}_F^j(n_L, n) = -\frac{(a-c)^2(n_L^2-+2n_L-2nn_L+n^2-2n-1)}{b(n-n_L)^2(n_L+1)^2(n-n_L+1)^2} > 0$ if $1-\sqrt{2} < n-n_L < 1+\sqrt{2}$. In the case of a merger of one leader and one follower, we get $\tilde{\pi}_L^i(n_L, n-1) - \tilde{\pi}_L^i(n_L, n) - \tilde{\pi}_F^j(n_L, n) = \frac{(a-c)^2}{b(n-n_L)^2(n_L+1)(n-n_L+1)^2} > 0$. Consider now the second experiment with $n-1$ firms with n_L+1 leaders. We get $\tilde{\pi}_L^i(n_L+1, n-1) - 2\tilde{\pi}_F^j(n_L, n) = -\frac{(a-c)^2(n_L+1)^2(n-n_L+1)^2-2(n_L+2)^2(n-n_L-1)}{b((n_L+1)(n_L+2)(n-n_L-1)(n-n_L+1)} > 0$ for $n \leqslant 28$ is also welfare enhancing.

The first experiment shows that merger between two firms who belong to the same cohort and have the same market power rarely have an incentive to merge, while merger between two firms who belong to two distinct cohorts and have different market power is always profitable since the leader firm incorporates the follower firm regardless of the number of rivals. The merger supplies the same amount as the leader did before the merging (here $X^*\frac{1}{n_L+1}$), but the joint payoff increases due to a rise in market price (here $\frac{a+c(n-n_L+1)^2}{(n_L+1)(n-n_L)(n-n_L+1)} > 0$). In the SOE the merger internalizes better the effect of the increase in price on payoffs than in the COE: the decrease in supply is lower than under Cournot quantity competition (see Huck et al., 2001b). Therefore, the implementation of competition policy should take into consideration the merger of two firms that are heterogeneous in terms of behavior and strategic market power (see Huck et al., 2002). The second experiment reveals that mergers can both individually and socially beneficial with a few leaders (Daughety, 1990).

We now turn to the related question of free entry in the SOE.

3.4 Free Entry

3.4.1 Exogenous versus endogenous entry in the SOE

The multiple leader–follower game delivers an interesting setup to study entry. Two mechanisms of entry with first-mover advantage together with their welfare effects are studied in Etro (2006, 2008). Etro (2008) considers a three-stage game with complete and perfect information. In stage 1, the leaders enter, pay a fixed cost, and simultaneously determine their individual strategies. In stage 2, knowing the strategies of the leaders, the potential entrants determine whether to enter or not (if a firm decides to enter, it pays the fixed cost). In stage 3, the firms that decide to enter in stage 2 are followers and simultaneously determine their individual strategies. The associated subgame equilibrium concept is Stackelberg equilibrium with free entry.[20] When entry is exogenous, i.e., the number of followers is fixed, the leaders are mainly concerned with the reactions of the potential entrants (such reactions are opposite under strategic substitutability or strategic complementarity). Indeed, the leaders are more aggressive when strategies are substitutes and more accommodating when strategies are complements. Nevertheless, when entry is endogenous, the leaders are primarily concerned with the effects of their choices on the entry decision (Etro, 2008). The induced entry would

[20] Two questions are not addressed. First, the existence of a Stackelberg equilibrium with free entry. Second, the problem of determining whether a follower might become a leader by entering endogenously in the process of entry.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

reduce the profitability of leaders. Therefore, the accommodating behavior is ineffective and the subsequent aggressive behavior of leaders limits entry.[21]

The following example shows that in case of homogeneous product, Stackelberg competition with endogenous entry always leads to entry-deterrence with only leaders as market participants. In addition, the exogenous entry SOE and the SOE with endogenous entry and active followers Pareto dominate the COE (Etro, 2008).[22]

Example 12 (linear model continued) Consider (10.25)–(10.30). Assume now there is a fixed cost of entry $f > 0$.

SOE with exogenous entry: Here the game displays two stages of decisions described by equations (10.17)–(10.21). The entrants are the n_F followers. The SOE strategies are given by $\tilde{x}_L^i = X^* \frac{1}{n_L+1}, i \in \mathcal{F}_L$, and $\tilde{x}_F^j = X^* \frac{1}{(n_L+1)(n_F+1)}, j \in \mathcal{F}_F$. In addition, the SOE with exogenous entry Pareto dominates the COE. Indeed, as $\widetilde{X} = X^* \frac{n_L(n_F+1)+n_F}{(n_L+1)(n_F+1)} > \widehat{X} = X^* \frac{n_L+n_F}{n_L+n_F+1}$ (the SOE price is lower), the SOE surplus of consumers is higher. In addition, for each $i \in \mathcal{F}_L$, $\tilde{s}_L^i = \frac{n_F+1}{n_L(n_F+1)+n_F} > \frac{1}{n_L+n_F} = \hat{s}_L^i$. Finally, as the profits of followers are positive, and for any leader i, we have $\tilde{\pi}_L^i = \frac{(a-c)^2}{b} \frac{1}{(n_L+1)^2(n_F+1)} - f > \hat{\pi}_L^i = \frac{(a-c)^2}{b} \frac{1}{(n_L+n_F+1)^2} - f$, $i \in \mathcal{F}_L$, whenever $n_L < \sqrt{n_F+1}$: then, the social surplus in the SOE is higher than in the COE provided the number of leaders in not too high.

SOE with endogenous entry: Now the game displays three stages of decisions, and only a limited number of firms can enter and produce in equilibrium. Therefore, the number of active firms n is now endogenous. Given X_L any follower will supply $x_F^j = \frac{a-c-bX_L}{b(n_F+1)}, j \in \mathcal{F}_F$. Therefore, the payoff of follower j is given by $\pi_F^j = \frac{1}{b}\left(\frac{a-c-bX_L}{n_F+1}\right)^2 - f, j \in \mathcal{F}_F$. The zero-profit condition for entry, i.e., for all $j \in \mathcal{F}_F$, $\pi_F^j - f = 0$, leads to the optimal number of followers $\tilde{n}_F = \frac{a-c-bX_L}{\sqrt{bf}} - 1$, which decreases with the supplies of leaders. Then, the supply of follower j is $x_F^j = \frac{1}{b}\sqrt{bf}$, $j \in \mathcal{F}_F$. The effective demand is given by $p = c + \sqrt{bf}$, $j \in \mathcal{F}_F$. Then, in the first stage, as long as there are some follower entrants, i.e., for $\tilde{n}_F \geqslant 0$ or $\tilde{x}_L^i \leqslant \frac{a-c-\sqrt{bf}}{bn_L}$, the payoff of leader i is $\pi_L^i = b\sqrt{bf}x_L^i - f$, $i \in \mathcal{F}_L$, which increases with x_L^i, with x_L^i decreasing in n_F. If there is no entry, in which case $\tilde{n}_F = 0$, the SOE strategy, aggregate supply, market price and payoffs are given respectively by $\tilde{x}_L^i = \frac{1}{n_L}X^* - \frac{\sqrt{bf}}{bn_L}$, $i \in \mathcal{F}_L$, $\widetilde{X} = \widetilde{X}_L = X^* - \frac{1}{b}\sqrt{bf}$, $\tilde{p} = c + \sqrt{bf}$, and $\tilde{\pi}_L^i = \sqrt{bf}\frac{a-c-\sqrt{bf}}{bn_L} - f$, $i \in \mathcal{F}_L$. So $\tilde{\pi}_L^i = \sqrt{bf}\frac{a-c-\sqrt{bf}}{bn_L} - f > b\sqrt{bf}x_L^i - f$, $i \in \mathcal{F}_L$, $\forall \tilde{n}_F$: the best strategy of leaders is to prevent entry. In addition, the SOE with endogenous entry Pareto dominates the COE with endogenous entry. From the linear model subsection 2.2, we deduce that the COE profit with

[21] Potential entry is conditioned on the fundamentals of the model (the costs functions and the (im)perfect substitutability of commodities). Etro (2008) shows that the impact of endogenous entry is amplified in price competition models. Price quoted by leaders is higher under exogenous entry and lower under endogenous entry. In particular, under endogenous entry with logit demand or Dixit and Stiglitz (1977) demand, a leader sells its variety at a lower price, which "Pareto improves" the market outcome.

[22] This property is still true in the case of decreasing marginal costs. In the case of quadratic costs, the supply of any leader always exceeds the supply of any follower, but if the size of the market is not large enough (so fixed costs are small enough), i.e., the equilibrium number of firms is small, then there is an SOE with entry deterrence (see Etro, 2008).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

endogenous entry is $\hat{\pi}_L^i = \hat{\pi}_F^j = \frac{(a-c)^2}{b(n_L+n_F+1)^2} - f$, $i \in \mathcal{F}_L, j \in \mathcal{F}_F$, so the number of active firms is $\hat{n} = \frac{a-c}{\sqrt{bf}} - 1$. Then, we deduce the COE with endogenous entry, namely $\hat{x}_L^i = \hat{x}_F^j = \frac{1}{b}\sqrt{bf}$, $\widehat{X} = X^* - \frac{1}{b}\sqrt{bf}$, $\hat{p} = c + \sqrt{bf}$, $\hat{\pi}_L^i = \hat{\pi}_F^j = 0$, $i \in \mathcal{F}_L, j \in \mathcal{F}_F$. Indeed, we have $\widetilde{X} = \widehat{X}$ (and then $\hat{p} = \tilde{p}$), so the surplus of consumers is the same in both equilibria with endogenous entry. For each $i \in \mathcal{F}_L$, we have $\hat{s}_L^i = \frac{1}{X^*}\frac{1}{b}\sqrt{bf} - 1 < \frac{1}{n_L} = \tilde{s}_L^i$ (as $X^* > \frac{1}{b}\sqrt{bf}$). Then, the leaders are more aggressive under endogenous entry ($\tilde{x}_L^i > \hat{x}_L^i$, $i \in \mathcal{F}_L$). Then, leaders have higher payoffs as $\tilde{\pi}_L^i = \sqrt{bf}\frac{a-c-\sqrt{bf}}{bn_L} - f > 0$, $i \in \mathcal{F}_L$. Therefore, the social surplus increases ($\sqrt{bf}(a - c - \sqrt{bf}) - n_L f > 0$).

This example shows that when there is a low number of potential entrants, all followers are active, but when there are many potential entrants in a free-entry SOE, the only active firms are the leader firms. In addition, a SOE with free entry improves welfare since Stackelberg competition induces a reduction in the number of active firms (see Etro, 2008). The aggressive behavior of leaders may be either a result of market exploitation or the consequence of an increase in the competitiveness caused by the potential presence of new entrants (the followers). Therefore, the role of antitrust policy is not crystal clear (see Etro, 2008).

We now turn to the extended version of the SOE, which constitutes the generalized quantity-setting oligopoly game with homogeneous products.

3.4.2 The extended Stackelberg oligopoly game (T-stage game II)

Let us consider the T-stage decision linear game (Watt, 2002, Lafay, 2010, Julien, Musy, and Saïdi, 2012). There are T stages of decisions indexed by t, $t = 1, 2, \ldots, T$. Each stage embodies n_t firms, with $\sum_{t=1}^{T} n_t = n$. The full set of sequence of stages represents a hierarchy. The supply of firm i in stage t is denoted by x_t^i. The aggregate supply in stage t is given by $X_t \equiv \sum_{i=1}^{n_t} x_t^i$. The n_t firms behave as leaders (followers) with respect to all firms at stages $\tau > t$ ($\tau < t$). Let $p(X) = a - bX$, $a, b > 0$, where $X \equiv \sum_t X_t$. The costs functions are $C_t^i(x_t^i) = cx_t^i$, $i = 1, \ldots, n_t$, $t = 1, \ldots, T$, with $c < a$. The strategy sets are $\mathcal{S}_t^i = [0, X^*]$, $i = 1, \ldots, n_t$, $t = 1, \ldots, T$, where $X^* = \frac{a-c}{b}$ (with $p^* = c$ as the competitive price). The T-stage SOE is given by $\tilde{x}_t^i = X^* \prod\limits_{\tau=1}^{t} \frac{1}{n_\tau+1}$, $t = 1, \ldots, T$, $\tilde{p} = c + (a - c) \prod\limits_{\tau=1}^{t} \frac{1}{n_\tau+1}$ as $\widetilde{X} = X^* \sum_{t=1}^{T} n_t \prod_{\tau=1}^{t} \frac{1}{n_\tau+1}$, and $\tilde{\pi}_t^i = \frac{(a-c)^2}{b} \prod\limits_{\tau=1}^{t} \frac{1}{(n_\tau+1)^2} \prod\limits_{\tau=t+1}^{T} \frac{1}{n_\tau+1}$, $t = 1, \ldots, T-1$, and $\tilde{\pi}_T^i = \frac{(a-c)^2}{b} \prod\limits_{\tau=1}^{T} \frac{1}{(n_\tau+1)^2}$.

It is worth noticing that the T-stage Stackelberg linear economy reduces to a multi-stage Cournot game in which firms compete oligopolistically on the residual demands (see Julien et al., 2012). Therefore, each firm within a given stage behaves as if there were no following stages, i.e., as if its direct followers did not matter, generalizing the T-stage monopoly property of Boyer and Moreaux (1986).

The welfare properties are explored in Julien et al. (2011). To this end, they define ω, the index of social welfare $\widetilde{X}$, as $\omega = \prod_{\tau=1}^{T} \eta_t n_t = 1 - \prod_{\tau=1}^{T} \frac{1}{n_\tau+1} = 1 + \eta_{1,T}$. Then, they obtain two results. First, *when the number of firms becomes arbitrarily large, either vertically (when T tends to infinity) or horizontally (when n_t tends to infinity) the T-stage SOE aggregate supply*

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

converges toward the CE aggregate supply. Second, *the social welfare can be maximized either by enlarging the hierarchy or by changing the size of existing stages by relocating firms from the most populated stage until equalizing the size of all stages*. The relocation echoes the merger analysis provided in Daughety (1990). There are two implications. First, a sequential market structure with one firm per stage Pareto dominates any other market structure, among which the COE (Watt, 2002). Second, the firms' surplus may be dominated by the COE surplus when $T \geqslant 3$. Indeed, unlike the preceding games, the leaders might be better off if they are supplying simultaneously!

These results may be useful to analyze how entry effects affect welfare. Indeed, when a new firm enters in stage t it causes a decrease in market price as $\tilde{p}'(\sum_{t=1}^{t} \widetilde{X}_\tau + \tilde{x}_t^{n_t+1}) - \tilde{p}(\sum_{t=1}^{t} \widetilde{X}_\tau) = \frac{a-c}{n_t \prod_{\tau=1}^{T}(n_t+1)} < 0$. In addition, as in Daughety (1990), the maximization of welfare implies the most asymmetric distribution of market power. Nevetheless, if costs are different, entry is affected by some relocations or extensions (for instance by ranking the entry by costs differences as in Pal and Sarkar, 2001). Lafay (2010) uses a T-stage game in which firms enter at different times or have different commitment abilities (firms bear different constant marginal costs). *The game confirms the positive effect of entry on welfare*. But, the salient feature is how firms must forecast future entries in the market. *Indeed, asymmetric costs could make entry inefficient*. When the firm reasons backwards, and the price is lower when there is further entry, it enters the market provided its costs do not exceed the resulting market price.[23]

4 THE MULTIPLE LEADER–FOLLOWER BILATERAL MARKET GAME

The SDE and SOE games feature strategic interactions on the supply side of a market with perfect competition on the demand side. When thinking about the market participants, there are two shortcomings. First, the market demand is exogenously specified and not derived from preferences. Second, the market is asymmetric: some agents behave strategically, while others behave competitively.[24] This view corresponds to the Cournot tradition, but might miss some features regarding the working and the consequences of market power caused by imperfectly competitive behavior. Therefore, we now consider a class of games in which all agents, consumers and producers, have preferences for all commodities. Preferences are endogenized, and market demand stems from strategic behavior. In addition, we allow the producers to be consumers. Let us notice that the story that will be developed throughout this section is far from being independent from the preceding ones. In the SDE and SOE games the presence of two commodities was left implicit: the consumers held some units of numéraire they were willing to exchange against some units of the produced commodity. The buyers have a utility function (not necessarily quasi-linear) and they behave strategically.

[23] When constant marginal costs differ among firms, the price contribution to an additional entrant may not be negative since the strategies of all firms are modified when a firm no longer enters the market (see Lafay, 2010).

[24] Models of monopolistic competition study imperfect competition in a multi-market environment but each monopoly firm is strategically isolated; no strategic interactions exist (see Thisse and Ushchev, Chapter 5 in this volume). Otherwise, in games with differentiated commodities the market is asymmetric as the demand side is competitive and the supply side is strategic. See Appendix A.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

To model Stackelberg competition within interrelated markets we consider the bilateral oligopoly model.[25] This model features a market with two divisible commodities and a finite number of traders (Gabszewicz and Michel, 1997, Bloch and Ghosal, 1997, Bloch and Ferrer, 2001, Amir and Bloch, 2009, Dickson and Hartley, 2011). It constitutes an illustration with two commodities and corner endowments of the Shapley (1976)–Shubik (1973), Dubey and Shubik, (1978), Sahi and Yao (1989), and Amir et al. (1990) market games models (see Giraud, 2003). There are two types of traders: each type has a corner endowment on one good, but wants to consume both goods. There is a market price mechanism that connects both sides of the market. This mechanism captures strategic interactions within each side and between both sides of the market. A game is associated with this market in which the players are the traders, the strategies are their supply decisions, and the payoffs are the utility they reach for this market outcome. We extend the bilateral oligopoly model introduced by Gabszewicz and Michel (1997) by considering heterogeneous behavior. Thus, the multiple leader–follower model is now embedded in a complete strategic quantity-setting game.

We first describe the bilateral multiple leader–follower oligopoly game. Second, we characterize the optimal behavior and state the bilateral Stackelberg oligopoly equilibrium (BSOE thereafter) conditions. Third, we consider the market power and the welfare properties of the BSOE. To this end, we compare the SOE with the bilateral Cournot oligopoly equilibrium (BCOE) and with the CE. We also outline the main differences with the SDE and SOE games.

4.1 The Model

Consider an economy with two divisible homogeneous commodities labeled X and Y.[26] Let p_X and p_Y be the corresponding unit prices. We assume commodity X does not exist initially and must be produced. There are traders of two types, namely 1 and 2, such that the set of traders is partitioned into two subsets T_1 and T_2, with $T_1 \cap T_2 = \{\varnothing\}$. We assume $2 \leqslant |T_1| < \infty$ and $2 \leqslant |T_2| < \infty$, where $|T|$ denotes the cardinality of the set T. Traders who belong to T_1 (resp. T_2) are indexed by i (resp. by j). We assume there are n_L leaders and n_F followers of type I, with $T_1 = \{1, \ldots, n_L, n_L + 1, \ldots, n\}$. Similarly, $T_2 = \{1, \ldots, m_L, m_L + 1, \ldots, m\}$. We assume $n_L \geqslant 1$ (resp. $m_L \geqslant 1$) and $n_F \geqslant 1$ (resp. $m_F \geqslant 1$).

Commodity Y is spread among traders who belong to T_2: the endowment of trader j is denoted by ω^j, with $\omega^j > 0$, for all $j \in T_2$. Like in Gabszewicz and Michel (1997), traders of type 1 have no endowment but have inherited some technology that specifies how to produce some amount e^i of good X with some amount z^i of good Y. The production function of trader i is defined as $F_i : R_+ \to R_+$, such that $e^i = F_i(z^i)$, $i \in T_1$. We assume all F_i are continuously differentiable, strictly monotonic and concave (constant returns to scale). Therefore, like in Sections 2 and 3, for all $i \in T_1$, the cost function $p_Y F_i^{-1}(e^i)$ is convex. The utility function of any type 1 trader is $U^i : \mathbb{R}^2_+ \to \mathbb{R}$, $x^i \mapsto U^i(x^i, y^i)$, $i \in T_1$, while the utility function of any type 2 trader is $V^j : \mathbb{R}^2_+ \to \mathbb{R}$, $x^j \mapsto V^j(x^j, y^j)$, $j \in T_2$, where x and y represent the amount of

[25] The first transposition of Stackelberg competition in an L-commodity setting is provided in Julien and Tricou (2010) and in Julien (2013) in the context of a pure exchange economy with a finite number of traders, and in Julien and Tricou (2012) with a productive sector.

[26] We could consider good Y is commodity money. This interpretation is implicit in the Stackelberg games of Sections 2 and 3 in which each consumer is endowed with units of numéraire that she exchanges for the produced good.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

goods X and Y consumed. The utility functions are twice-continuously differentiable, strictly monotonic and strictly quasi-concave.

We associate now a noncooperative market game $\mathbf{\Gamma}$ with this economy. This game displays two stages of decisions and no discounting. The timing of positions is given. Information is assumed to be complete and perfect. Let $\mathcal{B}^i = \{b^i \in \mathbb{R}_+ : 0 \leqslant b^i \leqslant e^i\}$ be the strategy sets of trader $i \in T_1$. The quantity b^i denotes the pure strategy of any trader $i \in T_1$. The strategy b^i represents the amount of commodity X trader $i \in T_1$ sells. Likewise, let $\mathcal{Q}^j = \{q^j \in \mathbb{R}_+ : 0 \leqslant q^j \leqslant \omega^j\}$, where q^j is the pure strategy of trader $j \in T_2$. Therefore, trader i consumes the difference between the amount produced e^i and the amount sold b^i, i.e., $x^i = e^i - q^i$, $i \in T_1$. In addition, she buys the amount $y^i = \frac{\pi^i}{p_Y}$ of good Y, where π^i is her nominal profit, with $\pi^i\left(b^i, e^i\right) = p_X q^i - p_Y F_i^{-1}(e^i)$, $i \in T_1$. A strategy profile is represented by the vector $(\mathbf{b}; \mathbf{q}) = (b^1, b^2, \ldots, b^n; q^1, q^2, \ldots, q^m)$, with $(\mathbf{b}; \mathbf{q}) \in \prod_{i \in T_1} \mathcal{B}^i \times \prod_{j \in T_2} \mathcal{Q}^j$. Let $\mathbf{b}^{-i}$ $(\mathbf{q}^{-j})$ denote the strategy profile of all type 1 traders but i (j). In addition, let $\mathbf{b}_L$ and $\mathbf{b}_F$ be the strategy profiles of type 1 leaders and followers. We also consider $\mathbf{b}_L^{-i}$ and $\mathbf{b}_F^{-i}$. The same holds for $\mathbf{q}$, with $\mathbf{q} = (\mathbf{q}_L, \mathbf{q}_F)$ and $(\mathbf{q}_L^{-j}, \mathbf{q}_F^{-j})$.

There is a trading post that specifies the relative price at which exchange occurs. The relative price $\frac{p_X}{p_Y}$ is such that $p_X \mathbf{B} \equiv p_Y \sum_{i \in T_1} b^i$ must balance $p_Y \mathbf{Q} \equiv p_Y \sum_{j \in T_2} q^j$ (see Sahi and Yao, 1989). Therefore, given a price vector $\mathbf{p} = (p_X, p_Y)$ and a strategy profile $(\mathbf{b}; \mathbf{q})$, the market clearing price $\frac{p_X}{p_Y}(\mathbf{b}; \mathbf{q})$ obtains as:[27]

$$\frac{p_X}{p_Y}(\mathbf{b}; \mathbf{q}) = \frac{\mathbf{Q}}{\mathbf{B}}, \text{ if } \mathbf{B} > 0 \text{ and } \mathbf{Q} > 0. \tag{10.34}$$

So, after exchange, each trader obtains the following commodity bundle:

$$\left(x^i, y^i\right) = \left(e^i - b^i, \frac{\mathbf{Q}}{b^i + \mathbf{B}^{-i}} b^i - F_i^{-1}(e^i)\right), i \in T_1 \tag{10.35}$$

$$\left(x^j, y^j\right) = \left(\frac{\mathbf{B}}{q^j + \mathbf{Q}^{-j}} q^j, \omega^j - q^j\right), j \in T_2. \tag{10.36}$$

The corresponding utility levels may be written as payoffs:

$$\Pi^i(\mathbf{b}; \mathbf{q}) = U^i\left(e^i - b^i, \frac{\mathbf{Q}}{b^i + \mathbf{B}^{-i}} b^i - F_i^{-1}(e^i)\right), i \in T_1 \tag{10.37}$$

$$\Lambda^j(\mathbf{b}; \mathbf{q}) = V^j\left(\frac{\mathbf{B}}{q^j + \mathbf{Q}^{-j}} q^j, \omega^j - q^j\right), j \in T_2. \tag{10.38}$$

We now analyze the properties of this Stackelberg bilateral market game.

[27] Since we consider only active equilibria, i.e., market participation equilibria without autarky, we delete the case for which $\frac{p_X}{p_Y}(\mathbf{b}; \mathbf{q}) = 0$ if $\mathbf{B} = 0$ and/or $\mathbf{Q} = 0$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

4.2 Bilateral Stackelberg Oligopoly Equilibrium

We first define the BSOE of this market game. Second, we provide a characterization of the equilibrium conditions. Third, we study the welfare properties of the BSOE. To this end, we study the effects of strategic interactions on market power, which we compare with the games developed in Sections 2 and 3.

4.2.1 BSOE: definition

The BSOE is a noncooperative equilibrium of a two-step game where the players are the traders who belong to both sides of the market, the strategies are the quantity they send to the market for trade, and their payment is the utility levels they reach. To simplify we denote an allocation $(\mathbf{x}, \mathbf{y})$ by $\mathbf{a}$, with $(\tilde{x}^i, \tilde{y}^i) = \tilde{\mathbf{a}}^i$, for $i \in T_1$, and $(\tilde{x}^j, \tilde{y}^j) = \tilde{\mathbf{a}}^j$, for $j \in T_2$. In addition, like in Section 3, let $\mathbf{b}_F = \boldsymbol{\varphi}(\mathbf{b}_L; \mathbf{q}_L)$ be the vector of best responses, with component $b^i = \varphi^i(\mathbf{b}_L; \mathbf{q}_L)$, $i = n_L + 1, \ldots, n$. Similarly, let $\mathbf{q}_F = \boldsymbol{\sigma}(\mathbf{b}_L; \mathbf{q}_L)$, with component $q^j = \sigma^j(\mathbf{b}_L; \mathbf{q}_L), j = m_L + 1, \ldots, m$.

Definition 3 Given $\boldsymbol{\Gamma}$, a vector $((\tilde{\mathbf{b}}; \tilde{\mathbf{q}}); \tilde{\mathbf{a}})$, consisting of a strategy profile $(\tilde{\mathbf{b}}; \tilde{\mathbf{q}}) = (\tilde{b}^1, \ldots, \tilde{b}^n; \tilde{q}^1, \ldots, \tilde{q}^m)$ and an allocation $\tilde{\mathbf{a}}$ such that $\tilde{\mathbf{a}}^i(\tilde{b}^i, \frac{p_X}{p_Y}(\tilde{\mathbf{b}}; \tilde{\mathbf{q}}))$, for $i \in T_1$, and $\tilde{\mathbf{a}}^j(\tilde{q}^j, \frac{p_X}{p_Y}(\tilde{\mathbf{b}}; \tilde{\mathbf{q}}))$, for $j \in T_2$, constitutes a bilateral Stackelberg oligopoly equilibrium, with respect to a relative price $\frac{p_X}{p_Y}(\tilde{\mathbf{b}}; \tilde{\mathbf{q}})$, if:

$$\Pi^i(\tilde{\mathbf{a}}^i(\tilde{b}^i, \tfrac{p_X}{p_Y}(\tilde{b}^i, \tilde{\mathbf{b}}^{-i}; \tilde{\mathbf{q}}))) \geqslant \Pi^i(\mathbf{a}^i(b^i, \tfrac{p_X}{p_Y}(b^i, \tilde{\mathbf{b}}^{-i}; \tilde{\mathbf{q}}))), \forall b^i \in \mathcal{B}^i, i = n_L + 1, \ldots, n$$

$$\Lambda^j(\tilde{\mathbf{a}}^j(\tilde{q}^j, \tfrac{p_X}{p_Y}(\tilde{\mathbf{b}}; \tilde{q}^j, \tilde{\mathbf{q}}^{-j}))) \geqslant \Lambda^j(\mathbf{a}^j(q^j, \tfrac{p_X}{p_Y}(\tilde{\mathbf{b}}; q^j, \tilde{\mathbf{q}}^{-j}))), \forall q^j \in \mathcal{Q}^j, j = m_L + 1, \ldots, m$$

$$\Pi^i(\tilde{\mathbf{a}}^i(\tilde{b}^i, \tfrac{p_X}{p_Y}(\tilde{b}^i, \tilde{\mathbf{b}}_L^{-i}, \boldsymbol{\varphi}(\tilde{b}^i, \tilde{\mathbf{b}}_L^{-i}; \tilde{\mathbf{q}}_L); \tilde{\mathbf{q}}_L, \boldsymbol{\sigma}(\tilde{b}^i, \tilde{\mathbf{b}}_L^{-i}; \tilde{\mathbf{q}}_L)))) \geqslant$$

$$\Pi^i(\mathbf{a}^i(b^i, \tfrac{p_X}{p_Y}(b^i, \tilde{\mathbf{b}}_L^{-i}, \boldsymbol{\varphi}(b^i, \tilde{\mathbf{b}}_L^{-i}; \tilde{\mathbf{q}}_L); \tilde{\mathbf{q}}_L, \boldsymbol{\sigma}(b^i, \tilde{\mathbf{b}}_L^{-i}; \tilde{\mathbf{q}}_L)))), \ \forall \boldsymbol{\varphi}(.) \in \prod_{i>n_L} \mathcal{B}^i, \ \forall \boldsymbol{\sigma}(.) \in \prod_{j>m_L} \mathcal{Q}^j, \forall b^i \in \mathcal{B}^i, i = 1, \ldots, n_L$$

$$\Lambda^j(\tilde{\mathbf{a}}^j(\tilde{q}^j, \tfrac{p_X}{p_Y}(\tilde{\mathbf{b}}_L, \boldsymbol{\varphi}(\tilde{\mathbf{b}}_L; \tilde{q}^j, \tilde{\mathbf{q}}_L^{-j}); \tilde{q}^j, \tilde{\mathbf{q}}_L^{-j}, \boldsymbol{\sigma}(\tilde{\mathbf{b}}_L; \tilde{q}^j, \tilde{\mathbf{q}}_L^{-j})))) \geqslant$$

$$\Lambda^j(\mathbf{a}^j(q^j, \tfrac{p_X}{p_Y}(\tilde{\mathbf{b}}_L, \boldsymbol{\varphi}(\tilde{\mathbf{b}}_L; q^j, \tilde{\mathbf{q}}_L^{-j}); q^j, \tilde{\mathbf{q}}_L^{-j}, \boldsymbol{\sigma}(\tilde{\mathbf{b}}_L; q^j, \tilde{\mathbf{q}}_L^{-j})))), \ \forall \boldsymbol{\varphi}(.) \in \prod_{i>n_L} \mathcal{B}^i, \ \forall \boldsymbol{\sigma}(.) \in \prod_{j>m_L} \mathcal{Q}^j, \forall q^j \in \mathcal{Q}^j, j = 1, \ldots, m_L.$$

Thus, a BSOE is a $n + m$-tuple of strategies $(\tilde{\mathbf{b}}; \tilde{\mathbf{q}})$ chosen by the traders such that no trader has an advantage to deviate unilaterally from her choice. We here assume the existence and uniqueness of an active (nonautarkic) BSOE and we focus on the equilibrium properties of the game.[28]

[28] To avoid the trivial equilibrium, some restrictions are needed regarding the behavior of the indifference curves on the boundary of the commodity space ($\lim_{x^k \to 0} \frac{\partial u^k}{\partial x^k} = \lim_{y^k \to 0} \frac{\partial u^k}{\partial y^k} = \infty$). It implies that the indifference curves have no intersection with the axis in the quantity space. More generally, the existence of an active Cournot oligopoly equilibrium rises specific difficulties. Indeed, it is well known that the Cournot bilateral oligopoly model can have the trivial equilibrium as a possible outcome (see notably Cordella and Gabszewicz, 1998).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

4.2.2 BSOE: characterization

Let $s_X^i \equiv \frac{b^i}{\mathbf{B}}$ and $s_Y^j \equiv \frac{q^j}{\mathbf{Q}}$ as the market shares of trader i and trader j.

Proposition 6 *If the strategy profile* $(\tilde{\mathbf{b}}; \tilde{\mathbf{q}})$ *constitutes a BSOE, then:*

$$\frac{p_X}{p_Y}\left(1-(1+\nu_X)\tilde{s}_X^i+\eta_Y\frac{\tilde{b}^i}{\tilde{\mathbf{Q}}}\right)=\frac{dF_i^{-1}(e^i)}{de^i}_{|e^i=\tilde{e}^i},\ i=1,\ldots,n_L \tag{10.39}$$

$$\frac{p_X}{p_Y}\left(1-\tilde{s}_X^i\right)=\frac{dF_i^{-1}(e^i)}{de^i}_{|e^i=\tilde{e}^i},\ i=n_L+1,\ldots,n \tag{10.40}$$

$$\frac{p_Y}{p_X}\left(1-(1+\nu_Y)\tilde{s}_Y^j+\eta_X\frac{\tilde{q}^j}{\tilde{\mathbf{B}}}\right)=MRS_{Y/X}^j(\tilde{\mathbf{a}}^j),\ j=1,\ldots,m_L \tag{10.41}$$

$$\frac{p_Y}{p_X}\left(1-\tilde{s}_Y^j\right)=MRS_{Y/X}^j(\tilde{\mathbf{a}}^j),\ j=m_L+1,\ldots,m, \tag{10.42}$$

where $\nu_X = \frac{\partial\boldsymbol{\varphi}(\mathbf{b}_L;\mathbf{q}_L)}{\partial b^i}$, $\eta_Y = \frac{\partial\boldsymbol{\sigma}(\mathbf{b}_L;\mathbf{q}_L)}{\partial b^i}$, $\nu_Y = \frac{\partial\boldsymbol{\sigma}(\mathbf{b}_L;\mathbf{q}_L)}{\partial q^j}$, *and* $\eta_X = \frac{\partial\boldsymbol{\varphi}(\mathbf{b}_L;\mathbf{q}_L)}{\partial q^j}$ *represent the aggregate reactions of followers to a small change of each leader's strategy,* $\frac{dF_i^{-1}(e^i)}{de^i}$ *is the marginal cost of trader* $i \in T_1$*, and* $MRS_{Y/X}^j$ *is the marginal rate of substitution of good X for good Y of trader* $j \in T_2$.

Proof See Appendix B. ■

The conditions (10.39)–(10.42) state that, for each trader, the marginal revenue balances the rate of tradeoff between both commodities X and Y. In (10.39) and (10.40), the rate of tradeoff is the marginal cost, i.e., the marginal rate of transformation of good Y for good X, while in (10.41) and (10.42), it is the marginal rate of substitution between goods X and Y. A trader's real marginal revenue always depends upon her market share, which measures her weight with respect to the size of her market side. Is also depends upon the price elasticity(ies) of aggregate supply function(s), namely $\epsilon_X := \frac{d\log\mathbf{B}}{d\log(p_X/p_Y)} = -1$ and $\epsilon_Y := \frac{d\log\mathbf{Q}}{d\log(p_X/p_Y)} = 1$. Conditions (10.39) and (10.40) depend also upon the terms ν_X and ν_Y that feature strategic interactions between the leaders and the followers who belong to the same side of the market. The meaning of these terms receive the same interpretation as the one given in Sections 2 and 3. Indeed, when $\nu \geqslant 0$ ($\nu < 0$), the strategies are independent or complements (substitutes) in the same side of the market. But the terms η_X and η_Y feature (new) strategic interactions between the leaders and the followers who belong to opposite sides of the market. Indeed, when $\eta < 0$ ($\eta \geqslant 0$), the strategies are substitutes (complements) between both sides of the market. The complementarity stems from the substitutability between commodities (see Bloch and Ferrer, 2001 for the Cournot bilateral market). It is worth noticing that η has some connections with the parameter that measures the differentiation between commodities in the differentiated products game (see Appendix A).

Whilst these conditions share some similarities with those obtained in the games developed in Sections 2 and 3, they are derived from a two-commodity framework. Likewise, the price

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

clears the market, i.e., the aggregate strategic supply must be consistent with the market demand. But here the demand side (which includes the producers) is endogenous since it is derived from preferences. Then, this game explains how relative price is formed as the ratio of aggregate strategic supplies. Hence, the price is not obtained on the basis of a given market demand function. Therefore, the market outcome does not depend on specific assumptions regarding the demand function.

Remark 2 When $U^i(.) = U(.)$ and $F_i(.) = F(.)$, $i = 1, \ldots, n$, the (10.39)–(10.40) may be written as $\frac{p_X}{p_Y}\left(1 - \frac{1+\nu_X}{n_L} + \frac{\eta_Y}{n_L}\frac{p_Y}{p_X}\right) = \frac{dF_i^{-1}(e^i)}{de^i}$, $i < n_L+1$, $\frac{p_X}{p_Y}\left(1 - \frac{1}{n_F}\right) = \frac{dF_i^{-1}(e^i)}{de^i}$, $i \geqslant n_L+1$. The same holds for type 2 traders.

We now explore the welfare properties of the Stackelberg bilateral oligopoly model and we outline the market power at stake.

4.3 Market Power and Welfare

We now consider market power in the BSOE. More specifically, we compare the BSOE with the BCOE, and with the CE. Then, we consider the relation between market concentration and surplus, and also the relation between individual market power and payoffs. Do the welfare properties of the BSOE differ significantly from those of the SOE? Why or why not?

4.3.1 The BSOE, BCOE and CE aggregate market outcomes

It is possible to state the following propositions: the first extends the inefficiency property of single-industry games, while the second is specific to the multiple leader–follower noncooperative market game:

Proposition 7 *The noncooperative BSOE is inefficient.*

Proof Immediate from (10.39)–(10.42). For each pair (i,j), we get $(1 - (1 + \nu_X)\tilde{s}_X^i + \eta_Y \frac{\tilde{b}^i}{\tilde{\mathbf{Q}}})^{-1} \frac{dF_i^{-1}(e^i)}{de^i}{}_{|e^i=\tilde{e}^i} = (1 - \frac{\tilde{q}^j}{\tilde{\mathbf{Q}}})^{-1} \frac{\partial V^j/\partial x^j}{\partial V^j/\partial y^j}(\tilde{\mathbf{a}}^j)$, $i = 1, \ldots, n_L, j = m_L + 1, \ldots, m$. ■

The BSOE displays some market failures: the marginal rate of transformation of good Y into good X is not equal to the marginal rate of substitution between X and Y. Market power is caused by imperfectly competitive behavior: the market outcome within an industry is affected by the strategic interactions that prevail in the other industry (and vice versa)!

Proposition 8 *There is no Pareto domination between the BSOE and the BCOE aggregate market outcomes, i.e., the BSOE aggregate supplies (market price) are not necessarily higher (not necessarily lower) than the COE aggregate supplies (market price). But the CE aggregate supplies (market price) are higher (lower) than the BSOE and BCOE aggregate supplies.*

By contrast with the previous games, this market game illustrates that aggregate market outcomes cannot be Pareto ranked. Here every trader exerts her market power by restricting the amount of the good brought to the market. To highlight the difference with partial games, recall that commodity Y may be viewed as commodity money (Shapley and Shubik, 1977).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

The market price is determined as the ratio of total bids of traders who buy over total supply of traders who produce but do not sell all their production. Then, as all traders behave strategically, a lower price does not mean higher aggregate supply! Thus, unlike Stackelberg competition equilibria with one homogeneous product or two differentiated products, no general conclusion emerges when strategic interactions occur between (at least) two industries. Nevertheless, we state some connection with the partial framework by looking at two scenarios: first, the traders of type 2 behave à la Cournot, and second, they behave competitively. The next example illustrates these features:

Example 13 Consider $|T_1| = |T_2| = 2$, $\omega^j = (0,1)$, $j = 1,2$, $U^i(x^i, y^i) = \left((x^i)^\rho + (y^i)^\rho\right)^{\frac{1}{\rho}}$, $\rho \leqslant 1$, $i = 1,2$, $V^j(x^j, y^j) = x^j.y^j$, $j = 1,2$, $F_1(z^1) = \frac{1}{\beta}z^1$, $\beta \in (0,1)$ and $F_2(z^2) = \frac{1}{\gamma}z^2$, $\gamma \in (0,1)$. The CE is given by $(\frac{\hat{p}_X}{p_Y})^* = \min\{\beta,\gamma\}$, $(e^i)^* = \frac{1}{\min\{\beta,\gamma\}}$, $(\pi^i)^* = (\Pi^i)^* = 0$, $i = 1,2$, and $(\Lambda^j)^* = \frac{1}{4}\frac{1}{\min\{\beta,\gamma\}}$, $j = 1,2$. The BCOE is given by $(\hat{b}^1, \hat{e}^1) = \frac{2}{3}\frac{\gamma}{\beta+\gamma}(\frac{1}{\beta+\gamma}, \frac{\frac{1}{\beta+\gamma}+\left(\frac{1}{\beta}\right)^{\frac{1}{1-\rho}}}{1+\left(\frac{1}{\beta}\right)^{\frac{\rho}{1-\rho}}})$, $(\hat{b}^2, \hat{e}^2) = \frac{2}{3}\frac{\beta}{\beta+\gamma}(\frac{1}{\beta+\gamma}, \frac{\frac{1}{\beta+\gamma}+\left(\frac{1}{\gamma}\right)^{\frac{1}{1-\rho}}}{1+\left(\frac{1}{\gamma}\right)^{\frac{\rho}{1-\rho}}})$, and $\hat{q}^j = \frac{1}{3}$, $j = 1,2$, and $\left(\frac{\hat{p}_X}{p_Y}\right) = \beta + \gamma$. The computation of the BSOE strategies leads to $(\tilde{b}^1, \tilde{e}^1) = \frac{\sqrt{2(31+\sqrt{97})}}{12}\frac{1}{2}\frac{\gamma}{\beta}(\frac{1}{2\beta}, \frac{\frac{1}{2\beta}+\left(\frac{1}{\beta}\right)^{\frac{1}{1-\rho}}}{1+\left(\frac{1}{\beta}\right)^{\frac{\rho}{1-\rho}}})$, $(\tilde{b}^2, \tilde{e}^2) = \frac{\sqrt{2(31+\sqrt{97})}}{12}(1 - \frac{1}{2}\frac{\gamma}{\beta})(\frac{1}{2\beta}, \frac{\frac{1}{2\beta}+\left(\frac{1}{\gamma}\right)^{\frac{1}{1-\rho}}}{1+\left(\frac{1}{\gamma}\right)^{\frac{\rho}{1-\rho}}})$, and $(\tilde{q}^1, \tilde{q}^2) = (\frac{\sqrt{97}-5}{12}, \frac{5+\sqrt{2(31+\sqrt{97})}-\sqrt{97}}{12})$. So we deduce $\tilde{\mathbf{B}} = \frac{\sqrt{2(31+\sqrt{97})}}{12}\frac{1}{2\beta}$and $\tilde{\mathbf{Q}} = \frac{\sqrt{2(31+\sqrt{97})}}{12}$, then $(\frac{\tilde{p}_1}{p_2}) = 2\beta$. Therefore, if $\beta \leqslant \gamma$, then $(\frac{\hat{p}_X}{p_Y})^* < (\frac{\tilde{p}_X}{p_Y}) \leqslant (\frac{\hat{p}_X}{p_Y})$, whilst if $\beta > \gamma$, then, $(\frac{\hat{p}_X}{p_Y})^* < (\frac{\hat{p}_X}{p_Y}) < (\frac{\tilde{p}_X}{p_Y})$. In addition, $\mathbf{Q}^* > \hat{\mathbf{Q}} > \tilde{\mathbf{Q}}$, and $\mathbf{B}^* > \tilde{\mathbf{B}}$ and $\mathbf{B}^* > \hat{\mathbf{B}}$. And, if $\beta = \gamma$, then $\tilde{\mathbf{B}} > \hat{\mathbf{B}}$. But when $\beta \neq \gamma$ things are more complicated. Indeed, if $\beta < \gamma$, with $\gamma = a\beta$, $a > 1$, then $\tilde{\mathbf{B}} > \hat{\mathbf{B}}$ whenever $\beta < \frac{1}{16}\sqrt{2(31+\sqrt{97})}\frac{1+a}{a}$, while if $\beta > \gamma$, with $\beta = b\gamma$, $b > 1$, then $\tilde{\mathbf{B}} > \hat{\mathbf{B}}$ whenever $\beta < \frac{1}{16}\sqrt{2(31+\sqrt{97})}(1+b)$. Now, assume that *both traders of type 2 behave as followers*. Then, the best response of type 1 follower is the same. The best responses of type 2 followers are given by $\sigma^1(q^2) = -q^2 + \sqrt{(q^2)^2 + q^2}$ and $\sigma^2(q^1) = -q^1 + \sqrt{(q^1)^2 + q^1}$. Some computations yield $(\check{b}^1, \check{e}^1) = \frac{1}{3}\frac{\gamma}{\beta}(\frac{1}{2\beta}, \frac{\frac{1}{2\beta}+\left(\frac{1}{\beta}\right)^{\frac{1}{1-\rho}}}{1+\left(\frac{1}{\beta}\right)^{\frac{\rho}{1-\rho}}})$, $(\check{b}^2, \check{e}^2) = \frac{2}{3}(1 - \frac{1}{2}\frac{\gamma}{\beta})(\frac{1}{2\beta}, \frac{\frac{1}{2\beta}+\left(\frac{1}{\gamma}\right)^{\frac{1}{1-\rho}}}{1+\left(\frac{1}{\gamma}\right)^{\frac{\rho}{1-\rho}}})$, and $\check{q}^j = \frac{1}{3}$, $j = 1,2$. Then, we get $\check{\mathbf{B}} = \frac{1}{3\beta}$ and $\check{\mathbf{Q}} = \frac{2}{3}$, so $(\frac{\check{p}_X}{p_Y}) = 2\beta$. Then, $(\frac{\check{p}_X}{p_Y}) = (\frac{\tilde{p}_X}{p_Y})$. Finally, assume that *both traders of type 2 behave as price-takers*. We deduce $\bar{\mathbf{B}} = \frac{1}{2\beta}$ and $\bar{\mathbf{Q}} = 1$, so $(\frac{\bar{p}_X}{p_Y}) = 2\beta$. Therefore, $(\frac{\bar{p}_X}{p_Y}) = (\frac{\tilde{p}_X}{p_Y})$. Then, $(\frac{\hat{p}_X}{p_Y})^* < (\frac{\check{p}_X}{p_Y}) = (\frac{\bar{p}_X}{p_Y}) = (\frac{\tilde{p}_X}{p_Y})$. In addition, $\mathbf{Q}^* = \bar{\mathbf{Q}} > \check{\mathbf{Q}} > \tilde{\mathbf{Q}}$ and $\tilde{\mathbf{B}} < \check{\mathbf{B}} < \bar{\mathbf{B}} < \mathbf{B}^*$.

Therefore, in a BSOE the aggregate supply is lowest than in other strategic equilibria: the supply of good X increases when the market power of the agents endowed with commodity

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Y dilutes. The rate of tradeoff becomes more disadvantageous for these traders, who must consequently "spend more money" to consume the same amount of good X. When all traders endowed with commodity Y behave as followers, the corresponding market outcome results in the Stackelberg-Cournot equilibrium. But when they behave as price-takers the market outcome results in the Stackelberg-Walras equilibrium (see Julien and Tricou, 2012). Such a market structure is implicit in the games of Sections 2 and 3. The central message is as follows: *if one side of the market embodies strategic traders while the other side includes only price-taking traders, the BSOE aggregate market outcomes mimic the SDE and SOE aggregate market outcomes of the partial industry models.*[29]

4.3.2 Market power

The notion of surplus makes no sense here. In addition, a measure of concentration is difficult to capture as all traders behave strategically. We rather consider the Lerner indexes: $L_L^i \equiv \frac{\frac{p_X}{p_Y} - \frac{dF_i^{-1}(e^i)}{de^i}}{\frac{p_X}{p_Y}}$, $i = 1, \ldots, n_L$, $L_F^i \equiv \frac{\frac{p_X}{p_Y} - \frac{dF_i^{-1}(e^i)}{de^i}}{\frac{p_X}{p_Y}}$, $i = n_L + 1, \ldots, n$, $L_L^j \equiv \frac{\frac{p_Y}{p_X} - MRS_{Y/X}^j}{\frac{p_Y}{p_X}}$, $j = 1, \ldots, m_F$, and $L_F^j \equiv \frac{\frac{p_Y}{p_X} - MRS_{Y/X}^j}{\frac{p_Y}{p_X}}$, $j = m_F + 1, \ldots, m$. The following corollary echoes (10.13)–(10.14) and (10.33).

Corollary 2 *Let L_X^i, $i \in T_1$, and L_Y^j, $j \in T_2$, be the Lerner indexes. Then:*

$$L_L^i = (1 + \nu_X)\tilde{s}_X^i - \eta_Y \frac{\tilde{b}^i}{\tilde{\mathbf{Q}}}, i = 1, \ldots, n_L \tag{10.43}$$

$$L_F^i = \tilde{s}_X^i, i = n_L + 1, \ldots, n \tag{10.44}$$

$$L_L^j = (1 + \nu_Y)\tilde{s}_Y^j - \eta_X \frac{\tilde{q}^j}{\tilde{\mathbf{B}}}, j = 1, \ldots, m_L \tag{10.45}$$

$$L_F^j = \tilde{s}_Y^j, j = m_L + 1, \ldots, m. \tag{10.46}$$

Proof Immediate from (10.39)–(10.42) and by using the Lerner indexes. ■

These conditions show that strategic behavior of any trader consists in contracting her supply to manipulate the rate of exchange.

Remark 3 The markups are given by $\kappa_L^i = \frac{1}{1+(1+\nu_X)s_X^1 - \eta_Y \frac{b^1}{\mathbf{Q}}} - 1, i = 1, \ldots, n_L, \kappa_F^i = \frac{1}{1+s_X^1} - 1$, $i = n_L + 1, \ldots, n$. Similar expressions hold for all $j \in T_2$.

Proposition 9 *Let L_L^i, $i = 1, \ldots, n_L$, and L_L^j, $j = 1, \ldots, m_F$. Then, $\frac{\partial L_L^i}{\partial b^i} > 0$, $i = 1, \ldots, n_L$, whenever $\nu_X \leqslant 0$ and $\eta_Y < 0$; and, $\frac{\partial L_L^i}{\partial q^j} > 0, j = 1, \ldots, m_L$, whenever $\eta_X < 0$ and $\eta_Y > 0$. In*

[29] An alternative, but equivalent way to model such asymmetric positions in markets consists of considering Stackelberg competition with a mixed measure space of traders (Julien, 2012).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

addition, $\frac{\partial L_L^j}{\partial q^j} > 0, j = 1, \ldots, m_L$, *whenever* $\nu_Y \leqslant 0$ *and* $\eta_X < 0$*; and,* $\frac{\partial L_L^j}{\partial b^i} > 0$, $i = 1, \ldots, n_L$, *whenever* $\eta_X > 0$ *and* $\eta_Y < 0$.

Proof Immediate from (10.39) as $\frac{\partial L_L^i}{\partial b^i} = \frac{(1+\nu_X)}{\mathbf{B}}[1-(1+\nu_X)s_X^i] - \eta_Y \frac{1}{\mathbf{Q}} + \eta_Y^2 \frac{b^i}{\mathbf{Q}^2}$ and $\frac{\partial L_L^i}{\partial q^j} = -(1+\nu_X)\eta_X \frac{b^i}{\mathbf{B}^2} + \eta_Y(1+\nu_Y)\frac{b^i}{\mathbf{Q}^2}$. The same holds with (10.41). ■

Thus, any leader's market power increases with her strategy when strategies are substitutes in both industries (the other followers decrease their supply). In addition, her market power increases with the strategy of any other leader type when the strategies of type 2 traders are substitutes, and the strategies of leaders and of followers of two different types are complements. Again these effects depend critically on preferences. As the market power is relative, we give the definition:

Definition 4 Let $l^i = \frac{L_L^i}{L_F^i}$, $l^j = \frac{L_L^j}{L_F^j}$, $l^{ij} = \frac{L_L^i}{L_L^j}$ be the relative Lerner indexes.

Proposition 10 *Assume* $U^i(.) = U(.)$ *and* $V^j(.) = V(.)$. *Then:*

1. $l^i \gtreqless 1$ *iff* $\left(\frac{dF_i^{-1}(e^i)}{de^i}_{|e^i=\tilde{e}^i}\right)_{i<n_L+1} \lesseqgtr \left(\frac{dF_i^{-1}(e^i)}{de^i}_{|e^i=\tilde{e}^i}\right)_{i\geqslant n_{L+1}}$, *and;*

2. $l^j \gtreqless 1$ *iff* $\left(MRS^j(\tilde{\mathbf{a}}^j)\right)_{j<m_L+1} \lesseqgtr \left(MRS^j(\tilde{\mathbf{a}}^j)\right)_{j\geqslant m_L+1}$.

Proof Immediate from the definitions of the Lerner indexes. ■

The market power of any trader is higher when the ratio of sacrifice (either in production or in consumption) is lower. The market power depends not only on the number of traders, but also on preferences and/or endowments. Thus, industries are "connected" by preferences.

Remark 4 Consider $l^i = 1$ (or $l^j = 1$). Then, unlike the games of Sections 2 and 3 with constant identical marginal costs, the equilibrium strategies of leaders and followers are equal when costs are identical! Here the production decisions, and thereby the costs, are determined by preferences.

Corollary 3 *Assume* $U^i(.) = U(.)$ *and* $V^j(.) = V(.)$. *Then:*

1. $l^i = 1$ *iff* $(\tilde{b}^i)_{i<n_L+1} = (\tilde{b}^i)_{i\geqslant n_{L+1}}$, *and;*

2. $l^j = 1$ *iff* $\left(\tilde{q}^j\right)_{j<m_L+1} = \left(\tilde{q}^j\right)_{j\geqslant m_L+1}$.

Proof Immediate from (10.43)–(10.46). ■

Whilst the first implication always holds for isolated industries, the second no longer holds: equilibrium strategies may differ when marginal costs are equal!

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Proposition 11 *Assume all traders $i \in T_1$ (resp. $j \in T_2$) have the same utility function and the same technology (resp. the same endowments), and $\eta_X = \eta_Y = 0$. Then, $l^i = 1$ (resp. $l^j = 1$) if and only if $\nu_X = 0$ (resp. $\nu_Y = 0$). In particular, for each $i \in T_1$, $\tilde{b}^i = \hat{b}^i$, and for each $j \in T_2$, $\tilde{q}^j = \hat{q}^j$, where $\hat{b}^i$ and $\hat{q}^j$ are the BCOE strategies of trader i and trader j respectively.*

Proof Immediate by using (10.39)–(10.42), Corollary 3 and Remark 3. ■

This equivalence between the BSOE and the BCOE mimics the equivalences obtained in Propositions 2 and 5.

Proposition 12 *In a BSOE, if $\eta_X = 0$ and $\nu_X < 0$, then $\frac{\partial l^{ij}}{\partial b^i} > 0$, $i < n_L + 1$, $j < m_L + 1$. In addition, if $\eta_Y = 0$ and $\nu_Y < 0$, then $\frac{\partial l^{ij}}{\partial q^j} > 0$, $i < n_L + 1$, $j < m_L + 1$.*

Proof As $l^{ij} = \frac{b^i}{q^j} \frac{(1+\nu_X)\frac{p_X}{p_Y} - \eta_Y}{(1+\nu_Y) - \eta_X \frac{p_X}{p_Y}}$, then $\frac{l^{ij}}{\partial b^i} = \frac{1}{b^i} l^{ij} [1 - s^i \frac{(1+\nu_X)(1+\nu_Y) - \eta_X \eta_Y}{(1+\nu_Y) - \eta_X \frac{p_X}{p_Y}}]$. If $\eta_X = 0$ and $\nu_X < 0$, then $\frac{l^{ij}}{\partial b^i} = \frac{1}{q^j} \frac{(1+\nu_X)\frac{p_X}{p_Y} - \eta_Y}{(1+\nu_Y)} \left[1 - (1+\nu_X)s^i\right] > 0$. ■

Therefore, the relative market power of any leader increases with her strategy (the strategy of any type 2 leader) when the strategies are substitutes within each side of the market. Thus, it extends the properties of the partial model when the behavior of buyers and sellers is strategic.

4.3.3 Market power and payoffs

We can now state the following proposition that holds in a BSOE:

Proposition 13 *Assume that all traders have the same utility function. Then:*

1. $(\tilde{\Pi}^i)_{i<n_L+1} \gtreqless (\tilde{\Pi}^i)_{i \geqslant n_L+1}$ *iff* $\left(\frac{dF_i^{-1}(e^i)}{de^i}_{|e^i = \tilde{e}^i}\right)_{i<n_L+1} \lesseqgtr \left(\frac{dF_i^{-1}(e^i)}{de^i}_{|e^i = \tilde{e}^i}\right)_{i \geqslant n_L+1}$;
2. $(\tilde{\Lambda}^j)_{j<m_L+1} = (\tilde{\Lambda}^j)_{j \geqslant m_L+1}$ *iff* $\left(MRS^j(\tilde{a}^j)\right)_{j<m_L+1} = \left(MRS^j(\tilde{a}^j)\right)_{j \geqslant m_L+1}$.

Part 1 says that the traders who bear the lower (higher) cost reach higher (lower) payoffs. But part 2 says that a leader can have higher Lerner index and less payoffs even if they have the same endowments!

Corollary 4 *Assume that all traders have the same utility function. Then:*

1. $(\tilde{\Pi}^i)_{i<n_L+1} \gtreqless (\tilde{\Pi}^i)_{i \geqslant n_L+1}$ *iff* $l^i \gtreqless 1$, *and;*
2. $(\tilde{\Lambda}^j)_{j<m_L+1} = (\tilde{\Lambda}^j)_{j \geqslant m_L+1}$ *iff* $l^j = 1$.

Let us notice that Proposition 2 and Corollary 1 stated in Section 2 no longer hold when $l^i = 1$ or $l^j = 1$: the reason stems from the fact that in this game the real opportunity costs (production costs and marginal rates of substitution) are determined by preferences. Here equal market power means equal payoff.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

4.4 Free Entry

We investigate now the conditions under which more competition would affect market power. The BSOE market game displays some inefficiency caused by strategic interactions. Indeed, failure of optimality stems from the fact that each trader restricts the quantity she sends to the market. We would like to know whether market performance is linked to the thickness of the market. Hence, for a fixed number of traders, under which conditions does the BSOE coincide with the CE? More specifically, under exogenous free entry, does the sequence of bilateral Stackelberg equilibria converge asymptotically toward the CE?

4.4.1 When does strategic behavior become competitive?

First, under which conditions does the BSOE coincide with the CE?

Proposition 14 *If $\nu_X = \nu_Y = -1$ and $\eta_X = \eta_Y = 0$, then the BSOE coincides with the CE.*

Proof Consider (10.43) and (10.46). If $\nu_X = \nu_Y = -1$ and $\eta_X = \eta_Y = 0$, then $L_L^i = L_L^j = 0$, which from (10.39) and (10.42) means $\frac{p_X}{p_Y} = \frac{dF_i^{-1}(e^i)}{de^i}$, $i = 1, \ldots, n_L$, and $\frac{p_Y}{p_X} = MRS^j_{Y/X}$, $j = 1, \ldots, m_L$. In addition, if $\nu_X = \nu_Y = -1$, then $L_L^i = 0$, $i = n_L + 1, \ldots, n$, and $L_L^j = 0$, $j = m_L + 1, \ldots, m$. ■

When $\nu_X = \nu_Y = -1$, aggregate best responses have slopes equal to -1: any change in the strategies of leaders is exactly compensated by an equal reduction of followers' aggregate supply, which leaves the relative price unchanged (any increase is entirely "absorbed" by the rivals). In addition, when $\eta_X = \eta_Y = 0$, there is no strategic interaction between both sides of the market: both industries are "strategically" isolated. This result extends a result that holds in single-industry models (see Dixit, 1986).

4.4.2 Free entry: a replication exercise

We now study exogenous free entry. To this end, the economy is enlarged in such a way the number of leaders increases. Let r be an integer, with $r \geq 1$. The new bilateral market game $\mathbf{\Gamma}^r$ now includes rn_L leaders of type 1, each being indexed by ik, $i = 1, \ldots, n_L$, $k = 1, \ldots, r$, and rm_L leaders of type 2, each being indexed by jk, $j = 1, \ldots, m_L$, $k = 1, \ldots, r$. Therefore, there are now $rn_L + n_F$ traders of type 1 and $rm_L + m_F$ traders of type 2. Let $U^{ik}(x^{ik}, y^{ik})$ and $V^{jk}(x^{jk}, y^{jk})$ the utility functions of leader ik of type 1 and of leader jk of type 2 respectively. In addition, $e^{ik} = F_{ik}(z^{ik})$, $i = 1, \ldots, n_L$, $k = 1, \ldots, r$, and $e^i = F_i(z^i)$, $i = rn_L + 1, \ldots, rn_L + n_F$. Finally, $\mathbf{w}^{jk} = (0, 1)$, $j = 1, \ldots, m_L$, $k = 1, \ldots, r$, and $\mathbf{w}^j = (0, 1)$, $j = rm_L + 1, \ldots, rm_L + m_F$.

Given a strategy profile $(\mathbf{b}^r; \mathbf{q}^r)$, the market clearing price $\frac{p_X}{p_Y}(\mathbf{b}^r; \mathbf{q}^r)$ obtains as $\frac{p_X}{p_Y}(\mathbf{b}^r; \mathbf{q}^r) = \frac{\sum_{k=1}^r \sum_{j=1}^{m_L} q^{jk} + \sum_{j=rm_L+1}^{rm_L+m_F} q^j}{\sum_{k=1}^r \sum_{i=1}^{n_L} b^{ik} + \sum_{i=rn_L+1}^{rn_L+n_F} b^i}$. Following the same steps as in Appendix B and using Remark 2, we deduce:

$$\frac{p_X}{p_Y}\left(1 - (1 + \nu_X)\tilde{s}_X^{ik}(r) + \eta_Y \frac{\tilde{\mathbf{B}}(r)}{\tilde{\mathbf{Q}}(r)} s_X^{ik}(r)\right) = \frac{dF_{ik}^{-1}(e^{ik})}{de^{ik}}, i < n_L + 1, k \geqslant 1 \tag{10.47}$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

$$\frac{p_X}{p_Y}\left(1-\tilde{s}_X^i(r)\right)=\frac{dF_i^{-1}(e^i)}{de^i}, i=rn_L+1,\ldots,rn_L+n_F \quad (10.48)$$

$$\frac{p_Y}{p_X}\left(1-(1+\nu_Y)\tilde{s}_Y^{jk}(r)+\eta_X\frac{\tilde{\mathbf{Q}}(r)}{\tilde{\mathbf{B}}(r)}s_Y^{jk}(r)\right)=MRS_{Y/X}^{jk}, j<m_L+1, k\geqslant 1 \quad (10.49)$$

$$\frac{p_Y}{p_X}(1-\tilde{s}_Y^j(r))=MRS_{Y/X}^j, j=rm_L+1,\ldots,rm_L+m_F. \quad (10.50)$$

Proposition 15 *Assume: (H1) both commodities are normal for all traders and all the utility functions satisfy the gross substitutes property; and, (H2) all leaders and all followers of each type are identical. Then, when the number of leaders becomes arbitrarily large the BSOE of the replicated economy converges to the CE.*

Proof Consider (10.47)–(10.50) under (H2). We have $s_X^{ik}(r)=\frac{b^{ik}}{\mathbf{B}_L}\frac{\mathbf{B}_L}{\mathbf{B}}$, where $\mathbf{B}_L\equiv\sum_{k=1}^r\sum_{i=1}^{n_L}b^{ik}$, and $s_Y^{jk}(r)\equiv\frac{q^{jk}}{\mathbf{Q}_L}\frac{\mathbf{Q}_L}{\mathbf{Q}}$, where $\mathbf{Q}_L\equiv\sum_{k=1}^r\sum_{j=1}^{m_L}q^{jk}$. Using Remark 2, we get for the leaders $\lim_{r\to\infty}(1+\nu_X)s_X^{ik}(r)=\lim_{r\to\infty}\frac{1+\nu_X}{r}\frac{\tilde{\mathbf{B}}_L(r)}{\tilde{\mathbf{B}}(r)}=0$ as $0<\frac{\tilde{\mathbf{B}}_L(r)}{\tilde{\mathbf{B}}(r)}<1$, and $\lim_{r\to\infty}(1+\nu_Y)s_Y^{jk}(r)=\lim_{r\to\infty}\frac{1+\nu_Y}{r}\frac{\tilde{\mathbf{Q}}_L(r)}{\tilde{\mathbf{Q}}(r)}=0$ as $0<\frac{\tilde{\mathbf{Q}}_L(r)}{\tilde{\mathbf{Q}}(r)}<1$. In addition, $\lim_{r\to\infty}\eta_Y\frac{\tilde{\mathbf{B}}(r)}{\tilde{\mathbf{Q}}(r)}s_X^{ik}(r)=\lim_{r\to\infty}\frac{\eta_Y}{r}\frac{\tilde{\mathbf{B}}_L}{\tilde{\mathbf{B}}(r)}\frac{\tilde{\mathbf{B}}(r)}{\tilde{\mathbf{Q}}(r)}=0$ as $\lim_{r\to\infty}\frac{\tilde{\mathbf{Q}}(r)}{\tilde{\mathbf{B}}(r)}=(\frac{p_X}{p_Y})^*$, by using Proposition 3 in Amir and Bloch (2009). Indeed, under (H1), both $\tilde{\mathbf{B}}(r)$ and $\tilde{\mathbf{Q}}(r)$ increase with r, and converge monotonically to the CE. For the followers, we have $\lim_{r\to\infty}s_X^{ik}(r)=\lim_{r\to\infty}\frac{1}{r}\frac{\tilde{\mathbf{B}}_L(r)}{\tilde{\mathbf{B}}(r)}=0$ and $\lim_{r\to\infty}s_Y^{jk}(r)=\lim_{r\to\infty}\frac{1}{r}\frac{\tilde{\mathbf{Q}}_L(r)}{\tilde{\mathbf{Q}}(r)}=0$. Then, (10.47)–(10.50) may now be written $\frac{p_X}{p_Y}=\frac{dF_{ik}^{-1}(e^{ik})}{de^{ik}}, i<n_L+1, k\geqslant 1$, $\frac{p_X}{p_Y}=\frac{dF_i^{-1}(e^i)}{de^i}, i=rn_L+1,\ldots,rn_L+n_F$; and, $\frac{p_Y}{p_X}=MRS_{Y/X}^{jk}, j<m_L+1, k\geqslant 1$, and $\frac{p_Y}{p_X}=MRS_{Y/X}^j, j=rm_L+1,\ldots,rm_L+m_F$. ■

When the economy is replicated an infinite number of times the BSOE market outcome converges to the CE. This result may be explained as follows. When commodities are substitutes, for any given supply of good Y, an increase in the supply of good X by leaders decreases the relative price, and increases the purchasing power of the traders of type 2. Under normality of good, they increase their demand for commodity X and, by the gross substitutability property of the utility function, they substitute quantities of good Y for quantities of good X. Then, it puts up the supply of commodity Y. This effect is also based on the fact that the strategies of leaders and followers within each side of the market are strategic complements, in which case the aggregate best responses increase. These effects are similar to those described in Amir and Bloch (2009) who consider a bilateral oligopoly game with Cournot competition. They show that under (H1) and (H2) the equilibrium of the market game converges monotonically to the CE. On the other hand Dickson (2013) studies the conditions under which entry by new sellers raises the Cournot-Nash equilibrium payoffs of existing sellers. When the demand is sufficiently elastic, sellers with large enough market shares leads to profit-increasing competition. But here the main difference stems from the fact that the effects of entry are also driven by the followers' reactions. Such results confirm the point that the market demand must be endogenized since preferences play a critical role.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

5 CONCLUSIONS

This inquiry into Stackelberg competition under quantity-setting two-stage games with perfect and complete information reveals the following:

1. In basic duopoly games, the SDE social surplus is higher (lower) than the CDE (CE) social surplus. In addition, endogenous timing and experimental games suggest that the noncooperative SDE can emerge as a plausible market outcome. Finally, free entry may lead to the competitive market outcome.
2. In the multiple leader–follower games, the welfare property stated in (1) holds, but no longer holds in multi-stage decision settings, when the number of stages is greater than two. In addition, the asymmetry between firms can limit market inefficiencies, so welfare is not always correlated with concentration. Moreover, merging between firms who belong to two distinct cohorts may be welfare enhancing. Finally, the Stackelberg equilibria with exogenous entry and endogenous entry Pareto dominate the Cournot equilibria with entry.
3. In the two-stage multiple leader–follower bilateral market game there is no Pareto ranking between market outcomes. Thus, the study of market performance is more complex. Production decision and strategic behavior depend critically upon preferences. Unlike the differentiated products games in which prices are based on some specific utility function, a market-clearing mechanism determines the relative price. In addition, market power is relative, and some features are not captured in single-industry games. Indeed, the leaders and the followers reach the same payoff when they bear the same cost. Moreover, as preferences matter, the comparative statics effects depend upon whether commodities are complements or substitutes. Finally, the BSOE converges to the CE when the market is enlarged.

To conclude, the class of Stackelberg bilateral oligopoly games provides a richer set of strategic interactions. Stackelberg competition in a multi-commodity market deserves careful study to investigate merging and free entry, and it paves the way for future theoretical and applied research devoted to competition policy.

REFERENCES

Albaek S. (1990), Stackelberg leadership as a natural solution under cost uncertainty, *Journal of Industrial Economics*, 38, 335–347.

Alj A., Breton M., Haurie A. (1988), Sequential Stackelberg equilibria in two-person games, *Journal of Optimization Theory and Applications*, 59, 71–97.

Amir R. (1995), Endogenous timing in a two-player game: a counterexample, *Games and Economic Behavior*, 9, 234–237.

Amir R., Bloch F. (2009), Comparative statics in a simple class of strategic market games, *Games and Economic Behavior*, 65, 7–24.

Amir R., Grilo I. (1999), Stackelberg versus Cournot equilibrium, *Games and Economic Behavior*, 26, 1–21.

Amir R., Stepanova A. (2006), Second-mover advantage and price leadership in Bertrand duopoly, *Games and Economic Behavior*, 55, 1–20.

Amir R., Sahi S., Shubik M., Yao S. (1990), A strategic market game with complete markets, *Journal of Economic Theory*, 51, 126–143.

Anderson S.P., Engers M. (1992), Stackelberg versus Cournot oligopoly equilibrium, *International Journal of Industrial Organization*, 10, 127–135.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Anderson S.P., Engers M. (1994), Strategic investment and timing of entry, *International Economic Review*, 35, 833–853.

Anderson S.P., Erkal N., Piccinin D. (2016), Aggregative oligopoly games with entry, Working Paper.

Basu K., Singh N. (1990), Entry-deterrence in Stackelberg perfect equilibria, *International Economic Review*, 31, 61–71.

Bloch F., Ferrer H. (2001), Strategic complements and substitutes in bilateral oligopolies, *Economics Letters*, 70, 83–87.

Bloch F., Ghosal S. (1997), Stable trading structures in bilateral oligopolies, *Journal of Economic Theory*, 74, 368–384.

Boyer M., Moreaux M. (1986), Perfect competition as the limit of a hierarchical market game, *Economics Letters*, 22, 115–118.

Boyer M., Moreaux M. (1987a), Being a leader or a follower: reflections on the distribution of roles in duopoly, *International Journal of Industrial Organization*, 5, 175–192.

Boyer M., Moreaux M. (1987b), On Stackelberg equilibria with differentiated products: the critical role of the strategy space, *The Journal of Industrial Economics*, 36, 217–230.

Brander J.A., Spencer B.J. (1983), International R&D rivalry and industrial strategy, *The Review of Economic Studies*, 50, 707–722.

Carlton D.W., Perloff J.M. (2004), *Modern Industrial Organization*, Lebanon, IN: Pearson/Addison Wesley.

Colombo L., Labrecciosa P. (2008), When Stackelberg and Cournot equilibria coincide, *The B.E. Journal of Theoretical Economics*, 8, 1–7.

Cordella T., Gabszewicz J.J. (1998), "Nice" trivial equilibria in strategic market games, *Games and Economic Behavior*, 22, 162–169.

Cournot A.A. (1838), *Recherches sur les principes mathématiques de la théorie des richesses*, Paris: Chez L. Hachette.

Daughety A. (1990), Beneficial concentration, *American Economic Review*, 80, 1231–1237.

Daughety A. (2009), Introduction, purpose, and overview, in: Daughety, A. (ed.), *Cournot Oligopoly. Characterization and Applications*, New York: Cambridge University Press. pp. 3–44.

DeMiguel V., Xu H. (2009), A stochastic multiple-leader Stackelberg model: analysis, computation, and application, *Operations Research*, 57, 1220–1235.

De Wolf D., Smeers Y. (1997), A stochastic version of a Stackelberg-Nash-Cournot equilibrium model, *Management Science*, 43, 190–197.

Dickson A. (2013), The effects of entry in bilateral oligopoly, *Games*, 4, 283–303.

Dickson A., Hartley R. (2011), Bilateral oligopoly and quantity competition, *Economic Theory*, 52, 979–1004.

Dixit A.K. (1979), A model of duopoly suggesting a theory of entry-barriers, *Bell Journal of Economics*, 10, 20–32.

Dixit A.K. (1980), The role of investment in entry-deterrence, *The Economic Journal*, 90, 95–106.

Dixit A.K. (1986), Comparative statics for oligopoly, *International Economic Review*, 27, 107–122.

Dixit A.K., Stiglitz J. (1977), Monopolistic competition and the optimum product diversity, *American Economic Review*, 67, 297–308.

Dowrick S. (1986), Von Stackelberg and Cournot duopoly: choosing roles, *Rand Journal of Economics*, 17, 251–260.

Dubey P., Shubik M. (1978), The noncooperative equilibria of a closed trading economy with market supply and bidding strategies, *Journal of Economic Theory*, 17, 1–20.

Eaton C.B., Ware R. (1987), A theory of market structure with sequential entry, *Rand Journal of Economics*, 18, 1–16.

Economides N. (1993), Quantity leadership and social inefficiency, *International Journal of Industrial Organization*, 11, 219–237.

Ehrenmann A. (2004), Manifolds of multi-leader Cournot equilibria, *Operations Research Letters*, 32, 121–125.

Etro F. (2006), Aggressive leaders, *The Rand Journal of Economics*, 37, 146–154.

Etro F. (2008), Stackelberg competition with endogenous entry, *Economic Journal*, 118, 1670–1697.

Farell J., Shapiro C. (1990), Horizontal merger: an equilibrium analysis, *American Economic Review*, 80, 107–126.

Fonseca M.A., Müller W., Normann H.T. (2006), *International Journal of Game Theory*, 34, 443–456.

Freiling G., Jank G., Lee S.R. (2001), Existence and uniqueness of open loop Stackelberg equilibria in linear quadratic differential games, *Journal of Optimization Theory and Applications*, 110, 515–544.

Fudenberg D., Tirole J. (1990), *Game Theory*, Cambridge, MA: MIT Press.

Fukushima M., Pang J.S. (2005), Quasi-variational inequalities, generalized Nash equilibria, and multi-leader–follower games, *Computer Management Sciences*, 2, 21–56.

Gabszewicz J.J., Michel M. (1997), Oligopoly equilibria in exchange economies, in B.C. Eaton and R.G. Harris (eds), *Trade, Technology and Economics. Essays in Honour of R.G. Lipsey*, Cheltenham, UK and Lyme, NH, USA: Edward Elgar Publishing. pp. 217–240.

Gal-Or E. (1985), First mover and second mover advantages, *International Economic Review*, 26, 649–653.

Giraud G. (2003), Strategic market games: an introduction, *Journal of Mathematical Economics*, 39, 355–375.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Hahn F.H. (1962), The stability of the Cournot oligopoly solution, *The Review of Economic Studies*, 29, 329–333.
Hamilton J., Slutsky S. (1990), Endogenous timing in duopoly games: Stackelberg or Cournot equilibria, *Games and Economic Behavior*, 2, 29–46.
Heywood J., McGinty M. (2007), Convex costs and the merger paradox revisited, *Economic Inquiry*, 45, 342–349.
Heywood J., McGinty M. (2008), Leading and merging: convex costs, Stackelberg, and the merger paradox, *Southern Economic Journal*, 74, 879–893.
Huck S., Konrad K.A., Müller W. (2001a), Stackelberg beats Cournot: on collusion and efficiency in experimental markets, *The Economic Journal*, 111, 749–765.
Huck S., Konrad K.A., Müller W. (2001b), Big fish eat small fish: on merger in Stackelberg markets, *Economics Letters*, 73, 213–217.
Huck S., Müller W., Normann H.T. (2002), To commit or not to commit: endogenous timing in experimental duopoly markets, *Games and Economic Behavior*, 38, 240–264.
Jeffrey M., Karp L.S., Perloff J. (2007), *Estimating Market Power and Strategies*, Cambridge, UK: Cambridge University Press.
Jia J.W., He J., Xiang S. (2015), Existence and stability of weakly Pareto-Nash equilibrium for generalized multiobjective multi-leader–follower games, *Journal of Global Optimization*, 61, pp. 397–405.
Johansen L. (1982), On the status of the Nash type of noncooperative equilibrium in economic theory, *The Scandinavian Journal of Economics*, 84, 421–441.
Julien L.A. (2011), A note on Stackelberg competition, *Journal of Economics*, 103, 171–187.
Julien L.A. (2012), Stackelberg-Cournot and Cournot equilibria in a mixed markets exchange economy, *Theoretical Economics Letters,* 2, 300–306.
Julien L.A. (2013), On Stackelberg competition in strategic multilateral exchange, *Research in Economics*, 67, 59–75.
Julien L.A. (2017), On noncooperative oligopoly equilibrium in the multiple leader–follower game, *European Journal of Operational Research*, 256, 650–662.
Julien L.A., Tricou F. (2010), Oligopoly equilibria "*à la* Stackelberg" in pure exchange economies, *Louvain Economic Review*, 76, 175–194.
Julien L.A., Tricou F. (2012), Market price mechanisms and Stackelberg general equilibria: an example, *Bulletin of Economic Research*, 64, 232–259.
Julien L.A., Musy O., Saïdi A. (2011), Do followers really matter in Stackelberg competition?, *Lecturas de Economia*, 75, 11–27.
Julien L.A., Musy O., Saïdi A. (2012), Hierarchical competition in oligopoly, *Journal of Economics*, 107, 217–237.
Kaplan T.R., Wettstein D. (2000), The possibility of mixed-strategy equilibria with constant-returns-to-scale technology under Bertrand competition, *Spanish Economic Review*, 2, 65–71.
Kolstad C.D., Mathiesen L. (1987), Necessary and sufficient conditions for uniqueness of a Cournot equilibrium, *The Review of Economic Studies*, 54, 681–690.
Kydland F.E., Prescott E.C. (1977), Rules rather than discretion: the inconsistency of optimal plans, *Journal of Political Economy*, 85, 473–493.
Lafay T. (2010), A linear generalization of Stackelberg's model, *Theory and Decision*, 69, 317–326.
Leitmann G. (1978), On generalized Stackelberg strategies, *Journal of Optimization Theory and Applications*, 26, 637–643.
Mailath G. (1993), Endogenous sequencing of firms decisions, *Journal of Economic Theory*, 59, 169–182.
Matsumura T. (1999), Quantity setting oligopoly with endogenous sequencing, *International Journal of Industrial Organization*, 17, 289–296.
Mukherjee A., Zhao L. (2009), Profit raising entry, *Journal of Industrial Economics*, 57, 1–7.
Murphy F.H., Sherali H.D., Soyster A.L. (1983), Stackelberg-Nash-Cournot equilibria: characterizations and computations, *Operations Research*, 31, 253–276.
Normann H.T. (2002), Endogenous timing with incomplete information and with observable delay, *Games and Economic Behavior,* 39, 282–291.
Nti K.O., Shubik M. (1981), Noncooperative oligopoly with entry, *Journal of Economic Theory*, 24, 187–204.
Okuguchi, K. (1976), Expectations and stability in oligopoly models. *Lecture Notes in Economics and Mathematical Systems*, Berlin: Springer.
Pal D. (1991), Cournot duopoly with two production periods and cost differentials, *Journal of Economic Theory*, 55, 441–448.
Pal D., Sarkar J. (2001), A Stackelberg oligopoly with nonidentical firms, *Bulletin of Economic Research*, 53, 127–135.
Perry M.K., Porter R. (1985), Oligopoly and the incentive for horizontal merger, *American Economic Review*, 75, 219–227.
Rassenti S., Reynolds S.S., Smith V.L., Szidarovszky F. (2000), Adaptation and convergence of behavior in repeated experimental Cournot games, *Journal of Economic Behavior and Organization*, 41, 117–146.
Reinganum J.F. (1985), A two-stage model of research and development with endogenous second-mover advantages, *International Economic Review*, 3, 275–292.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Reynolds R.J., Salant S.W., Switzer S. (1983), Losses from horizontal merger: The effects of an exogenous change in industry structure on Cournot-Nash equilibrium, *The Quarterly Journal of Economics*, 98, 185–199.

Robson A. (1990a), Duopoly with endogenous strategic timing: Stackelberg regained, *International Economic Review*, 31, 263–274.

Robson A.J. (1990b), Stackelberg and Marshall, *The American Economic Review*, 80, 69–82.

Sadanand A., Sadanand V. (1996), Firm scale and the endogenous timing of entry: a choice between commitment and flexibility, *Journal of Economic Theory*, 70, 516–530.

Sahi S., Yao S. (1989), The noncooperative equilibria of a trading economy with complete markets and consistent prices, *Journal of Mathematical Economics*, 18, 325–346.

Saloner G. (1987), Cournot duopoly with two production periods, *Journal of Economic Theory*, 42, 183–187.

Salop S. (1979), Monopolistic competition with outside goods, *Bell Journal of Economics*, 10, 141–156.

Santos Pinto L. (2008), Making sense of the experimental evidence on endogenous timing in duopoly markets, *Journal of Economic Behavior and Organization*, 68, 657–666.

Schelling T. (1960), *The Strategy of Conflict*, Cambridge, MA: Harvard University Press.

Seade J.K. (1980), On the effects of entry, *Econometrica*, 48, 479–489.

Selten R. (1975), Re-examination of the perfectness concept for equilibrium points in extensive games, *International Journal of Game Theory*, 4, 25–55.

Shapley L.S. (1976), Noncooperative general exchange, in S. Lin (ed.), *Theory and Measurement of Economic Externalities*, New York: Academic Press.

Shapley L., Shubik M. (1977), Trade using one commodity as a means of payment, *Journal of Political Economy*, 85, 937–967.

Sherali H.D. (1984), A multiple leader Stackelberg model and analysis, *Operations Research*, 32, 390–405.

Shubik M. (1973), Commodity money, oligopoly, credit and bankruptcy in a general equilibrium model, *Western Economic Journal*, 11, 24–38.

Singh N., Vives X. (1984), Price and quantity competition in a differentiated oligopoly, *Rand Journal of Economics*, 15, 546–554.

Spence M.A. (1977), Entry, capacity, investment and oligopolistic pricing, *Bell Journal of Economics*, 8, 534–544.

Spencer B.J., Brander J.A. (1992), Strategic commitment with R&D: the symmetric case, *Bell Journal of Economics*, 14, 225–235.

Stackelberg H., von (1934), *Marktform und Gleichgewicht*, Berlin: Springer.

Tasnadi A. (1999), A two-stage Bertrand-Edgeworth game, *Economics Letters*, 65, 353–358.

Tobin R.L. (1992), Uniqueness results and algorithm for Stackelberg-Cournot-Nash equilibria, *Annals of Operations Research*, 34, 21–36.

Van Damme E., Hurkens S. (1999), Endogenous Stackelberg leadership, *Games and Economic Behavior*, 28, 105–129.

Van Damme E., Hurkens S. (2004), Endogenous price leadership, *Games and Economic Behavior*, 47, 404–420.

Vives X. (1988), Sequential entry, industry structure and welfare, *European Economic Review*, 32, 1671–1687.

Vives X. (1999), *Oligopoly Pricing. Old Ideas and New Tools*, Cambridge, MA: MIT Press.

Wang H.L., Yu J. (2008), An existence theorem for equilibrium points for multi-leader–follower games, *Nonlinear Analysis*, 69, 1775–1777.

Watt R. (2002), A generalized oligopoly model, *Metroeconomica*, 53, 46–55.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

APPENDIX A: A MULTIPLE LEADER–FOLLOWER GAME WITH TWO DIFFERENTIATED PRODUCTS

In Appendix A, we consider the basic game with differentiated products. Thus, we compare this model with the two homogeneous commodities models developed in Section 4. We consider a quantity-setting multiple leader–follower game with two differentiated products. Then, we compare the SOE with two differentiated products with the corresponding CE. We also compare the payoffs with the corresponding Cournot equilibrium payoffs. Moreover, we compare a duopoly differentiated version of this model with the BSOE game of Section 4. Finally, we consider the Bertrand leadership price-setting game, which we compare with the quantity-setting game.

Consider a market with a continuum $[0, 1]$ of identical consumers. The preferences of consumers are represented by the following quasi-linear utility function:

$$U(X_1, X_2, M) = X_1 + X_2 - \frac{1}{2}\left((X_1)^2 - 2\alpha X_1 X_2 + (X_2)^2\right) + M, \tag{A1}$$

where X_i, $i = 1, 2$, is the quantity of commodity i and M is the quantity of a numeraire good, whose price is equal to 1. The parameter α, with $\alpha \in (-1, 1)$, indicates whether the goods are substitutes, independent or complements, whenever $\alpha \lesseqgtr 0$ respectively.[1] The budget set of any consumer is $\mathcal{B}(p_1, p_2, I) = \left\{(X_1, X_2) \in \mathbb{R}^2_+ : p_1 X_1 + p_2 X_2 + M \leqslant I)\right\}$, where $0 < I < \infty$ is her income.

Let $L(X_1, X_2, M, \lambda) := U(X_1, X_2, M) + \lambda(I - p_1 X_1 - p_2 X_2 - M)$ be the Lagrangian, where $\lambda \geqslant 0$ is the Lagrange multiplier. By using the fact that the third optimality condition yields $\lambda^* = 1$, and since the measure of consumers over $[0, 1]$ is equal to 1, the system of inverse market demand functions may be obtained as:

$$\begin{aligned} p_1 &= 1 - X_1 + \alpha X_2 \\ p_2 &= 1 - X_2 + \alpha X_1, \end{aligned} \tag{A2}$$

with $p_1 \geqslant 0$ and $p_2 \geqslant 0$, i.e., $1 + \alpha X_2 \geqslant X_1$ and $1 + \alpha X_1 \geqslant X_2$: the marginal utility of each consumer for each commodity must be non-negative for exchange to hold.

Commodity 1 is produced by n_1 leaders, while commodity 2 is produced by n_2 followers. To simplify and without loss of generality, costs are assumed to be zero. Define $\pi_1^i(x_1^i, X_1^{-i}, X_2) := (1 - x_1^i - X_1^{-i} + \alpha X_2)x_1^i$ as the payoff function of leader i, $i = 1, \ldots, n_1$, where x_1^i is her supply, with $X_1^{-i} \equiv \sum_{-i \neq i} x_1^{-i}$, and $X_2 \equiv \sum_j x_2^j$. Similarly, for any follower j, we define $\pi_2^j(X_1, x_2^j, X_2^{-j}) := (1 - x_2^j - X_2^j + \alpha X_1)x_2^j$, with $X_1 \equiv \sum_i x_1^i$, and $X_2^{-j} \equiv \sum_{-j \neq j} x_2^{-j}$.

[1] More generally, $U(X_1, X_2) = \alpha_1 X_1 + \alpha_2 X_2 - \frac{1}{2}\left(\beta_1 (X_1)^2 - 2\gamma X_1 X_2 + \beta_2 (X_2)^2\right)$, where $\alpha_i, \beta_i > 0$, $i = 1, 2$, $\gamma \in [-1, 1]$, with $\beta_1\beta_2 - \gamma^2 > 0$ in which case U is strictly concave, and with $\alpha_i\beta_j - \alpha_j\gamma > 0$, $i \neq j$, $i = 1, 2$. The commodities are substitutes, independent or complements whenever $\gamma \lesseqgtr 0$. In addition, when $\alpha_1 = \alpha_2$ and $\beta_1 = \beta_2 = \gamma$, the goods are perfect substitutes. Finally, when $\alpha_1 = \alpha_2$, the ratio $\rho \equiv \frac{\gamma^2}{\beta_1\beta_2}$ measures the product differentiation, with $\rho \in [0, 1]$, where $\rho = 0$ and $\rho = 1$ corresponds to independency and perfect substitutability respectively (Dixit, 1979, Singh and Vives, 1984, and Vives, 1999).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

The SOE with Two Differentiated Products

There are the n_1 leaders and n_2 followers. The strategy sets are $\mathcal{S}^i = [0,\infty), i = 1,\dots,n_1$, and $\mathcal{S}^j = [0,\infty), j = 1,\dots,n_2$. The CE is given by $((x_1^i)^*, (x_2^j)^*) = \frac{1}{1-\alpha}(\chi_1^i, \chi_2^j)$, $0 \leqslant \chi_1^i, \chi_2^j \leqslant 1$, $(p_1^*, p_2^*) = (0,0)$, and $((\pi_1^i)^*, (\pi_2^j)^*) = (0,0)$, $i = 1,\dots,n_1, j = 1,\dots,n_2$. The COE with differentiated products is given by: $\hat{x}_1^i = \frac{1+(1+\alpha)n_2}{(n_1+1)(n_2+1)-\alpha^2 n_1 n_2}$, $\hat{p}_1 = \frac{1+(1+\alpha)n_2}{(n_1+1)(n_2+1)-\alpha^2 n_1 n_2}$, and $\hat{\pi}^i = \left(\frac{1+(1+\alpha)n_2}{(n_1+1)(n_2+1)-\alpha^2 n_1 n_2}\right)^2$, $i = 1,\dots,n_1$; and $\hat{x}_2^j = \frac{1+(1+\alpha)n_1}{(n_1+1)(n_2+1)-\alpha^2 n_1 n_2}$, $\hat{p}_2 = \frac{1+(1+\alpha)n_1}{(n_1+1)(n_2+1)-\alpha^2 n_1 n_2}$, and $\hat{\pi}^j = \left(\frac{1+(1+\alpha)n_1}{(n_1+1)(n_2+1)-\alpha^2 n_1 n_2}\right)^2$, $j = 1,\dots,n_2$. Then, we deduce $\frac{\hat{p}_1}{\hat{p}_2} = \frac{1+(1+\alpha)n_2}{1+(1+\alpha)n_1}$, with $\lim_{(n_1,n_2)\to(\infty,\infty)} \frac{\hat{p}_1}{\hat{p}_2} = 1$. Assuming symmetry among followers, i.e., $x_2^j = x_2^{-j}$, $-j \neq j$, the best response $\varphi^j(X^1)$ of follower j may be written:

$$\varphi^j(X_1) = \frac{1}{n_2+1} + \frac{\alpha}{n_2+1} X_1, j = 1,\dots,n_2. \tag{A3}$$

We remark that for all j we get $\frac{\partial \varphi^j(X^1)}{\partial x_1^i} = \frac{\alpha}{n_2+1} \lesseqqgtr 0$, whenever $\alpha \lesseqqgtr 0$, $\forall i$. In the first stage of the game, leader i's supply is the solution to:

$$\max\ \pi_1^i(x_1^i, X_1^{-i}, X_2(x_1^i, X_1^{-i})), i = 1,\dots,n_1, \tag{A4}$$

where $\pi_1^i(x_1^i,.) := \left(1 - x_1^i - X_1^{-i} + \alpha\frac{n_2}{n_2+1}(1 + \alpha(x_1^i + X_1^{-i}))\right) x_1^i$. We deduce:

$$\tilde{x}_1^i = \frac{(1+\alpha)n_2 + 1}{(n_1+1)(n_2+1) - \alpha^2(n_1+1)n_2}, i = 1,\dots,n_1. \tag{A5}$$

Using (A3) we deduce the equilibrium strategy of follower j:

$$\tilde{x}_2^j = \frac{1}{n_2+1} + \frac{\alpha n_1}{n_2+1}\frac{(1+\alpha)n_2 + 1}{(n_1+1)(n_2+1) - \alpha^2(n_1+1)n_2}, j = 1,\dots,n_2. \tag{A6}$$

The equilibrium market prices follow from (A2):

$$\tilde{p}_1 = \frac{(1+\alpha)n_2 + 1}{n_2+1} - \frac{(1-\alpha)(1+\alpha)n_2 + 1}{n_2+1}\widetilde{X}_1 \tag{A7}$$

$$\tilde{p}_2 = \frac{1}{n_2+1} + \frac{\alpha}{n_2+1}\widetilde{X}_1. \tag{A8}$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Then, $\frac{\tilde{p}_1}{\tilde{p}_2} = \frac{(1+\alpha)n_2+1-[(1-\alpha)(1+\alpha)n_2+1]\widetilde{X}_1}{1+\alpha\widetilde{X}_1}$. The payoffs are given by:

$$\tilde{\pi}_1^i = \left(\frac{(1+\alpha)n_2+1}{n_2+1} - \frac{[(1-\alpha)(1+\alpha)n_2+1]\widetilde{X}_1}{n_2+1}\right)\frac{\widetilde{X}_1}{n_1}, i = 1,\ldots,n_1 \tag{A9}$$

$$\tilde{\pi}_2^j = \left(\frac{\alpha\widetilde{X}_1+1}{n_2+1}\right)^2, j = 1,\ldots,n_2. \tag{A10}$$

Let us notice that when $n_1 = n_2 = 1$, the market outcome coincides, up to a scalar multiple due to normalization, with the market outcome of the Stackelberg duopoly equilibrium with two differentiated products of Boyer and Moreaux (1987b).

We now compare the Stackelberg oligopoly equilibrium with two differentiated products with the CE. Some calculations lead to $\lim_{(n_1,n_2)\to(\infty,\infty)} = (\tilde{p}_1, \tilde{p}_2) = (0,0) = (p_1^*, p_2^*)$ as $\lim_{(n_1,n_2)\to(\infty,\infty)} \widetilde{X}_i = \frac{1}{1-\alpha} = X_i^*$, $i = 1,2$. Then, free entry leads to the CE market prices, but the relative price is indeterminate, which is not the case in the bilateral oligopoly model of Section 4. To compare, first, the equilibrium strategies, prices and payoffs of the leaders and of the followers in the SOE, and, second, the SOE with the COE with differentiated products, let $n_1 = n_2 = n$ in (A5)–(A10). We get $\widetilde{X}_1 = \frac{n[(1+\alpha)n+1]}{(n+1)^2-\alpha^2 n(n+1)}$, $\tilde{p}_1 = \frac{[(1+\alpha)n+1][(1-\alpha^2)n+1]}{(n+1)[(n+1)^2-\alpha^2 n(n+1)]}$, and $\tilde{\pi}_1^i = \frac{(1-\alpha^2)n+1}{n+1}\left(\frac{(1+\alpha)n+1}{(n+1)^2-\alpha^2 n(n+1)}\right)^2$, $i = 1,\ldots,n_1$; and $\widetilde{X}_2 = \frac{n}{n+1}\frac{(n+1)^2+\alpha n(2-\alpha)}{(n+1)^2-\alpha^2 n(n+1)}$, $\tilde{p}_2 = \frac{(n+1)^2+\alpha n(n+1-\alpha)}{(n+1)[(n+1)^2-\alpha^2 n(n+1)]}$, and $\tilde{\pi}_2^j = \frac{[(n+1)^2+\alpha n(2-\alpha)][(n+1)^2+\alpha n[n+1-\alpha]}{((n+1)[(n+1)^2-\alpha^2 n(n+1)])^2}$, $j = 1,\ldots,n_2$. Some tedious calculations lead to $\widetilde{X}_1 > \widetilde{X}_2$, $\tilde{p}_1 < \tilde{p}_2$, and $\tilde{\pi}_1^i > \tilde{\pi}_2^j$, as $-1 < \alpha < 1$, $i = 1,\ldots,n_1$, $j = 1,\ldots,n_2$.

In addition, in the Cournot oligopoly equilibrium we get $\widehat{X}_1 = \frac{n[(1+\alpha)n+1]}{(n+1)^2-\alpha^2 n^2}$, $\hat{p}_1 = \frac{(1+\alpha)n+1}{(n+1)^2-\alpha^2 n^2}$, and $\hat{\pi}_1^i = \left(\frac{(1+\alpha)n+1}{(n+1)^2-\alpha^2 n^2}\right)^2$, $i = 1,\ldots,n_1$; and $\widehat{X}_2 = \frac{n[(1+\alpha)n+1]}{(n+1)^2-\alpha^2 n^2}$, $\hat{p}_2 = \frac{(1+\alpha)n+1}{(n+1)^2-\alpha^2 n^2}$, and $\hat{\pi}_2^j = \left(\frac{(1+\alpha)n+1}{(n+1)^2-\alpha^2 n^2}\right)^2$, $j = 1,\ldots,n_2$. Then, we deduce $\widetilde{X}_1 > \widehat{X}_1$, $\tilde{p}_1 < \hat{p}_1$, and $\tilde{\pi}_1^i > \hat{\pi}_2^i$, $i = 1,\ldots,n_1$; and $\widetilde{X}_2 > \widehat{X}_2$, $\tilde{p}_2 < \hat{p}_2$, and $\tilde{\pi}_1^j > \hat{\pi}_2^j$, as $-1 < \alpha < 1$, $i = 1,\ldots,n_1, j = 1,\ldots,n_2$. We also have $\frac{\tilde{p}_1}{\tilde{p}_2} < \frac{\hat{p}_1}{\hat{p}_2} = 1$. *Therefore, unlike the BSOE of Section 4, the Stackelberg and Cournot equilibria with differentiated products do not coincide even if marginal costs are identical.* This result stems from the linearity of the market demands, each of which is derived from competitive behavior. Indeed, as outlined in the Introduction, this result stems from the asymmetric behavior of the first type: the demand side embodies competitive buyers, whilst the supply side includes strategic firms.

A Variation Around the Stackelberg Game with Two Differentiated Products

To refine the comparison between the BSOE provided in Example 13, and the Stackelberg oligopoly with two differentiated products, let $n_1 = n_2 = 2$. From (A5)–(A10), we get $\tilde{x}_1^i = \frac{2\alpha+3}{3(3-2\alpha^2)}$, $\tilde{p}_1 = \left(\frac{2\alpha+3}{3}\right)^2$, and $\tilde{\pi}_1^i = \frac{1}{3-2\alpha^2}\left(\frac{2\alpha+3}{3}\right)^3$, $i = 1,2$; and $\tilde{x}_2^j = \frac{9+6\alpha-2\alpha^2}{9(3-2\alpha^2)}$, $\tilde{p}_2 = \frac{9+6\alpha-2\alpha^2}{9(3-2\alpha^2)}$, and $\tilde{\pi}^j = \left(\frac{9+6\alpha-2\alpha^2}{9(3-2\alpha^2)}\right)^2$, $j = 1,2$. Consider now that, for each commodity, there is one leader and one follower who compete

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

on the same commodity. Therefore, the two followers solve respectively $\max \pi_1^2 := \left(1 - x_1^L - x_1^F + \alpha X_2\right) x_1^F$ and $\max \pi_2^2 := \left(1 + \alpha X_1 - x_2^L - x_2^F\right) x_1^2$, where $X_1 \equiv x_1^1 + x_1^2$ and $X_2 \equiv x_2^1 + x_2^2$. The best responses are $\varphi\left(x_1^1, x_2^1\right) = \frac{1}{2-\alpha} - \frac{2-\alpha^2}{(2-\alpha)(2+\alpha)} x_1^1 + \frac{\alpha}{(2-\alpha)(2+\alpha)} x_2^1$ and $\sigma(x_1^1, x_2^1) = \frac{1}{2-\alpha} + \frac{\alpha}{(2-\alpha)(2+\alpha)} x_1^1 - \frac{2-\alpha^2}{(2-\alpha)(2+\alpha)} x_2^1$. The leaders' equilibrium strategy profile is $(\bar{x}_1^1, \bar{x}_2^1) = \left(\frac{\alpha+2}{2(2-\alpha^2)-\alpha}, \frac{\alpha+2}{2(2-\alpha^2)-\alpha}\right)$, from which we deduce the followers' profile $\left(\bar{x}_1^2, \bar{x}_2^2\right) = \left(\frac{2-\alpha^2}{(2-\alpha)\left[2(2-\alpha^2)-\alpha\right]}, \frac{2-\alpha^2}{(2-\alpha)\left[2(2-\alpha^2)-\alpha\right]}\right)$. The equilibrium prices are given by $\bar{p}_1 = \bar{p}_2 = \frac{2-\alpha^2}{(2-\alpha)[2(2-\alpha^2)-\alpha]}$, and the payoffs by $\left(\bar{\pi}_1^1, \bar{\pi}_2^1\right) = \frac{(\alpha+2)(2-\alpha^2)}{(2-\alpha)[2(2-\alpha^2)-\alpha]^2}(1,1)$ and $(\bar{\pi}_1^2, \bar{\pi}_2^2) = \left(\frac{2-\alpha^2}{(2-\alpha)[2(2-\alpha^2)-\alpha]}\right)^2 (1,1)$, with $\left(\bar{\pi}_1^1, \bar{\pi}_2^1\right) >> \left(\bar{\pi}_1^2, \bar{\pi}_2^2\right)$.

Then, $\tilde{\pi}_1^i > \bar{\pi}_1^1 = \bar{\pi}_2^1$, $i = 1, 2$, and $\tilde{\pi}_2^j > \bar{\pi}_1^2 = \bar{\pi}_2^2$, $j = 1, 2$ (prices are higher in the double Stackelberg case). Leaders (resp. followers) always prefer interacting only with leaders (resp. followers) in their own industry, whichever commodities are substitutes, independent or complements. The leaders' market power is higher when they face symmetric Cournot competitors who interact in the other industry. Thus, this comparison displays a Pareto domination between the two duopoly equilibria. This result does not necessarily hold in the BSOE since the allocations depend on the endogenous relative price. The welfare implications of endogenous entry are more difficult to handle. Nevertheless, price reduction by leaders may more than compensate for the reduction in the number of varieties, so consumers' surplus increase (see Etro, 2008).

The Stackelberg Price-setting Game with Two Differentiated Products
Let $n_1 = n_2 = 1$. The demand obtained from (A2) are given by:

$$X_1 = \frac{1}{1-\alpha} - \frac{1}{1-\alpha^2} p_1 - \frac{\alpha}{1-\alpha^2} p_2 \tag{A11}$$
$$X_2 = \frac{1}{1-\alpha} - \frac{1}{1-\alpha^2} p_2 - \frac{\alpha}{1-\alpha^2} p_1.$$

The best response is $\varphi(p_1) = \frac{1+\alpha}{2} - \frac{\alpha}{2} p_1$, so the leader solves $\max \pi_1(p_1, \varphi(p_1)) = (\frac{2-\alpha}{2(1-\alpha)} - \frac{2-\alpha^2}{2(1-\alpha^2)} p_1) p_1$. We deduce:

$$(\tilde{p}_1, \tilde{p}_2) = \left(\frac{(1+\alpha)(2-\alpha)}{2(2-\alpha^2)}, \frac{(1+\alpha)(2-2\alpha+\alpha^2)}{4(2-\alpha^2)}\right). \tag{A12}$$

From (A12) we deduce the quantity traded:

$$(\widetilde{X}_1, \widetilde{X}_2) = \frac{1}{4(1-\alpha)(2-\alpha^2)} (4 - 2\alpha^2 - \alpha^3, 6 - 2\alpha - 3\alpha^2). \tag{A13}$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Then, the payoffs are given by:

$$\tilde{\pi}_1 = \left(\frac{1}{2(2-\alpha^2)}\right)^2 \frac{(1+\alpha)(2-\alpha)(4-2\alpha^2-\alpha^3)}{2(1-\alpha)} \tag{A14}$$

$$\tilde{\pi}_2 = \left(\frac{1}{4(2-\alpha^2)}\right)^2 \frac{(1+\alpha)(2-2\alpha+\alpha^2)(6-2\alpha-3\alpha^2)}{1-\alpha}. \tag{A15}$$

We get $\tilde{p}_1 > \tilde{p}_2$, $\widetilde{X}_1 < \widetilde{X}_2$, and $\tilde{\pi}_1 < \tilde{\pi}_2$. Some calculations show that both firms achieve higher payoffs than in the simultaneous-move Bertrand game: the leader expects that its follower will lower its price to increase its higher market share, so the leader chooses a high price, which leads both firms to set prices above the Bertrand single period game. The main difference with the corresponding quantity duopoly game is that the leader's profit is higher whenever the products are substitutes (with identical costs). To see this consider (A5)–(A10) with $n_1 = n_2 = 1$. Then, we deduce $(\tilde{x}_1, \tilde{x}_2) = \left(\frac{2+\alpha}{2(2-\alpha^2)}, \frac{4+2\alpha-\alpha^2}{4(2-\alpha^2)}\right)$, $(\tilde{p}_1, \tilde{p}_2) = \left(\frac{2+\alpha}{4}, \frac{4+2\alpha-\alpha^2}{4(2-\alpha^2)}\right)$, and $(\tilde{\pi}_1, \tilde{\pi}_2) = \left(\frac{1}{2(2-\alpha^2)}\left(\frac{2+\alpha}{2}\right)^2, \left(\frac{4+2\alpha-\alpha^2}{4(2-\alpha^2)}\right)^2\right)$, so $\tilde{x}_1 > \tilde{x}_2$, $\tilde{p}_1 \lesseqgtr \tilde{p}_2$, and $\tilde{\pi}_1 \gtreqless \tilde{\pi}_2$, whenever $\alpha \lesseqgtr 0$. Therefore, the quantity-setting and price-setting differentiated products games are dual to each other when strategies are substitutes. Otherwise, the quantity-setting game has the same welfare property as the basic Stackelberg duopoly game since the parameter α plays the same role as the term η in the model of Section 4, with the main difference that η is endogenously determined.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

APPENDIX B: PROOF OF PROPOSITION 6

In Appendix B, we provide a proof for Proposition 6. The optimality conditions enables the capture of some salient features of the BSOE. The followers' optimal decision mappings are solutions to:

$$\max_{\phi^i(\mathbf{b}^{-i};\mathbf{q})} U^i\left(e^i - q^i, \frac{\mathbf{Q}}{b^i + \mathbf{B}^{-i}} b^i - F_i^{-1}(e^i)\right), i = n_L + 1, \ldots, n \tag{B1}$$

$$\max_{\psi^j(\mathbf{b};\mathbf{q}^{-j})} V^j\left(\frac{b^i + \mathbf{B}^{-i}}{\mathbf{Q}} q^j, \omega^j - q^j\right), j = m_L + 1, \ldots, m. \tag{B2}$$

Differentiating (B1) with respect to b^i and y^i, and expressing the two conditions as a single one; and differentiating (B2) with respect to q^j, lead respectively to:

$$\frac{p_X}{p_Y}(\mathbf{b};\mathbf{q}) + \frac{\partial(\frac{p_X}{p_Y})}{\partial \mathbf{B}} b^i = \frac{dF_i^{-1}(e^i)}{de^i}, i = n_L + 1, \ldots, n \tag{B3}$$

$$\frac{p_Y}{p_X}(\mathbf{b};\mathbf{q}) + \frac{\partial(\frac{p_Y}{p_X})}{\partial \mathbf{Q}} q^j = \frac{\partial V^j/\partial y^j}{\partial V^j/\partial x^j}, j = m_L + 1, \ldots, m. \tag{B4}$$

The optimal decision functions of each type of follower may be written:

$$b^i = \phi^i\left(\mathbf{b}^{-i};\mathbf{q}, \frac{dF_i^{-1}(e^i)}{de^i}\right), i = n_L + 1, \ldots, n \tag{B5}$$

$$q^j = \psi^j\left(\mathbf{b};\mathbf{q}^{-j}, \frac{\partial V^j/\partial y^j}{\partial V^j/\partial x^j}\right), j = m_L + 1, \ldots, m. \tag{B6}$$

We assume the system made up of equations (B5)–(B6) is consistent. Therefore, we assume the existence of best responses, which are given by:

$$b^i = \varphi^i\left(\mathbf{b}_L;\mathbf{q}_L, \bigtriangledown(F_i^{-1})_{i \geqslant n_L+1}\right), i = n_L + 1, \ldots, n \tag{B7}$$

$$q^j = \sigma^j\left(\mathbf{b}_L;\mathbf{q}_L, \bigtriangledown(F_i^{-1})_{i \geqslant n_L+1}\right), j = m_L + 1, \ldots, m, \tag{B8}$$

where $\bigtriangledown\left(F_i^{-1}\right)_{i \geqslant n_L+1} = \left(\frac{dF_{n_L+1}^{-1}(e^{n_L+1})}{de^{n_L+1}}, \ldots, \frac{dF_n^{-1}(e^n)}{de^n}\right)$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Then, leader i, $i = 1, \dots, n_L$, and leader j, $j = 1, \dots, m_L$, solve:

$$\max_{b^i} U^i \left(e^i - b^i, \frac{p_X}{p_Y} \left(b^i, \mathbf{b}_L^{-i}, \boldsymbol{\varphi}\left(b^i, .\right); \mathbf{q}_L, \boldsymbol{\sigma}\left(b^i, .\right) \right) b^i - F_i^{-1}(e^i) \right) \tag{B9}$$

$$\max_{q^j} V^j \left(\frac{p_Y}{p_X} \left(\mathbf{b}_L, \boldsymbol{\varphi}(., q^j, .); q^j, \mathbf{q}_L^{-j}, \boldsymbol{\sigma}(., q^j, .) \right) q^j, \omega^j - q^j \right), \tag{B10}$$

with $\boldsymbol{\varphi}(b^i, .) = (\varphi^{n_L+1}(b^i, .), \dots, \varphi^n(b^i, .))$, where for each $i = n_L + 1, \dots, n$, the function $\varphi^i(b^i, .)$ is defined by (B7); and with $\boldsymbol{\sigma}\left(., q^j\right) = (\sigma^{m_L+1}(., q^j, .), \dots, \sigma^m(., q^j, .))$, where for each $j = m_L + 1, \dots, m$, the function $\sigma^j(., q^j, .)$ is defined by (B8).

Let the market price be $\frac{p_X}{p_Y} = \frac{\sum_{j=1}^{m_L} q^j + \sum_{j=m_L+1}^{m} \sigma^j \left(\mathbf{b}_L; \mathbf{q}_L, \nabla\left(F_i^{-1}\right)_{i \geqslant n_L+1} \right)}{\sum_{i=1}^{n_L} b^i + \sum_{i=m_L+1}^{m} \varphi^i \left(\mathbf{b}_L; \mathbf{q}_L, \nabla\left(F_i^{-1}\right)_{i \geqslant n_L+1} \right)}$. Then, partially differentiating (B9) with respect to b^i and y^i respectively, and expressing the two conditions obtained as a single one; and partially differentiating (B10) with respect to q^j, lead respectively to the two following first-order (sufficient) conditions:

$$\frac{p_X}{p_Y} + \frac{\partial\left(\frac{p_X}{p_Y}\right)}{\partial \mathbf{B}} (1 + \nu_X) b^i + \frac{\partial\left(\frac{p_X}{p_Y}\right)}{\partial \mathbf{Q}} \eta_Y b^i = \frac{dF_i^{-1}(e^i)}{de^i}, i = 1, \dots, n_L \tag{B11}$$

$$\frac{p_Y}{p_X} + \frac{\partial\left(\frac{p_Y}{p_X}\right)}{\partial \mathbf{Q}} (1 + \nu_Y) q^j + \frac{\partial\left(\frac{p_Y}{p_X}\right)}{\partial \mathbf{B}} \eta_X q^j = \frac{\partial V^j / \partial y^j}{\partial V^j / \partial x^j}, j = 1, \dots, m_L. \tag{B12}$$

These expressions may be written in equilibrium as:

$$\frac{p_X}{p_Y} \left(1 - \frac{1 + \nu_X}{|\epsilon_X|} \frac{\tilde{b}^i}{\tilde{\mathbf{B}}} + \frac{\eta_Y}{|\epsilon_Y|} \frac{\tilde{b}^i}{\tilde{\mathbf{Q}}} \right) = \frac{dF_i^{-1}(e^i)}{de^i}_{|e^i = \tilde{e}^i}, i = 1, \dots, n_L \tag{B13}$$

$$\frac{p_Y}{p_X} \left(1 - \frac{1 + \nu_Y}{|\epsilon_Y|} \frac{\tilde{q}^j}{\tilde{\mathbf{Q}}} + \frac{\eta_X}{|\epsilon_X|} \frac{\tilde{q}^j}{\tilde{\mathbf{B}}} \right) = \frac{\partial V^j / \partial y^j}{\partial V^j / \partial x^j} (\tilde{\mathbf{a}}^j), j = 1, \dots, m_L, \tag{B14}$$

where $|\epsilon_X| = \frac{\partial \mathbf{B}}{\partial\left(\frac{p_X}{p_Y}\right)} \frac{\frac{p_X}{p_Y}}{\mathbf{B}} = 1$ and $|\epsilon_Y| = \frac{\partial \mathbf{Q}}{\partial\left(\frac{p_X}{p_Y}\right)} \frac{\frac{p_X}{p_Y}}{\mathbf{Q}} = 1$.

The equations (B13) and (B14) determine together the equilibrium strategy profile of leaders $(\tilde{\mathbf{b}}_L; \tilde{\mathbf{q}}_L) = (\tilde{b}^1, \dots, \tilde{b}^{n_L}; \tilde{q}^1, \dots, \tilde{q}^{m_L})$ and the equilibrium production vector $\tilde{\mathbf{e}}_L = (\tilde{e}^1, \dots, \tilde{e}^{n_L})$. Then, from (B7) and (B8) we deduce the equilibrium strategy profile of the

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

followers $(\tilde{\mathbf{b}}_F;\tilde{\mathbf{q}}_F) = (\tilde{b}^{n_L+1},\ldots,\tilde{b}^n;\tilde{q}^{m_L+1},\ldots,\tilde{q}^m)$, which in the bilateral oligopoly market game must satisfy (B3) and (B4), so:

$$\frac{p_X}{p_Y}\left(\tilde{\mathbf{b}};\tilde{\mathbf{q}}\right)\left(1-\frac{1}{|\epsilon_X|}\frac{\tilde{b}^i}{\tilde{\mathbf{B}}}\right)=\frac{dF_i^{-1}(e^i)}{de^i}_{|e^i=\tilde{e}^i}, i=1,\ldots,n_L \tag{B15}$$

$$\frac{p_Y}{p_X}\left(\tilde{\mathbf{b}};\tilde{\mathbf{q}}\right)\left(1-\frac{1}{|\epsilon_Y|}\frac{\tilde{q}^j}{\tilde{\mathbf{Q}}}\right)=\frac{\partial V^j/\partial y^j}{\partial V^j/\partial x^j}(\tilde{\mathbf{a}}^j), j=1,\ldots,m_L. \tag{B16}$$

Therefore (B13)–(B16) coincide with (10.39)–(10.42).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

11. Entry games and free entry equilibria*

Michele Polo

1 INTRODUCTION

What are the elements that may explain why certain industries are populated by a large number of firms, each covering a small fraction of total output, whereas other markets are dominated by a few large firms that supply a relevant fraction of customers? These questions have been at the core of the topics studied in industrial organization (IO) from the very beginning.[1] These research topics have been approached in the early phases of industrial economics mostly from an empirical perspective[2] within the structure–conduct–performance paradigm, while the theoretical foundations of endogenous market structures have been explored more rigorously in the game-theoretic framework of the new IO literature. The analytical framework that has been developed looks at market entry and exit as the process that endogenously determines the number and characteristics of active firms in the long run. In this setting, then, other research questions emerge. How do these market structures change in reaction to a variation in some key parameters? Are we able to identify a set of robust comparative statics properties in oligopoly markets, despite the rich variety of models in the IO literature? And finally, on the normative side, does entry into the market, a key component of the competitive process, lead to a welfare-maximizing outcome, or might the number and characteristics of firms exceed or fall short of the level of efficiency?

This chapter deals with the theories of market equilibria when the number and characteristics of active firms are endogenously determined through the process of entry. More precisely, we shall review the literature on entry games and free entry equilibria in a multi-stage game framework. A large number of potential entrants decide first whether to enter or not; once all the firms have undertaken their entry decisions, the active firms compete according to some oligopoly game. The chapter is organized as follows. In Section 2 we present the general analytical framework. In Section 3 we analyze a wide range of symmetric oligopoly models to identify the relationship between the number of firms and the market equilibria: we start with homogeneous products and competition in strategic substitutes (Section 3.1), moving then to differentiated products and competition in strategic complements (Section 3.2), next offering a general explanation of the comparative statics properties (Section 3.3) and concluding with cartels (Section 3.4). We then move to free entry equilibria and the determinants of the maximum number of firms (Section 4). Finally, we consider symmetric entry games under a normative perspective (Section 5), looking at the comparison between the free entry and the welfare-maximizing number of firms. In Section 6 we review asymmetric

* I thank Simon Anderson, Emilio Calvano, Chiara Fumagalli and Michele Grillo for very useful discussions and suggestions. Usual disclaimers apply.

1 See Bain (1956) and Scherer (1980).

2 See Schmalensee (1989) for a comprehensive survey of the empirical literature.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

free entry equilibria that exploit the aggregative nature of most oligopoly models. We then present the case of endogenous sunk cost and persistent concentration (Section 7) and the case of frictionless entry and contestable markets (Section 8). Concluding remarks follow.

2 ENTRY GAMES

There are several ways to model the entry process and market interaction among active firms. The various set-ups allow us to highlight different issues, focussing on distinct effects that interact in the overall market dynamics. We can draw a key distinction between the environments in which the entry decisions precede the market strategies, and those where some firms undertake entry decisions after observing their competitors' market strategies.

In the former case, the *market strategies* of individual firms cannot be chosen with the purpose of affecting the entry decisions of any firm, since entry already occurred, although the features of the *market equilibria* that result from the aggregate process of entry affect the early decision to enter the market. In this perspective, multi-stage games represent a suitable formal framework. There is a large group of m potential entrants $j \in I_m$ that choose whether to enter, incurring a fixed set-up cost $F > 0$, or not; then, once they have taken their decision and the set of $n \leq m$ entrants $i \in I_n$ is common knowledge, the active firms play a market game. This set-up is usually adopted to study long-run free entry equilibria, in which a set of exogenous variables referring to the primitives of technology and preferences explains the long-run market structure.

Alternatively, in a second class of strategic environments, a subset of early entrants (incumbents) commit to observable market strategies before the other firms (entrants) decide whether to enter or not. The incumbents' initial strategy, then, may affect the entry decisions of the latecomers, explaining why this set-up is widely used to study strategic entry deterrence and foreclosure. In this environment, the market structure is explained by foreclosure strategies, based on a rich set of strategic tools, rather than by market fundamentals.

The two set-ups are useful to explore different and complementary issues and they are characterized by a different time horizon. Sequential entry with incumbents and entrants is a more realistic representation of short-run market dynamics, since entry is typically an ongoing process where already established and new firms interact. The possibility of foreclosure, then, is an empirically relevant issue that characterizes the evolution of markets. At the same time, multi-stage entry games allow us to move away from these short-run phenomena and focus on the underlying features of preferences and technology as long-run drivers of market evolution. By shifting attention to this complementary perspective we can identify fundamental forces that, despite the frictions that in the short run may slow down the process and foreclose the market, push towards a more or less concentrated market. Since in this chapter the focus is on long-run market structures rather than foreclosure, we will consider several and different specifications of multi-stage entry games.

A second relevant feature recurring across models is the assumption of symmetric firms. Supply-side symmetry is a natural assumption in a long-run perspective, since we may think that any barrier to adopting best practice technologies, such as patent protection or private know-how, tends to vanish in the long run. Demand-side symmetry, consistent with homogeneous products or horizontal product differentiation and different varieties, is a

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

convenient assumption when we want to analyze the number of entrants and the distribution of market shares.[3]

The different models considered in the following sections make use of the symmetry assumption at different levels, either by applying it to the whole population of potential entrants, or to a subset of them identified as marginal entrants, while allowing for asymmetries across major market players. We shall see that the symmetry assumption is also at the core of the analysis of potential competition and contestable markets.

3 SYMMETRIC OLIGOPOLY MARKETS

We start our analysis of entry games by considering the (second-stage) market games where n firms are active, having decided to enter in the first stage. In this section we consider symmetric market games where all the n firms share the same (best-practice) technology and no one has an advantage on the demand side, e.g. a higher-quality product. In this setting, when firms adopt the same strategies $a_i = a$, $i \in I_n$, then they obtain the same level of profits. A symmetric environment greatly simplifies the analysis of free entry equilibria, since the equilibrium profits, as well as the equilibrium strategies, consumers' surplus and welfare, all depend on a vector x of parameters related to the properties of costs (technology) and demand (preferences), and on the number of firms n: $\Pi_i(a_i^*, a_{-1}^*) = \Pi^*(n; x)$. Market equilibria, once the entry process has been completed, therefore can be analyzed simply in terms of the number of firms n. The individual equilibrium profits $\Pi^*(n; x)$ are therefore the object that potential entrants consider when, at the initial stage of the game, they choose whether to enter or not, given their expectation of the number of firms that will enter.

Oligopoly theory offers a very rich set of models that describe market interaction among n competitors, ranging from homogeneous to differentiated products and distinguishing competition in strategic substitutes or complements. In all these environments, moreover, demand and cost functions can be specified differently. Finally, beyond static, possibly multi-stage games, the literature on tacit collusion adds to the toolkit for the analysis of cartels. A general theory of free entry equilibria has to encompass all these classes of models, admitting a variety of business strategies, modes of strategic interaction and features of demand and costs. In this perspective, then, the key point is whether there exist some regularities across different models in the relationship between the number of (symmetric) active firms n and the equilibrium profits they obtain $\Pi^*(n; x)$. A first, relevant result that we are going to present in the following sections, is that, despite the significant differences in oligopoly equilibria across models, we can establish under very general conditions a negative relationship between the equilibrium profits and the number of firms.

We organize the discussion by considering three different cases: homogeneous products and strategic substitutes, differentiated products and strategic complements, and repeated games.

[3] As will be clear in the following sections, this approach does not prevent us from also considering environments where, for instance, firms offer goods of different quality, which are therefore attractive to consumers in different ways. What we maintain is that, even in these cases, there is a further dimension of (horizontal) product differentiation such that for each level of quality several firms may further differentiate their products by variety. In this case, symmetry is preserved at each layer of quality.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

3.1 Homogeneous Products and Strategic Substitutes

Our first look at symmetric oligopoly equilibria refers to a market with n firms producing a homogeneous product and competing in strategic substitutes, usually associated with the Cournot model. Since the pioneering work of Cournot (1838) a large number of contributions have explored the conditions for the existence of and characteristics of the equilibria when n firms compete in quantities. McManus (1962, 1964) and Roberts and Sonnenschein (1976), independently proved the existence of a symmetric equilibrium in symmetric Cournot games with convex costs. Novshek (1985) showed that an n-oligopoly has a Nash equilibrium if each firm's marginal revenue is decreasing in the other firms' aggregate output. A step forward in proving the existence of Cournot equilibria under general conditions is in Vives (1990), who showed in the duopoly case the relationship between the assumptions of the previous literature and the submodularity of Cournot games. Supermodular games and the techniques of monotone comparative statics,[4] have proved to be extremely useful tools to explore the properties of Cournot oligopolies and to identify the general conditions under which the comparative statics of equilibria can be analyzed. We summarize here the main results following this approach as in Amir and Lambson (2000).

Consider an oligopoly with n firms offering a homogeneous product and producing with the same cost function $C(q_i)$ and incurring no capacity constraint over the relevant output range. Market inverse demand $P(Q)$ is a continuous and differentiable function of total output $Q = \sum_{i=1}^{n} q_i$. The profit function of firm i, then, is:

$$\Pi_i(q_i, \mathbf{Q}_{-i}) = P(Q)q_i - C(q_i)$$

where $\mathbf{Q}_{-i} = \{q_j\}_{j \neq i}$ is the vector of outputs of the other firms. In this traditional specification, each firm maximizes its profits by choosing a level of output for given strategies of the other firms, $\mathbf{Q}_{-i}$. It is well recognized that under standard assumptions, firm i's best reply $\widehat{q}_i(\mathbf{Q}_{-i}) = \arg\max_{q_i} \Pi_i(q_i, \mathbf{Q}_{-i})$ is downward sloping in the other firms' output, implying a submodular game and competition in strategic substitutes.

Let us define

$$\Delta(q_i, Q) := -P'(Q) + C''(q_i). \tag{11.1}$$

Then, Amir and Lambson (2000) prove that if $\Delta(q_i, Q) > 0$ on the relevant range of outputs and the inverse demand function is log-concave, there exists a unique and symmetric equilibrium, with individual output $q^*(n)$ nonincreasing in n and total output $Q^*(n)$ (market price $P(Q^*(n))$) nondecreasing (nonincreasing) in n.[5] This condition holds, for instance, in the set-up adopted in the works of McManus (1962, 1964), Roberts and Sonnenschein (1976) and Novshek (1985) quoted above and is consistent with the framework proposed in Vives (1999).

---

4 See Milgrom and Roberts (1990, 1994) and Milgrom and Shannon (1994).

5 Amir and Lambson (2000) prove (Theorem 2.2) a more general result that does not require log-concavity of the inverse demand function and that allows for multiplicity of Cournot equilibria. In this case the comparative statics properties with respect to n of total equilibrium output and the equilibrium output of $n-1$ firms are preserved by considering the values of the extremal equilibria. We focus in the text on uniqueness to ease the exposition.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

To illustrate this result with an example let us consider the linear Cournot model: market demand is $Q = S * \left[\alpha - \beta p\right]$, where S measures market size, e.g. the number of consumers. Then, the inverse demand is $P(\frac{Q}{S}) = a - b\frac{Q}{S}$ where $a = \frac{\alpha}{\beta}$, $b = \frac{1}{\beta}$ and Q is total supply. Firms produce at constant marginal cost $c \in (0, a)$ and compete in quantities. Then, each firm selects its optimal output by solving $q_i^* = \arg\max_{q_i} \left(P(\frac{Q}{S}) - c\right) q_i$. The symmetric equilibrium quantity $q^*(n)$ satisfies for all firms the first-order conditions:

$$\left(P\left(\frac{nq^*}{S}\right) - c\right) - P'\frac{q^*}{S} = 0, \tag{11.2}$$

Substituting and solving for the symmetric equilibrium we get:

$$q^*(n) = S\frac{a-c}{b(n+1)}, \quad p^*(n) = \frac{a+nc}{n+1} \geq c, \quad \Pi^*(n) = \frac{S}{b}\left(\frac{a-c}{n+1}\right)^2. \tag{11.3}$$

When the number of firms increases, therefore, the individual quantity decreases, whereas total output increases. Consequently, the market clearing price falls and tends to the marginal cost when the number of firms increases indefinitely. Finally, the equilibrium profits, gross of the fixed entry costs, decrease in n and tends to zero at the limit, due to the combined quantity and price effects.

This pattern characterizes the so-called *Cournotian paradigm*, a representation of the market equilibrium that depends on the number of firms and that moves from the monopoly to the perfectly competitive equilibrium as n increases from 1 to infinity. Perfect competition, in this setting, corresponds to the limiting case when each firm supplies an infinitesimal amount of output in a market populated by an infinite number of negligible firms.

This structural view of perfect competition can be easily derived from the first-order conditions that guarantee a profit-maximizing solution for any number of firms. Equation (11.2), indeed, implies that the market clearing price tends to the marginal cost when the last term vanishes. There are two possible explanations why $P'\frac{q^*}{S} \to 0$. One argues that when firms are small with respect to the market, they follow a *price-taking behavior*; that is, they expect the market price not to react to any change in their individual output. This case corresponds to assuming $P' = 0$ in a perfectly competitive market. The other explanation, which is consistent with the structuralist view of the Cournotian paradigm, instead focusses on the fact that it is the individual quantity that vanishes as n becomes indefinitely large, whereas $P' < 0$ even at the limit. In this latter case, indeed, $\lim_{n\to\infty} q^*(n) = 0$, as evident from (11.3).

It is interesting to notice that the last term in (11.2) also represents the negative externality that characterizes strategic interaction in a Cournot game, i.e. $\frac{\partial \Pi_i}{\partial q_j} = P'\frac{q^*}{S}$. In other words, with Cournot competition each firm affects the rivals' profits when it increases its quantity since it makes the price fall and reduces the revenues that the competitors obtain from their production. The level of individual production, therefore, multiplicatively affects this externality, which vanishes when each firm produces a negligible output. Then, a perfectly competitive market in a Cournotian perspective is also characterized at the limit by vanishing externalities across firms. This result confirms the idea that in a perfectly competitive market no externality occurs, a feature that is driven by the same effect ($\lim_{n\to\infty} q^*(n) = 0$) that explains why the competitive price tends to the marginal cost.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Finally, market size S increases individual and total quantities as well as the equilibrium profits.

3.2 Differentiated Products and Strategic Complements

A different class of oligopoly models moves into the realm of differentiated products and assumes that firms compete in prices, a framework that entails strategic complementarities. In the product differentiation literature, moreover, we can assume that either differentiation does not break the intrinsic symmetry of firms' market positions, or alternatively that product differentiation introduces a competitive advantage for some firms with respect to the others. The former case recalls the idea of (horizontal) differentiation by variety, where products differ in terms of characteristics, each one being more suited to a specific subset of customers. The latter, instead, captures the idea of (vertical) differentiation in quality. Given our focus on symmetric equilibria, in this section we shall consider several approaches to differentiation by variety. We shall consider entry and differentiation by quality in Section 7.

There are three main ways to model the demand side when products are (horizontally) differentiated: the representative consumer approach characterized by preference for variety; the discrete choice model where the external observer is able to reconstruct consumers' behavior up to a random component related to unobservable individual characteristics; and the address approach that assumes heterogeneous consumers with inelastic demand.[6]

Let $q_i = S * D_i(p_i, \mathbf{p}_{-i})$ be the demand for product $i \in I_n$, where S measures the size of the market and $\mathbf{p}_{-i}$ is the vector of prices other than p_i. Let us further assume $D_i(.)$ is continuous and differentiable and $C_i(D_i(.)) = cD_i(p_i, \mathbf{p}_{-i})$. Finally, let us assume that each firm offers only one variety.[7] Each firm solves the following problem: $\max_{p_i}(p_i - c)D_i(p_i, \mathbf{p}_{-i})$. Under standard assumptions on the strategy space being compact and convex, and the profit function being quasi-concave, the following equation identifies the necessary and sufficient conditions for a maximum:

$$\frac{p_i^* - c}{p_i^*} = \frac{D_i\left(p_i^*, \mathbf{p}_{-i}\right)}{p_i^* \frac{\partial D_i}{\partial p_i}} = \frac{1}{\varepsilon_i} \tag{11.4}$$

where ε_i is the price elasticity of demand for product i. In a symmetric equilibrium $p_i^* = p^*(n)$, $i \in I_n$, and

$$\varepsilon^*(n) = \frac{p^*(n)\frac{\partial D_i}{\partial p_i}}{D_i\left(p^*(n), \mathbf{p}^*(\mathbf{n})\right)}. \tag{11.5}$$

---

[6] For a detailed analysis of these three approaches and the relationships among them see Anderson, De Palma and Thisse (1992). On the representative consumer models see, for instance, the constant elasticity of substitution (CES) representation adopted in Spence (1976) and Dixit and Stiglitz (1977) and the linear representation in Shubik and Levitan (1980) and Singh and Vives (1984). On the interpretations of random utility models, we find two approaches: Manski (1977) assumes that utility is deterministic but it is not perfectly observed by the other agents, with a random term capturing the unobserved component; Quandt (1956) instead assumes the individual behavior to be intrinsically probabilistic. Finally, the address model approach was first proposed in Hotelling (1929). See also Salop (1979) and d'Aspremont et al. (1979).

[7] As we shall discuss in Section 4, assuming single-product firms makes the analysis of the maximum number of varieties and that of firms equivalent. With multi-product firms, instead, the maximum number of varieties will be larger than the number of active firms in a free entry equilibrium.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Hence, the pattern of equilibrium prices $p^*(n)$ when the number of firms increases depends inversely on the corresponding pattern of $\varepsilon^*(n)$. If $\lim_{n\to\infty} \varepsilon^*(n) = \infty$, then at the limit the price converges to the marginal cost, and we replicate the perfectly competitive equilibrium already found in the case of Cournot competition. When, instead, $\lim_{n\to\infty} \varepsilon^*(n) = \overline{\varepsilon}$ with $\overline{\varepsilon}$ finite, a positive mark-up persists at the limit, a pattern associated to Chamberlinian monopolist competition.[8] As we shall see, the limiting properties of the different approaches to product differentiation are consistent with either of the two alternatives.

Let us consider first the case of convergence to competitive equilibria. Generalizing the duopoly linear model originally proposed by Singh and Vives (1984) and further developed in Häckner (2000), the utility function of the representative consumer is quasi-linear according to the expression:

$$U(q_1, \ldots, q_n; I) = \alpha \sum_{i=1}^{n} q_i - \frac{1}{2}\left(\sum_{i=1}^{n} q_i^2 + 2\gamma \sum_{j\neq i} q_i q_j\right) + O \tag{11.6}$$

where $\gamma \in [0, 1)$ measures product substitutability and O is the money spent on outside goods. The demand system, then, is:

$$D_i(p_i, \mathbf{p}_{-i}) = S * \frac{\alpha(1-\gamma) + \gamma \sum_{j\neq i} p_j - \left[\gamma(n-2)+1\right] p_i}{(1-\gamma)\left[\gamma(n-1)+1\right]} \tag{11.7}$$

where S measures the size of the market, i.e. the number of representative consumers. Notice that in a symmetric price configuration $p_i = p$ for $i \in I_n$, firm i's demand

$$D_i(p, \mathbf{p}) = S * \frac{\alpha - p}{\left[\gamma(n-1)+1\right]}$$

decreases in the number of firms, since consumers spread their purchases over a larger set of varieties. The demand elasticity in a symmetric price equilibrium is:

$$\varepsilon^*(n) = \frac{\left[\gamma(n-2)+1\right] p^*(n)}{(\alpha - p^*(n))(1-\gamma)}. \tag{11.8}$$

Hence, $\lim \varepsilon^*(n) = \infty$ being $p^*(n) < \alpha$. Indeed, the equilibrium price

$$p^*(n) = \frac{\alpha(1-\gamma) + c\left[\gamma(n-2)+1\right]}{\gamma(n-3)+2} \tag{11.9}$$

[8] See Vives (1999), pp. 160–64 for a detailed discussion.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

tends to the marginal cost when $n \to \infty$. Moreover, the equilibrium quantity and profits

$$q^*(n) = S * \frac{(\alpha - c)\left[\gamma(n-2)+1\right]}{\left[\gamma(n-1)+1\right]\left[\gamma(n-3)+2\right]} \tag{11.10}$$

and

$$\Pi^*(n) = S * \frac{(\alpha - c)^2(1-\gamma)\left[\gamma(n-2)+1\right]}{\left[\gamma(n-1)+1\right]\left[\gamma(n-3)+2\right]^2} \tag{11.11}$$

are decreasing in the number of firms n.

A similar pattern can be obtained within the address models of product differentiation. Following Salop (1979) we can extend the original linear Hotelling duopoly to encompass n active firms by considering a circular market of length 1 where S consumers are uniformly distributed according to their individual preferred version t. Firms $i \in I_n$ produce at constant marginal cost c and sell at price p_i horizontally differentiated varieties that are evenly distributed at $x_i = i/n$ around the circle. Finally, a consumer of type t purchasing variety i has a net utility $u^* - p_i - (x_i - t)^2/\gamma$. We also use in this class of address models parameter γ to positively affect product substitutability. When γ is large the utility mostly depends on the price and the consumers are ready to switch to a more convenient, although more distant, variety. The demand system, in this setting, is given by:

$$D_i(p_i, p_{i-1}, p_{i+1}) = S\left[\frac{1}{n} - n\gamma p_i + \frac{n\gamma}{2}(p_{i+1} + p_{i-1})\right] \tag{11.12}$$

and displays localized competition between neighboring varieties, a notable feature of the address approach. The demand elasticity in a symmetric equilibrium is

$$\varepsilon^*(n) = \gamma n^2 p^*(n) \tag{11.13}$$

and $\lim_{n\to\infty} \varepsilon^*(n) = \infty$, implying convergence to the marginal cost. Notice also that, for given n, the elasticity is increasing in the substitutability parameter γ.

The symmetric equilibrium price, quantity and profits, indeed, are given by:

$$p^*(n) = c + \frac{1}{\gamma n^2}, \quad q^*(n) = \frac{S}{n} \quad \Pi^*(n) = \frac{S}{\gamma n^3}. \tag{11.14}$$

Comparing the symmetric equilibria in the Singh and Vives (1984) and in the Salop (1979) models of product differentiation with those obtained in the Cournot linear model we find significantly similar comparative statics properties, with price and individual quantity falling in the number of firms and the price approaching the marginal cost as the number of firms

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

tends to infinity. Indeed, the driving effect we highlighted in Cournot, based on vanishing individual quantities still applies. In the Salop model, however, an additional interesting effect is at work. When n increases indefinitely the market is completely covered with (locally) almost identical varieties. Localized competition between adjacent varieties reproduces a Bertrand environment, leading to marginal cost pricing. This latter effect corresponds to an increasingly intense price competition between closer and closer variety. In other words, in the localized competition model of product differentiation an increase in n produces at the same time a vanishing quantity externality and an increasing price externality, both pushing towards convergence to a competitive outcome.

We can now turn to the case of monopolistic competition, when positive mark-ups are associated with a market populated by a very large (i.e. infinite) number of infinitesimal firms. We illustrate this case referring to the multinomial logit model, thereby also covering the discrete choice approach to product differentiation. Let the utility of a consumer be described by a deterministic component $U(p_i) = \alpha - p_i$ and an additive random independent and identically distributed (i.i.d.) component η_i that is distributed according to the double exponential distribution $F(x) = \exp - \left[\exp -(\gamma x + \epsilon\right]$ where ϵ is Euler's constant and γ a positive constant that negatively affects the variance. Then, the resulting probability of choosing product i given the vector of prices $(p_1, \ldots, p_n)$ is

$$P_i(p_i, \mathbf{p}_{-i}) = \frac{\exp(-\gamma p_i)}{\sum_{j=1}^{n} \exp(-\gamma p_j(\mu)}. \tag{11.15}$$

Then firm i's expected profits are:

$$\Pi_i(p_i, \mathbf{p}_{-i}) = S * (p_i - c) P_i(p_i, \mathbf{p}_{-i}).$$

We can observe that $\frac{\partial P_i}{\partial p_i} = \gamma P_i(1 - P_i)$ and that, therefore, parameter γ, once again, captures product substitutability. Moreover, in a symmetric equilibrium $P_i(p, \mathbf{p})=\frac{1}{n}$. Then, the elasticity of demand is

$$\varepsilon^*(n) = \frac{\gamma(n-1)p^*(n)}{n}, \tag{11.16}$$

with $\lim_{n\to\infty} \varepsilon^*(n) = \gamma p^*(n)$ finite.[9] Hence, the firms obtain a positive mark-up when n tends to infinity. The equilibrium price, quantity and profits are:

$$p^*(n) = c + \frac{n}{\gamma\,(n-1)}, \quad q^*(n) = \frac{S}{n}, \quad \Pi^*(n) = \frac{S}{\gamma\,(n-1)}. \tag{11.17}$$

9 Parameter γ, as in the previous models, positively affects price elasticity for given n.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

The multinomial logit model[10] presents a different pattern of price adjustment, with the equilibrium price decreasing in the number of firms and converging to a mark-up $1/\gamma$ when $n \to \infty$. Despite the positive mark-up, the firm's profits vanish at the limit, since the individual output becomes negligible, as it is in a monopolistic competition environment. We can also notice that the basic channel of interaction across firms vanishes as well at the limit: $\frac{\partial P_i}{\partial p_j} = \gamma P_i P_j = \frac{\gamma}{n^2}$. Hence, the "competitive" component of monopolistic competition is associated with vanishing externalities, as already observed when discussing the Cournot model.

To sum up, the different models of product differentiation display similar comparative static properties with respect to the number of firms, with the equilibrium price, quantity and profits falling in n. The main difference rests on the convergence of the equilibrium prices to the marginal cost, as in a perfectly competitive market, or instead to a positive mark-up over costs that characterizes monopolistic competition. Moreover, the size of the market, in all cases, pushes up profits.

The results of the product differentiation literature provide an additional insight related to the intensity of price competition and its effect on n-firms market equilibria. In the three models, with a little abuse of notation, we have represented product substitutability through parameter γ, with the price elasticity increasing and the price and profits falling in γ.

3.3 Explaining the Comparative Statics in a Unified Framework

In the previous sections we have shown that the market equilibria, described by prices and quantities, share similar comparative statics properties across a wide range of different oligopoly models and features of preferences and technology. This raises a natural question of whether this common pattern may be accounted for through a unified explanation. The theory of monotone comparative statics developed by Milgrom and Roberts (1990, 1994) and Milgrom and Shannon (1994) offers an enlightening perspective. Their approach allows the development of new tools with which to study how equilibria change in reaction to a variation in the parameters and constraints of the maximization problem, moving beyond the tradition approach based on the implicit function theorem.[11] Quoting Amir (2003, p. 2), "if in a maximization problem, the objective reflects a complementarity between an endogenous variable and an exogenous parameter, in the sense that having more of one increases the marginal return to having more of the other, then the optimal value of the former will be increasing in the latter. In the case of multiple endogenous variables, then all of them must also

[10] A similar result is obtained, within the representative consumer approach, assuming Cobb-Douglas preferences between a numéraire good q_0 and a set of differentiated products q_i with CES preferences:

$$U(q_o, q_{1,\ldots}, q_n) = q_0^{1-\beta}\widetilde{q}^{\beta} \quad \text{with } \gamma \in (0,1)$$

and

$$\widetilde{q} = \left(\sum_{i=1}^{n} q_i^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}.$$

See Spence (1976), Dixit and Stiglitz (1977) and Anderson et al. (1992) pp. 226–9.

[11] Importantly, the new tools help to deal with the comparative statics of multiple equilibria, by studying how extremal equilibria move in reaction to a change in exogenous variables. For the purpose of our discussion, however, we shall focus on the case of unique equilibria.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

be complements so as to guarantee that their increases are mutually reinforcing". The former property corresponds to increasing differences (between the endogenous and the exogenous variables, and more in general between two variables), whereas the latter qualifies the function to be maximized as supermodular.[12]

When a game is supermodular and characterized by increasing differences, an increase in the strategy of one player increases the marginal payoff of the strategy of the other players, inducing them to adjust their optimal choice upwards. This case, therefore, corresponds to upward-sloping reaction functions or, in the classification of Bulow, Geanakoplos and Klemperer (1985), strategic complementarity. Moreover, increasing differences between the endogenous variables and the exogenous variable implies that an increase in the exogenous variable increases the marginal payoff of the strategy of the players, with an upwards shift in the best-reply functions.

Increasing differences then can be easily turned into decreasing differences by reverting the sign of the adjustment or defining a new exogenous variable that is the negative of the original one. In this case, an increase in the exogenous variable induces a contraction in the endogenous one.[13]

We can borrow from the theory of monotone comparative statics two conditions, described in the statements of Theorem 5 and 6 in Milgrom and Shannon (1994) that, in our setting, fit the problem. The exogenous variable[14] is the number of firms n whereas the endogenous variables are, depending on the model specification, the quantities q_i or prices p_i. Then, we require the profit functions to be supermodular in the strategic variables and to display decreasing differences in the number of firms. Since we consider continuous and differentiable functions, the two conditions correspond to $\frac{\partial \Pi_i}{\partial a_i \partial a_j} > 0$ and $\frac{\partial \Pi_i}{\partial a_i \partial n} < 0$ for $i, j = I_n$, $i \neq j$, where a_i describes firm i's strategy, i.e. quantity or price. Moreover, in order to focus on the comparative statics, we take for granted that an equilibrium exists and is unique, by assuming that the profit function is strictly quasi-concave in the choice variable and that the best-reply slope meets the contraction mapping requirement.

Starting with the Cournot case, a first problem arises since in the traditional description competition is in strategic substitutes, and the game is submodular rather than supermodular.[15] A way out of this problem borrows from an early intuition in Novshek (1985) and is developed in Amir and Lambson (2000). Indeed, a notable property of the Cournot model is that the profits can be expressed as a function of own output q_i and of the aggregate level of output of the other $n - 1$ firms $Q_{-i} = \sum_{j \neq i} q_j$, i.e.

$$\Pi_i(q_i, Q_{-i}) = P(q_i + Q_{-i})\, q_i - C(q_i).$$

[12] See Vives (1999), Chapter 2. When the payoff functions are smooth and the strategy space of each firm and the exogenous parameters space are one-dimensional, supermodularity and increasing differences boil down to the condition that the second cross-partials between each firm's strategic variable and the other firm's strategic variable and with the exogenous parameter are positive.

[13] Increasing differences is a cardinal property and can be replaced by the ordinal Spence-Mirlees single-crossing property considered in Milgrom and Shannon (1994). When this property holds, if an increase in the choice variable is profitable when the exogenous variable is low it is still profitable when the exogenous variable is high, although it is not required, as in the case of increasing differences, that the profitability is higher in the latter case.

[14] Here for convenience we measure the number of firms n as a continuous variable defined on the positive reals.

[15] While in a Cournot duopoly this issue is easily adjusted by describing one of the strategies as $-q$, transforming the setting into a supermodular game, with $n > 2$ firms this trick cannot be applied anymore.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Moreover, we can equivalently describe firm i's strategy, rather than refer to the choice of its own output q_i, as the selection of a certain level of total output Q for given output Q_{-i} supplied by the competitors. In this alternative formulation

$$\widehat{\Pi}_i(Q, Q_{-i}) = P(Q)(Q - Q_{-i}) - C\left((Q - Q_{-i})\right). \tag{11.18}$$

Then,

$$\frac{\partial^2 \widehat{\Pi}_i}{\partial Q \partial Q_{-i}} = C''(Q - Q_{-i}) - P'(Q) = \Delta, \tag{11.19}$$

which corresponds to (11.1). Then, the condition $\Delta > 0$ implies the supermodularity of the modified Cournot game. Decreasing differences can be easily established by noting that when the other $n-1$ firms choose the same output q then $Q_{-i} = (n-1)q$. Then, substituting in the first-order conditions for the choice of Q in the modified Cournot problem we have:

$$\frac{\partial \widehat{\Pi}_i}{\partial Q} = P'(Q)\left(Q - (n-1)q\right) + P(Q) - C'\left(Q - (n-1)q\right). \tag{11.20}$$

Hence,

$$\frac{\partial^2 \widehat{\Pi}_i}{\partial Q \partial n} = q\Delta > 0 \tag{11.21}$$

when the game is supermodular. We conclude that the equilibrium total output $Q^*(n)$ is increasing in the number of firms. In a symmetric equilibrium $Q^*_{-i}(n) = \frac{n-1}{n} Q^*(n)$, and therefore the output of the firms other than i is increasing in n as well, since both terms $\frac{n-1}{n}$ and $Q^*(n)$ are positive and increasing in n. Moreover, since firm i's best reply in the original Cournot problem is downward sloping and $Q^*_{-i}(n)$ is increasing in n, the individual output $q_i^*(n)$ is decreasing in the number of firms. Finally, since demand is bounded, when $n \to \infty$ we must have $Q^*(n) = nq^*(n)$ finite and therefore $\lim_{n\to\infty} q^*(n) = 0$. Then, given the first-order conditions of the original Cournot problem, $p^*(n) \to C'\left(q^*(n)\right)$.

Our discussion offers a clear insight into the advantages of the techniques of monotone comparative statics. A single and general condition, $\Delta = C''(q_i) - P'(Q) > 0$, generates supermodularity of the modified Cournot problem and $Q^*(n)$ and $Q^*_{-i}(n)$ increasing in the number of firms, while the comparative statics on individual output $q_i^*(n)$ and the limiting competitive result on the price derive from the first-order conditions of the original Cournot problem. Interestingly, the condition $\Delta > 0$ includes elements of demand and costs, and both jointly define the relevant condition. This extends with respect to previous contributions that explored the properties of Cournot equilibria by making specific assumptions on costs or demand.[16]

---

[16] See Amir and Lambson (2000) for a general analysis of equilibria in Cournot games.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Turning to the models of product differentiation and price competition, in an n-firm oligopoly each one solves $\max_{p_i} p_i D_i(p_i, \mathbf{p}_{-i}; n) - C(D_i(.))$ where we emphasize that, differently from the homogeneous product case, the number of substitute products n may directly enter into the expression of the demand for product i. Moreover, notice that in our symmetric environment we assume that all firms have the same cost structure, i.e. $C_i(D_i(.)) = C(D_i(.))$.

If

$$\frac{\partial^2 \Pi_i}{\partial p_i \partial p_j} = \frac{\partial D_i}{\partial p_j} + (p_i - C')\frac{\partial^2 D_i}{\partial p_i \partial p_j} - C''\frac{\partial D_i}{\partial p_j}\frac{\partial D_i}{\partial p_i} > 0, \tag{11.22}$$

for any $i,j = I_n$, $i \neq j$, the game is in strategic complements, that is the condition for supermodularity is met. Then, the equilibrium prices fall in the number of firms if

$$\frac{\partial^2 \Pi_i}{\partial p_i \partial n} = \frac{\partial D_i}{\partial n} + (p_i - C')\frac{\partial^2 D_i}{\partial p_i \partial n} - C''\left(\frac{\partial D_i}{\partial p_i}\right)^2 < 0.$$

Substituting the first-order conditions $p_i - C' = -\frac{D_i}{\partial D_i/\partial p_i}$ and rearranging we get:

$$\frac{\partial^2 \Pi_i}{\partial p_i \partial n} = \frac{\partial D_i}{\partial n} + \frac{p_i}{\varepsilon_p}\frac{\partial^2 D_i}{\partial p_i \partial n} - C''\left(\frac{\partial D_i}{\partial p_i}\right)^2. \tag{11.23}$$

Differentiating the elasticity of demand with respect to n, we obtain:

$$\frac{\partial \varepsilon_p}{\partial n} = -\frac{\varepsilon_p}{D_i}\left[\frac{\partial D_i}{\partial n} + \frac{p_i}{\varepsilon_p}\frac{\partial^2 D_i}{\partial p_i \partial n}\right].$$

Hence, we can rewrite (11.23) as

$$\frac{\partial^2 \Pi_i}{\partial p_i \partial n} = -\frac{D_i}{\varepsilon_p}\frac{\partial \varepsilon_p}{\partial n} - C''\left(\frac{\partial D_i}{\partial p_i}\right)^2. \tag{11.24}$$

Then, if (11.22) holds and (11.24)<0 for all $i \in I_n$, the symmetric equilibrium prices fall in the number of firms. We can notice that the conditions (11.22) and (11.24) display a combination of demand and cost elements, a feature already noticed in the Cournot model. For instance, if the marginal costs are not decreasing and the demand elasticity is increasing in the number of firms, then the conditions are met.

Turning to our three examples of differentiated products models referred to in the different approaches, we have directly derived the equilibrium prices and observed that they fall in the number of firms. It is easy to check that the two conditions (11.22) and (11.24) are satisfied in our examples. Indeed, we assumed in the examples linear costs, i.e. $C'' = 0$. Moreover, it can be easily verified that when the other $n-1$ firms set the same price p, the elasticity of demand

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

is increasing in n. Hence, the game features supermodularity and increasing differences and the prices fall in n.

3.4 Collusive Equilibria

We conclude our review of n-firms' oligopolies by considering the case of collusive equilibria. We refer to the infinite horizon repeated game approach pioneered by Friedman (1971) and further developed in Fudenberg and Maskin (1986). Since we are considering symmetric oligopolies, we assume that the basic market interaction can be represented in each period $t = 1, .., T$ by a symmetric and stationary constituent game $\Gamma^t = \{I_n, a_i^t \in A, \pi_i^t = \pi(\mathbf{a}^t)\}$, where I_n is the set of n firms, $\mathbf{a}^t = (a_i^t, \mathbf{a}_{-i}^t)$ is the vector of actions chosen by firm i and the other $n - 1$ firms at time t, A is the set of feasible actions and $\pi_i^t = \pi(\mathbf{a}^t)$ the per-period payoff. We further assume that Γ^t has a unique symmetric Nash equilibrium $\widehat{\mathbf{a}} = (\widehat{a}, .., \widehat{a})$ that is Pareto dominated by other market configurations $A^{n*} = \{(a_i^*, \mathbf{a}_{-i}^*) \in A^n \,|\pi(a_i^*, \mathbf{a}_{-i}^*) \geq \pi(\widehat{\mathbf{a}})\ \forall i \in I_n\}$. Let $\overline{\mathbf{a}}^*$ be the maximal collusive symmetric configuration. The firms maximize the discounted sum of profits $V_0 = \sum_{t=0}^{T} \delta^t \pi_i^t$, where $\delta = 1/(1 + r)$ is the discount factor. Each firm observes the other firms' actions with a one-period lag. The set of observed actions at time t, the history of the game, then, is $H^t = \{\mathbf{a}^0, .., \mathbf{a}^{t-1}\}$.

In what follows we concentrate on symmetric collusive equilibria, in the spirit of the overall section. Let a^C be firm i's collusive action, $\mathbf{a}^C \in A^{n*}$ be the vector of collusive actions, and $\pi^C = \pi(\mathbf{a}^C)$ the corresponding individual profits. Notice[17] that $\mathbf{a}^C \in [\overline{a}^*, \widehat{\mathbf{a}}]$; that is, the collusive symmetric allocation is in between the Nash equilibrium and the maximal collusive allocation. Further, define $a^P = \widehat{a}$ firm i's action during the punishment phase, corresponding to the symmetric Nash equilibrium action in the constituent game, and $\pi^P = \pi(\widehat{\mathbf{a}})$ the punishment individual profits. Finally, let $a^D = \arg\max_{a_i} \pi(a_i, \mathbf{a}_{-i}^C)$ be firm i's optimal deviation when the other firms stick to the collusive action, yielding $\pi^D = \pi(a^D, \mathbf{a}_{-i}^C)$. Our previous discussion implies that $\pi^P \leq \pi^C \leq \pi^D$ with strict inequalities if $a^C < \widehat{a}$. We focus on closed-loop grim-trigger strategies:

$$\sigma_i^* = \begin{cases} a_i^t = a^C & \text{for } t = 0 \\ a_i^t = a^C & \text{for } t > 0 \text{ and } H^t = \{\mathbf{a}^C, .., \mathbf{a}^C\}. \\ a_i^t = a^P & \text{for } t > 0 \text{ and } H^t \neq \{\mathbf{a}^C, .., \mathbf{a}^C\} \end{cases}$$

When $T = \infty$ (infinite horizon), given the strategy followed by the other firms and the stationarity of the repeated game each firm chooses to collude if the following incentive compatibility constraint holds:

$$V^C = \frac{\pi^C}{1 - \delta} \geq V^D = \pi^D + \frac{\delta}{1 - \delta}\pi^P.$$

[17] We implicitly assume in this notation that $\widehat{a} > a^*$, as is the case if the action corresponds to an output level. If, instead, the action corresponds to a price, the boundaries of the interval should be inverted.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Then, a well-known result (Folk theorem) states that any allocation $\mathbf{a}^* \in A^{n*}$ can be implemented as a subgame perfect equilibrium in the game repeated indefinitely ($T = \infty$) if the following condition holds[18] for all firms $i \in I_n$:

$$\delta \geq \delta^* = \frac{\pi^D - \pi^C}{\pi^D - \pi^P}. \tag{11.25}$$

We can now address the key issue of whether the price(s), quantities and profits change, and in which direction, when the number of firms increases. To answer these questions we can consider two examples of market interaction when firms offer homogeneous products, characterizing the constituent game Γ^t as a price-setting Bertrand game or a quantity-setting Cournot game. Let $\Pi^C = n\pi^C$ be the total profits of the cartel. Then, in a Bertrand setting $\pi^C = \Pi^C/n$, $\pi^D = \Pi^C$ and $\pi^P = 0$. Then, the condition (11.25) boils down to

$$\delta \geq \delta^*(n) = \frac{n-1}{n}$$

that is increasing in n. In other words, if the basic market interaction takes the form of Bertrand competition with homogeneous products, the incentive compatibility constraint becomes tighter the larger the number of firms. The economic intuition is pretty simple: a cartel with more members distributes the overall profits Π^C among a larger number of participants, making the individual profits fall. Deviation and punishment profits, in this setting, are instead unaffected by the number of cartel members, making the condition for cartel sustainability harder to meet. We can further observe that the incentive compatibility constraint does not depend on the specific (symmetric) collusive allocation $\mathbf{a}^C$ the cartelists agree upon, since the gains from deviations are always proportional to the collusive profits. Then, a focal outcome would be to mimic the monopoly price p^m. Our prediction, then, is that the market price will be p^m if the number of firms is $n \leq \frac{1}{1-\delta}$, falling to the Nash equilibrium price $p = c$ thereafter. To sum up, individual profits are strictly decreasing and the market price is weakly decreasing in the number of firms.

Turning to the Cournot model, we can identify a further element in the comparative statics. Indeed, in a Cournot setting the profits in the different states vary nonproportionally in the collusive allocation Q^C the firms choose to implement. More precisely, the incentive compatibility constraint becomes tighter when the firms coordinate on an allocation, summarized by total output Q^C; that is, closer to the monopoly output Q^m. Hence, in a Cournot setting the critical discount factor $\delta^*(Q^C, n)$ is decreasing in the collusive output Q^C, whereas it continues to be increasing in the number of firms n.[19] The most collusive sustainable output in a symmetric cartel, $\overline{Q}^C$, then, is (weakly) increasing in the number of firms: if we start from $\overline{Q}^C = Q^m$, we can find a number of firms $n(Q^m, \delta)$ such that $\delta^*(Q^m, n(Q^m, \delta)) = \delta$. For a larger number of firms the cartel would collapse. However, the firms can coordinate on a less

[18] Notice that, having assumed symmetric firms, the incentive compatibility constraint and the threshold discount factors are the same for each and every firm.

[19] For instance, it is easy to show that, in the linear Cournot model when firms implement the monopoly output the critical discount factor is $\delta^* = \frac{n^2+2n+1}{n^2+6n+1}$ and is therefore increasing in n.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

collusive output (i.e. $\overline{Q}^C > Q^m$) such that the incentive compatibility constraint is satisfied. In general, when (11.25) holds as an equality, for given δ we have

$$\frac{d\overline{Q}^C}{dn} = -\frac{\frac{\partial \delta^*}{\partial n}}{\frac{\partial \delta^*}{\partial \overline{Q}^C}} \geq 0.$$

Hence, for $n \leq n(Q^m, \delta)$ the individual profits are decreasing in n while the market price is p^m, whereas for $n > n(Q^m, \delta)$ both the individual profits and the market price are falling in n.

Finally, an informal argument that is often put forward refers to the impact of a larger and larger cartel on the monitoring activity that the firms have to perform to prevent cheating. It seems realistic that such activity may take more time the higher the number of firms to be scrutinized. We can include this further argument by considering that the length of the period in the repeated game framework may increase when more firms participate in the agreement and have to be monitored. A longer period, then, corresponds to a lower discount factor δ, leading to a decreasing relationship $\delta(n)$. In this latter case, the incentive compatibility constraint would become $\delta^*(\overline{Q}^C, n) \geq \delta(n)$ and the effect of the number of firms on the maximal collusive allocation would be

$$\frac{d\overline{Q}^C}{dn} = -\frac{\frac{\partial \delta^*}{\partial n} - \frac{\partial \delta}{\partial n}}{\frac{\partial \delta^*}{\partial \overline{Q}^C}} \geq 0,$$

implying a stronger expansion in the cartel output when n increases. Finally, when $n \to \infty$ both π^P and π^C tend to zero and the only sustainable output $\overline{Q}^C$ becomes the competitive one.

The effect of market size S on collusive equilibria is twofold. Under constant marginal costs, market size and the scale of production multiplicatively affect the profits in each of the relevant states. Then, S cancels out in the expression of the critical discount factor. In other words, under constant marginal costs the incentive compatibility constraints are unaffected by market size. On the other hand, the level of collusive equilibrium profits π^C increase with market size.

To sum up, even the cartel equilibria display comparative statics properties similar to those already highlighted: the individual profits decrease, as does the market price, when the number of firms increases, and they tend to the perfectly competitive output when $n \to \infty$. Market size positively affects collusive profits while being neutral on the conditions for sustainability of the cartel. Moreover, the level of profits in a cartel are higher, for a given number of firms, than those of the oligopoly equilibria analyzed in the previous sections.

4 FREE ENTRY SYMMETRIC EQUILIBRIA

We can now endogenize the entry decision that determines how many of the m potential entrants will decide to become active, sinking the entry cost F. In a symmetric setting, the post-entry profits depend on the number of active firms n and are decreasing in it, as analyzed in detail across a wide set of models in the previous section. We can summarize the main findings in the relationship $\Pi(n, S, \gamma)$ between the individual profits, the number of firms n,

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

the market size S and the variable γ that captures the intensity of price competition. This latter, therefore, can be referred to as the degree of substitutability among differentiated products, as in Section 3.2, as well as the mode of competition (price, quantity, collusion). Hence, the individual profits are decreasing in the number of firms, increasing in market size and decreasing in the intensity of competition.

The maximum number of firms n^* in a symmetric free entry equilibrium (SFEE) is then captured by the two conditions:

$$\Pi(n^*, S, \gamma) \geq F \qquad (11.26)$$
$$\Pi(n^* + 1, S, \gamma) < F$$

The former ensures that all the active firms make non-negative net profits, whereas the latter implies that in a market equilibrium with $n^* + 1$ firms each one would not cover the sunk entry costs. Given the monotonicity of the individual profits in n we can therefore write[20]

$$n^* = n(S, F, \gamma), \qquad (11.27)$$

where

$$\frac{\partial n^*}{\partial S} = -\frac{\partial \Pi/\partial S}{\partial \Pi/\partial n} > 0, \ \ ,\frac{\partial n^*}{\partial F} = \frac{1}{\partial \Pi/\partial n} < 0 \ \text{ and } n^*(\gamma') < n^*(\gamma) \ \text{ if } \gamma' > \gamma. \qquad (11.28)$$

Hence, our main predictions state that the number of firms in a symmetric free entry equilibrium is increasing in market size, decreasing in the sunk entry costs (economies of scale) and decreasing in the intensity of competition.[21] Interestingly, relaxed competition (a lower γ), as may arise if products are weak substitutes, or in the case that the industry is cartelized, is concomitant with an increased number of firms. We can further notice that if marginal costs are constant, market size multiplicatively increases the profits and therefore the number of firms depends on the ratio F/S that captures the relevance of economies of scale with respect to market size. Then, an increase in market size, as it may derive from free trade agreements, leads to an increase in the number of firms and a fall in prices, making consumers better off.

The SFEE identifies the maximum number of firms sustainable given market fundamentals and the prevailing strategic behavior. More specifically, in differentiated products markets we have identified the maximum number of varieties sustainable in an SFEE, assuming that each variety requires to sink a cost F to be produced, whereas the number of firms may be lower if some of them offer a portfolio of different varieties.[22]

---

[20] We consider here for convenience n as defined on $\mathbb{R}^+$, ignoring the integer issue. Then, given the monotonicity of profits in n the two conditions for an SFEE boil down to $\Pi(n^*, S, \beta) = F$.

[21] We express the relationship between n^* and γ to encompass both the case when γ is defined over a compact interval (the substitutability parameter in the differentiated products models) and when it is a discrete index measuring the intensity of competition (as when comparing collusive and non-cooperative equilibria).

[22] This statement should be further qualified according to the different models of product differentiation. In general, if in a symmetric multi-product setting each firm offers k varieties some cross-variety effects are internalized, and therefore the market price should be different (higher) than in the case of single-product firms. With higher individual profits in the symmetric k-varieties firms equilibrium some further entry should be profitable. Therefore, the number of multi-product firms should be larger than n^*/k, where n^* is the SFEE number of single-product firms.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

5 FREE ENTRY AND SOCIAL EFFICIENCY

Moving from the positive to the normative analysis, we are interested in evaluating whether the entry process leads to an optimal, excessive or insufficient number of firms. A frequent presumption is that guaranteeing conditions of free entry is desirable from a social point perspective. The analysis we have developed in the previous sections allows us to address this issue and to verify whether and under which conditions free entry leads to socially desirable outcomes. Spence (1976) and Dixit and Stiglitz (1977) have explored the issue in a monopolist competition set-up, finding that the number of varieties in a free entry equilibrium falls short of the social optimum. In a homogeneous product environment, instead, Von Weizsäcker (1980) and Perry (1984) established an opposite result, with too many firms entering with respect to the social optimum.

We discuss the social efficiency of SFEE following Mankiw and Whinston (1986) and Amir, De Castro and Koutsougeras (2014) and adopting the same two-stage game of the previous sections. We analyze a second-best welfare maximization problem where the social planner is assumed to control the number of firms but to be unable to affect or determine the behavior of the active firms once they enter. In the comparison of the equilibrium and the socially optimal number of firms we focus on the case when the fixed costs are non-negligible given market size, and the number of firms in either solution is finite.

We start with the case of homogeneous products and quantity competition and then move to a product differentiation and price competition environment. We can borrow from the analysis of symmetric market equilibria three conditions that we proved to hold under fairly general conditions in the Cournot model:[23]

1. In the symmetric equilibrium the individual output is decreasing in n: $q(n) > q(n')$ for $n' > n$.
2. Total output is increasing in the number of firms: $Q(n) = nq(n) < Q(n') = n'q(n')$ for $n < n'$.
3. The price cost margin is non-negative for any number of firms, and strictly positive for a finite number of firms: $P\left(Q(n)\right) - C'\left(q(n)\right) \geq 0$ for all n and $P\left(Q(n)\right) - C'\left(q(n)\right) > 0$ for n finite.

Given these features, the social planner maximizes total welfare by choosing the number of firms:

$$\max_n W(n) = \int_0^{Q(n)} P(s)ds - nC\left(q(n)\right) - nF \tag{11.29}$$

[23] In their paper, Mankiw and Whinston do not model explicitly the post-entry game and assume that certain features characterize the firm and aggregate pattern of the equilibrium strategies. We can, instead, explicitly refer to the properties of the equilibria developed in the previous sections. A similar approach can be found in Amir et al. (2014).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Let us define n^W as the solution. Then, under 1–3, the SFEE number of firms is higher than the social optimum, that is $n^* > n^W$. The result can be easily proved by noting that the first-order conditions in problem (11.29) are:

$$W'(n) = P(.)\left[n\frac{\partial q}{\partial n} + q(n)\right] - C(q) - nC'(q)\frac{\partial q}{\partial n} - F = \tag{11.30}$$

$$= \Pi(n) - F + n\left[P(Q(n)) - C'(q(n))\right]\frac{\partial q}{\partial n}.$$

Since in SFEE $\Pi(n^*) = F$, $\frac{\partial q}{\partial n} < 0$ by condition 1 and $P(Q(n^*)) - C'(q(n^*)) > 0$ for n^* finite given condition 3, it follows that $W'(n^*) < 0$ and therefore $n^* > n^W$.

The economic intuition of the excessive entry result is straightforward: when an additional firm enters, it adds to the social welfare the profit $\Pi(n) - F$ but, at the same time, it steals output, and therefore profits, from the other firms, the last term in the derivative (11.30), second line. The *business-stealing effect*, captured by condition 1 above, creates a wedge between the private incentives of the entrant, and the social effect of entry, explaining why too many firms enter in an SFEE.[24] We can observe that when F (or F/S) tends to zero then $P(Q(n^*)) - C'(q(n^*))$ and $\Pi(n^*)$ vanish, implying that an infinite number of firms enter in equilibrium and maximize welfare. In other words, the excessive entry result applies to the case of significant fixed costs and a finite number of firms, whereas it vanishes when fixed costs become negligible. A policy that expands markets, as it is a free trade approach, therefore can fix the excessive entry distortion and realign competitive market outcomes and social optimality.

The case of differentiated products adds an additional effect of entry on welfare, since more firms imply a larger set of varieties available to the consumers. Following Spence (1976) we capture this effect by assuming that the gross consumers' benefit is

$$CS(\mathbf{q}) = G\left[\sum_{i=1}^{n} f(q_i)\right] \tag{11.31}$$

where $\mathbf{q}$ is the vector of outputs, $f(0) = 0$, $f'(.) > 0$ and $f''(.) \leq 0$ for all $q_i \geq 0$ implies a preference for variety and $G'(z) > 0$, $G''(z) < 0$ for all $z \geq 0$ qualifies products as substitutes.[25] The social planner then solves the problem

$$\max_n W(n) = G\left[nq(n)\right] - nC(q(n)) - nF.$$

Contrary to the case of homogeneous products, when products are differentiated in general we cannot rank the number of firms in an SFEE and the socially optimal one. The reason is immediately evident from the first-order conditions of the problem:

$$W'(n) = G'\left(nf'\frac{\partial q}{\partial n} + f\right) - C(q) - nC'(q)\frac{\partial q}{\partial n} - F = \tag{11.32}$$

$$= \Pi(n) - F + n\left(G'f' - C'\right)\frac{\partial q}{\partial n} + G'\left(f - f'q\right)$$

[24] Mankiw and Whinston show that, when the integer problem is taken into account, $n^* \geq n^W - 1$.

[25] Consumers' utility maximization implies that in a symmetric equilibrium the price is equal to $G'(nf(q))f'(q)$ and therefore the profits can be written as $\Pi = G'(nf(q))f'(q)q - C(q) - F$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Condition (11.32) shows that an additional firm adds to total welfare the profits generated, $\Pi(n) - F$, and further affects total welfare with two additional terms. The first corresponds to the business-stealing effect already identified in the case of homogeneous products, and captures the fact that the new firm subtracts output and profits to the competitors, with a lower net social gain than the private firm and a bias towards excessive entry.

The last term is new and refers to the impact of an additional variety on consumers' surplus. $G'f$ is the marginal social effect of the new variety, whereas $G'f'q$ is the firm revenue. Since the firm does not internalize all the social benefit of the additional variety, the private incentives are lower than the social ones, leading to underprovision of varieties.

Without specific assumptions on preferences the two terms with opposite signs in (11.32) do not allow the identification of $W'(n^*)$. Therefore, we may have an excessive, insufficient or optimal number of firms entering in an SFEE. Under more specific assumptions on the utility function, we can generate examples where the ranking can be established. For instance, Dixit and Stiglitz (1977), using a CES utility function, obtain that the SFEE number of firms is lower than the welfare-maximizing one, reverting the case of excessive entry that characterizes a homogeneous product environment.

6 FREE ENTRY EQUILIBRIA WITHOUT SYMMETRY

Although a symmetric environment is a natural reference when analyzing long-run free entry equilibria, we may be interested in the effects of free entry in oligopoly markets when some kind of asymmetry has long-lasting effects. This may come from the existence of patents or other frictions in the adoption of process innovations that prevent the equalization of production techniques, from persisting advantages on the demand side coming from quality or brand image, to institutional features that affect the behavior of firms, such as, for instance, the coexistence of different ownership structures or the presence of state-owned firms. Since free entry equilibria suggest the pattern of adjustment when the entry process unfolds, asymmetric oligopolies are an interesting and relevant case to be addressed.

Once firms intrinsically differ, the number of firms is no longer a relevant statistic with which to describe, in a positive or normative sense, the long-run equilibria. However, many of the oligopoly models we have already considered in a symmetric setting share a particular property: that of being aggregative games, which allows us to deal easily with asymmetric environments.[26]

The profits of firm i in an aggregative oligopoly game can be written as a function of a choice variable (action) a_i and of the sum of the actions of *all* market participants $A = \sum_{j=1}^{n} a_j$; that is, $\Pi_i(a_i, A)$. A very simple illustration is the Cournot model already considered in Section 3.1. Setting $q_i = a_i$ we can write $\Pi_i(a_i, A) = P(A)a_i - C_i(a_i)$. We also recognize an aggregative structure in some of the models of product differentiation.[27] In the Singh and

[26] See Anderson, Erkal and Piccinin (2015) on free entry equilibria with aggregative oligopoly games.

[27] One can notice that address models with $n > 3$, such as the Salop circular road model described above, are not aggregative games, since the profits of each firm depend only on a subset of prices.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Vives (1984) linear model the prices enter additively in the demand function and therefore, setting $p_i = a_i$, the profits are written as:

$$\Pi_i(a_i, A) = (a_i - c)\frac{\alpha(1-\gamma) + \gamma A - \left[\gamma(n-1)+1\right]a_i}{(1-\gamma)\left[\gamma(n-1)+1\right]}.$$

Even the logit model shares the feature of an aggregative game, once we define $a_i = \exp(-\gamma p_i)$: the profits can be written as

$$\Pi_i(a_i, A) = (-\log(a_i)/\gamma - c_i)\frac{a_i}{A}.$$

To illustrate the main features of aggregative games, we use here the linear Cournot model $\Pi_i(q_i, Q) = (a - bQ - c_i)q_i$ as an example. The traditional setting describes the profit function as depending on own output and the aggregate of other firms' production $Q_{-i} = \sum_{j \neq i} q_j$; that is, $\Pi_i(q_i, Q_{-i}) = (a - b(q_i + Q_{-i}) - c_i)q_i$ and identifies the best reply

$$\widehat{q}_i(Q_{-i}) = \frac{a - c_i}{2b} - \frac{Q_{-i}}{2}.$$

Alternatively, following the aggregative setting we can identify the inclusive best reply first introduced by Selten (1970), where the optimal individual output is consistent with a given aggregate level of production:[28]

$$\widetilde{q}_i(Q) = \frac{a - c_i}{b} - Q.$$

Notice that an equilibrium exists only if $\sum_{i=1}^{n} \widetilde{q}_i(Q) = Q$; that is, if the sum of the inclusive best replies has a fixed point.[29] Further we can define firm i's profits, when it and all firms choose their inclusive best reply, as a function of total output Q:

$$\Pi_i(Q, \widetilde{q}_i(Q)) = \Pi_i^*(Q) = \frac{(a - c_i - bQ)^2}{b} \tag{11.33}$$

that is strictly decreasing in Q. The function (11.33) plays a fundamental role in the analysis of free entry equilibria when asymmetries are admitted. Indeed, it allows the mapping of the total equilibrium output – in general the aggregate A – into the profits of the individual firms,

[28] One can notice that both expressions come directly from the first-order conditions

$$\frac{\partial \Pi_i}{\partial q_i} = a - c_i - b(q_i + Q_{-i}) - bq_i = 0.$$

[29] Anderson et al. (2015) introduce a set of assumptions that guarantee the existence and uniqueness of an equilibrium in inclusive best replies. Moreover, under these assumptions the nature of interaction (strategic substitutability or complementarity) of the original best replies translates into an analogous feature of the inclusive best replies.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

where therefore Q replaces the number of firms as the key driver of equilibrium profits in an asymmetric setting.

Continuing with our Cournot example, a free entry equilibrium (FEE) can be defined as a set of quantities $\left\{(q_i^*)_{i \in I}\right\}$ and a set of entrants $I \subseteq I_m$, where I_m is the set of all m potential entrants, such that

$$\Pi_i(Q_I^*) \geq F_i \quad \text{for all } i \in I \tag{11.34}$$
$$\Pi_j(Q_I^* + q_j^*) < F_j \quad \text{for all } j \notin I$$

where $Q_I^* = \sum_{i \in I} \widetilde{q}_i(Q_I^*)$ is the aggregate output of the entrants I. Notice that we are not imposing symmetry in gross profits Π_i or in the sunk costs F_i. As a final step, it is often argued that the marginal entrant in a free entry equilibrium gains zero profit, a condition that is shared by all firms in a symmetric equilibrium. Anderson et al. (2015) assume that, among the potential entrants, there is a subset $e \subset I_m$ of *symmetric* marginal firms[30] with identical profit function $\Pi_i = \Pi_e(q_i, Q)$ and entry cost $F_i = F_e$ for all $i \in e$. Some of these marginal firms may be active, belonging to the set $e_a \subset I$.

In a zero-profit free entry equilibrium (ZPFEE) a nonempty set of marginal firms e_a is active and gains zero profit. More formally, a ZPFEE is an FEE with a set I of active firms such that $e_a \subset I$ and $\Pi_i = \Pi_i^*(Q_I^*) = F_i$ for all $i \in e_a$, where $\Pi_i^*(.)$ is given by (11.33). The existence of a fringe of symmetric active marginal entrants allows the combination of the zero-profit condition of the marginal firms with a unique level of aggregate output Q_I^* and with a variety of profit levels of the inframarginal (asymmetric) firms. Indeed, since $\Pi_i^*(Q)$ is decreasing in Q, from the zero-profit condition for the active marginal firms we obtain Q_I^*, and this latter determines the profits of the other inframarginal firms $\Pi_i^*(Q_I^*)$. The number of active marginal firms e_a is then adjusted through the entry process to find the ZPFEE.

To illustrate these properties it is interesting to analyze how the ZPFEE varies when exogenous changes in the set of inframarginal firms occur, modifying their profit structure and, consequently, the optimal output they deliver to the market. Let us consider an exogenous shock that affects a subset I_C of inframarginal firms (the changed firms), such as, for instance, a process innovation, or a merger, or a privatization, while leaving the other inframarginal firms in subset I_U (the unchanged firms) unaffected. Hence, in the initial state, $I = I_C \cup I_U \cup e_a$.

Then, after the shock the set of active firms in a ZPFEE moves from I to I'. All the changed and unchanged inframarginal firms remain active both before and after the shock, i.e. $I_C = I'_C$ and $I_U = I'_U$. The adjustment to the new ZPFEE works through a variation in the set of active marginal entrants: $e_a \neq e'_a$. Since $e'_a \neq \emptyset$ in the new equilibrium, $\Pi_i = \Pi_e^*(Q_{I'}^*) = F_e$ must hold for $i \in e'_a$ and therefore total output remains the same; that is, $Q_{I'}^* = Q_I^*$. Consequently, the profits of the unchanged inframarginal firms do not vary. Hence, for instance, a reduction in the marginal cost of the changed inframarginal firms I_C leads them to produce more in the new ZPFEE, whereas the unchanged inframarginal firms I_U maintain the same level of production. Since total output does not vary, the set of marginal firms shrinks as does their

[30] A possible justification of this key assumption rests on the following argument. The industry is populated by a set of larger firms that display rich strategies and, through them, are able to introduce some competitive advantage, i.e. asymmetry. Then, there is a fringe of small firms (the marginal entrants) that are not strategically sophisticated and adopt a standard and similar technology and are therefore less efficient than the larger ones.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

overall production, adjusting the larger production of the changed inframarginal firms and maintaining total output Q_I^* at the initial level.[31]

This property of the ZPFEE also encompasses the case of the "aggressive leaders" in Etro (2006), where one firm, the leader, is the inframarginal agent and the other symmetric firms, the followers, belong to the active marginal entrant group e_a. A change in the profits of the leader, for instance due to some investment, as Etro (2006, p. 150) writes, "does not affect the equilibrium strategies of the other firms, but it reduces their equilibrium number". Interestingly, in this setting with an endogenous number of followers, if the investment increases the marginal profit of the leader, this latter has an incentive to over-invest, no matter whether competition is in strategic complements or substitutes. Indeed, if the market equilibrium output does not change with its investment, whereas its market share and profits increase, the leader will over-invest. At the limit, if the investment is not costly, the leader has the incentive to produce more than the usual Stackelberg leader's output and to monopolize the market, preventing the entry of the followers.

This result of generalized over-investment is strikingly different from what happens when the number of followers (entrants) is given and exogenous. In the taxonomy proposed by Fudenberg and Tirole (1984), when the investment increases the marginal profit, the leader over-invests (top dog) if competition is in strategic substitutes but it under-invests (puppy dog) when it competes in strategic complements.

Aggregative games also greatly simplify the normative analysis of asymmetric environments. Starting with the case of homogeneous products, we observe that consumers' surplus depends on aggregate output only,[32] i.e. $CS = CS(Q)$, with $CS(0) = 0$, $CS'(.) > 0$ and $CS''(.) \leq 0$ for all $Q \geq 0$. Then, when a shock affects a subset of inframarginal firms while leaving total output Q_I^* unchanged, consumers' surplus also does not vary. The only impact on social welfare comes from the variation in profits of the changed inframarginal firms I_C. Indeed, the profits of the unchanged inframarginal firms I_U do not vary and the change in the number of active marginal firms from e_a to e_a' does not affect welfare, since they gain zero profits. We conclude that if a shock induces a profitable adjustment in a subset of firms and a change in their market shares, the only effect on welfare comes from the variation in the profits of the affected firms, quite in contrast with the impact in the short run when the number of firms does not vary.

To appreciate the result, let us consider the welfare impact of a merger between two firms absent any efficiency gain. The short-run effects are well known in the IO literature: the merged entity internalizes the negative externalities and contracts output; the outsiders react by expanding their production. The net effect is a fall in total output, consumers' surplus and total welfare, an increase in outsiders' profits and, in the case of constant returns to scale, a fall in insiders' profits.[33]

Once we consider entry and ZPFEE, however, the effects change significantly. Since additional active marginal firms enter in reaction to the short-run adjustments, total output,

[31] This neutrality outcome recalls a case of a competitive market where a fringe of identical firms with constant marginal costs makes the supply curve flat at some price p. Any efficiency improvement of the inframarginal firms affects the supply curve but the market equilibrium is always determined by $(p, D(p))$. The reduction in costs, then, is cashed in by the inframarginal firms as increased profits. These latter, in a sense, are Ricardian rents.

[32] This is true if firms' activities do not entail any externality, such as, for instance, different levels of pollution. If this were the case, the composition, and not only the total level of output would matter from a welfare point of view. In our discussion we are assuming that these composition effects do not arise in a homogeneous product market.

[33] See Salant, Switzer and Reynolds (1983).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

consumers' surplus and outsiders' profits (I_U and e_a) do not change. The insiders' (I_C) profits, due to their output contraction, are weakly lower. If, however, the merger allows the realization of efficiencies, insiders' profits, as well as their incentive to merge, increase, as does total welfare. This result brings with it a strong policy implication in favor of lifting ex ante merger control and authorization policies. Indeed, since the long-run private and social effects of a merger coincide, if private firms have an incentive to merge, then social welfare will rise, whereas socially damaging mergers would never be implemented given the lack of private incentives.[34]

In the welfare analysis of homogeneous product markets, we assumed that consumers' surplus depends only on total output but not on its allocation among the active firms. Moving to a differentiated products environment a similar assumption may be more problematic. Indeed, Anderson et al. (2015) show that in aggregative oligopoly games with differentiated products, a reallocation of a given aggregate among the different varieties, although neutral on the ZPFEE conditions, may affect total surplus and welfare. In other words, it may be that consumers' surplus not only depends on the aggregate, but also on its composition.

They show that the dependence of consumers' surplus on the aggregate only still persists with differentiated products if the demand functions satisfy the independence of irrelevant alternatives (IIA) property; that is, if the ratio of any two demands depends only on their own prices and not on the prices of other, unconsidered, alternatives. Notably, the logit model, as well as the demand functions derived from the CES utility function, satisfy the IIA and therefore the corresponding oligopoly game is not only aggregative, but also allows the expression of consumers' surplus as a function of the sum of the prices only.[35]

7 ENDOGENOUS SUNK COSTS AND PERSISTENT CONCENTRATION

The entry decision in the previous sections involved sinking a fixed set-up cost F that was related to some initial indivisible investment. We have not further specified the nature of these outlays. Assuming that the level of the sunk cost F is an exogenous parameter with respect to the entry and market strategies may be explained referring to technology (e.g. investment in a minimum efficient scale plant) or institutions (e.g. the payment of a license fee needed to operate). The sunk cost may vary, allowing us to extrapolate comparative statics properties, but for reasons orthogonal to the market strategies adopted by the active firms once entered. In this sense we can label the environments considered so far as characterized by *exogenous sunk costs.*

In this setting, the amplitude of the sunk costs F compared to the size of the market S is a fundamental driver in determining the maximum number of firms sustainable in a free entry equilibrium. The limiting case, when F becomes negligible with respect to S, leads

[34] Notice that the hands-off policy implications of free entry on merger control are much stronger than the usual argument that low entry barriers may constitute a favorable element when analyzing a merger. In this latter case easy entry conditions may mean that pros are balanced with the cons of enhanced market power in the evaluation of a merger. In the ZPFEE case, free entry is instead sufficient to generate mergers only when they are welfare enhancing.

[35] It should be stressed that aggregative product differentiation models do not necessarily satisfy the IIA, as is evident, for instance, considering the linear model drawn from Singh and Vives (1984). In this case consumers' surplus depends not only on the aggregate price but also on its composition.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

to convergence to a competitive equilibrium with an infinite number of firms, vanishing externalities and price converging to the marginal cost.

Although this paradigm can apply to several industries, there are many other sectors where a relevant part of the sunk costs arise due to specific market strategies of the firms, which in general we may connect to the effort of attaining a competitive advantage and market leadership. This is the case with investments in advertising that enhance the perceived quality of the product, or with R&D expenditures aimed at improving the efficiency of the technology or the quality of the products.[36] Similar effects take place in industries such as media and entertainment, where market leadership can be reached by securing specific, nonreproducible inputs such as, for instance, talent and premium content.[37] In all these examples, a competitive advantage is reached through enhanced efforts and, therefore, higher sunk costs. We label this second class of economic environments *endogenous sunk costs.*

When sunk costs react to market incentives, we may expect that the entry process, which is constrained by the need to repay all the sunk outlays, is affected. Indeed, market size, which drives the tendency to fragmentation in an exogenous sunk cost industry, has the additional effect of increasing the marginal return to market dominance, incentivizing leadership and endogenous sunk costs. A central result of the endogenous sunk cost case claims that if the incentives for effort are sufficiently high, an increase in market size does not lead to an increasingly fragmented market structure. There exists an upper bound to fragmentation such that, even at the limit, large firms and concentration persist.

We illustrate this result through a very simple model due to Schmalensee (1992)[38] that conveys the main ideas and intuition. In this setting we set the price $p > c$ fixed and concentrate on the investment in advertising A_i. The demand for product i has a similar structure to that in discrete choice models: $D_i(A_i, A_{-i}) = S * P_i(A_i, A_{-i})$ where S is market size and P_i firm i's market share. Moreover,

$$P_i(A_i, A_{-i}) = \frac{A_i^\gamma}{\sum\limits_{j=1}^{n} A_j^\gamma} \tag{11.35}$$

where $\gamma \in [0, 2]$ is a parameter that measures the mobility of consumers in reaction to advertising outlays. Notice that $\frac{\partial D_i}{\partial A_i} = \frac{\gamma}{A_i} P_i * (1 - P_i)$.

The profit function of firm i, then, is

$$\Pi_i(A_i, A_{-i}) = (p - c)S \frac{A_i^\gamma}{\sum\limits_{j=1}^{n} A_j^\gamma} - A_i - F \tag{11.36}$$

[36] A pathbreaking contribution in the theory and empirical analysis of these industries is due to Sutton (1991, 1998), the former referring to advertising-intensive industries and the latter to R&D-intensive sectors. See also Sutton (2007) for a comprehensive review.

[37] See on these examples Motta and Polo (1997, 2003).

[38] A full-fledged model based on quantity competition and investments in quality can be found in Sutton (1991, and 2007, Appendix B).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

where the last two terms refer to endogenous sunk costs in advertising (A_i) and exogenous sunk entry costs (F). In this setting there exists a symmetric Nash equilibrium in advertising levels

$$A^* = (p - c)S\gamma \frac{n-1}{n^2} \tag{11.37}$$

that is increasing in market size S and in consumers' reactivity to advertising γ.

Plugging into the profit function and taking into account that in a symmetric equilibrium $P_i = 1/n$, the zero-profit condition can be rewritten as:

$$\frac{1-\gamma}{n^*} + \frac{\gamma}{n^{*2}} - \frac{F}{S(p-c)} = 0, \tag{11.38}$$

where n^* is a solution of the above equation; that is, the SFEE number of firms.

The last term refers to exogenous sunk costs F and vanishes as the size of the market S increases indefinitely. However, the first two terms, which are directly related to the endogenous sunk costs in advertising outlays, present a different pattern: they do not depend on market size.[39]

When $\gamma \leq 1$, corresponding to consumers poorly reacting to advertising, and therefore a weak competitive pressure for market leadership, the single positive solution n^* of (11.38) increases indefinitely in market size S, reproducing a pattern we already observed in pure exogenous sunk cost models. However, for $\gamma \in (1, 2]$ the incentives to invest in market leadership bite and advertising increases in larger markets, pushing up the endogenous sunk costs. In this latter case

$$\lim_{S \to \infty} n^* = \frac{\gamma}{\gamma - 1}.$$

The entry process in this case is predominantly governed by the endogenous sunk costs, and the number of sustainable firms is bounded above for any market size, implying persistent concentration.[40] Moreover, the endogenous sunk costs tend to rise more quickly when consumers are more responsive to advertising, increasing concentration. Interestingly, in exogenous sunk costs environments more intense competition is associated with a lower n^* and a more concentrated market, although these features dilute and vanish when the market size increases indefinitely. This pattern of higher concentration when competition is harsher, instead persists in endogenous sunk cost industries even with growing market size.

[39] This feature, literally speaking, depends on the specific set-up of the very simple model we adopt. However, a general property of this class of models is that when market size increases indefinitely, gross profits and investment costs once we reach a certain number of firms tend to increase at the same rate. In this case, when S increases, boosting the gross profits, the incentives to invest in market leadership increase accordingly and the endogenous sunk costs increase at the same rate, preventing entry of additional firms.

[40] Shaked and Sutton (1983) identify a second case when the number of firms does not increase when market size rises. When firms offer different qualities $x_i \in [\underline{x}, \overline{x}]$ and the burden of quality improvements falls on fixed rather than marginal costs, price competition squeezes the margins. With relatively similar prices the demand for lower-quality products vanishes and a limited number of firms survives (finiteness property).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

8 FRICTIONLESS ENTRY AND CONTESTABILITY

The general result in the endogenous and exogenous sunk costs cases claims that there exists a maximum number of firms sustainable in a free entry equilibrium, and that it is decreasing in the amplitude of the sunk costs F compared with market size S. A concentrated market, in turn, is associated with noncompetitive mark-ups and allocative inefficiency. At the limit, when the economies of scale are particularly relevant, then we might find that only one firm can operate in the market: a case of natural monopoly. The firm will set the monopoly price p^m, being able to cover the high fixed costs with the monopoly margins. A second, symmetric entrant, pushing the market price down to $p(2) = P(Q(2))$ with its additional output, would make losses, since by definition in a natural monopoly it would be unable to cover the fixed costs. Then, there is a range of fixed costs such that the monopoly price is charged and no entry occurs. Similar cases can be generated where a small number of firms can be sustained in a free entry equilibrium.

The contestable markets approach[41] challenges this view, arguing that when entry is frictionless, structural monopoly or oligopoly environments do not lead to monopoly or oligopoly pricing and the associated allocative distortions. Indeed, potential competition may exert a sufficient corrective effect on the incumbent, inducing it to set a (second-best) efficient price to prevent temporary (hit-and-run) entry. Allocative efficiency is therefore ensured by (potential) competition even when economies of scale are so relevant to preventing actual competition.

This striking result re-establishes in a free entry environment a central feature of the Bertrand result, which claims that no relationship exists between the number $n > 1$ of active firms and the (socially efficient) oligopoly equilibrium. Indeed, as the exogenous sunk cost paradigm extends to the free entry case the Cournotian result of smooth convergence to competitive equilibria, the contestable market approach brings to the stage of the free entry story a Bertrand-type flavour.

It is now time to specify in more detail what we mean by frictionless entry. As a general point, the incumbent firm and the (potential) entrant are, under any respect, perfectly identical.

Since we are considering a natural monopoly, the first issue to address is the nature and amplitude of the fixed costs. Let us consider the following example. On the supply side, suppose that, in order to operate in the industry, it is necessary to bear a total investment F for an indivisible capital good that provides production services over a time horizon T. Let us divide this total time into t periods, whose length we are going to specify below. The incumbent firm I, then, has to cover a fraction $f = F/t$ of the fixed costs in each of the t periods it is active in the market, and has variable costs $C_I(q_I)$. Let us consider the case $f \in (\Pi_2, \Pi^m]$, where Π_2 are the gross profits from duopoly and Π^m the monopoly gross profits. Under this assumption the number of firms sustainable in the market is $n^* = 1$; that, is the market is a natural monopoly.

The potential entrant E, if it is willing to enter, has to pay $F = t * f$ to purchase the capital good. If, after one period, E decides to exit, the residual value of the capital good is $(t - 1) * f$.

[41] See Baumol, Panzar and Willig (1982). To ease the exposition we present here the case of a contestable natural monopoly. The authors generalize the contestable market approach to natural monopolies, showing that second-best efficient allocations arise also in these cases when entry is frictionless. The case of multi-product firms and economies of scope is a third, relevant extension of the analysis.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Let $\alpha \in [0,1]$ be the fraction of the residual value that can be cashed back by reselling the capital good or by using it in other markets. This parameter measures the sunkness of the initial investment, with $\alpha = 1$ corresponding to the case when the capital good can be efficiently recovered after exit and $\alpha < 1$ to some level of sunkness. If E enters and produces, its costs are $C_E(q_E)$. It is evident that, since the incumbent can efficiently use the capital good in the market for the entire length of its economic life, the entrant is in a symmetric position on the supply side only if $\alpha = 1$ and $C_E(q) = C_I(q)$.

Turning to demand, for a given price p the entrant's demand is $D_E(p) \leq D_I(p)$ where the equals sign corresponds to a symmetric position towards the customers, who are uncommitted and can switch to the entrant if the price p_E is more attractive than the incumbent's price p_I.

The timing of the game is as follows: at $s = 0$ the incumbent sets a price p_I *that cannot be changed* for a period of length T/t; just after p_I is set the entrant posts its own price p_E; once the two prices are set, the customers choose which of the two firms to patronize and are supplied immediately; at $s = T/t$, before the incumbent changes its price, the entrant exits and resells (or reuses) the capital good, collecting $\alpha(t-1)f$.

Once the contestable market story is unbundled, two key ingredients become evident:

1. There is no administrative restriction on entry, as licenses or authorizations.
2. Demand and supply quantities adjust instantaneously while price changes take time.

In this environment, the incumbent sets a (limit) price that prevents the temporary entry of the competitor:

$$\widehat{p}_I = \frac{C_E\left(D_E(\widehat{p}_I)\right) + f\left[\alpha + t(1-\alpha)\right]}{D_E(\widehat{p}_I)}. \tag{11.39}$$

If we compare (11.39) with the second-best Ramsey price

$$p^{sb} = \frac{C\left(D(p^{sb})\right) + f}{D(p^{sb})}$$

we can immediately notice that the limit price set by the incumbent is second-best efficient if three further conditions hold:

3. The entrant has access to the same technology as the incumbent, with no restrictions coming from patents or privately owned know-how: $C_E(q) = C_I(q)$; moreover, it can instantaneously change the level of production at the desired level.
4. The customers see the entrant and the incumbent as offering perfect substitutes and have no restrictions or costs in switching from one to the other: $D_E(p) = D_I(p)$.
5. The fixed indivisible investment is not sunk and the entrant recovers the residual value of the capital good entirely: $\alpha = 1$.

Under assumptions 1–5 potential competition is able to discipline the incumbent and induces second-best efficient outcomes in markets plagued by substantial economies of scale and concentration. Intuitively, perfect symmetry of the incumbent and the entrant and frictionless entry allow the market to be supplied, indifferently, by either of the two firms. If the incumbent commits to a profitable price, it is temporarily replaced by the entrant through undercutting. In this case, the identity of the provider changes for a period, although the

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

market remains a monopoly. To avoid undercutting, the incumbent is forced to adopt the efficient limit price equal to the average costs. This remarkable result is derived under a set of specific assumptions, and can be evaluated both with respect to their empirical relevance and theoretical robustness. On theoretical grounds, the limit price expression (11.39) clearly shows that substantial departures from the second-best efficient price occur when any of the assumptions are weakened.

Turning to empirical relevance, the contestable market approach inspired the liberalization of the airline industry in the USA in the late 1970s.[42] In this sector a market corresponds to a route, and therefore the large investments in aircrafts are not specific to a market: the aircrafts can be moved to other routes or resold in an efficient market. Alternatively, the carriers can lease the aircrafts. The other fixed costs, check-in and handling services, are specific to airports, and therefore to the routes served. In the market reform, the airports, rather than the carriers, supplied these services, leasing them to the carriers on a variable cost basis. Hence, Assumption 5 of no sunkness seems consistent with the empirical data, as well as the access to the same technology (Assumption 3). Price stickiness may derive from contractual constraints on fares posted in advance (Assumption 2), and lifting authorizations was a key measure of the reform (Assumption 1). However, Assumption 4 was the Achilles' heel of the reform, since slots were assigned under grandfather rights, and the peak-hour more profitable ones remained in the portfolio of the incumbents. Moreover, in the years after the reform the carriers reorganized the routes from a spoke-to-spoke to a hub-and-spoke pattern, enhancing their dominant role in large hubs and achieving high load factors. With $D_E(p) < D_I(p)$, after an initial phase of turbulence, the incumbents were able to profitably prevent entries and maintain dominance in their key hubs.

Hence, although intellectually brilliant, the contestable market approach can hardly be considered a general theory of free market equilibria due to its lack of robustness. Although potential competition is an important ingredient in entry games, its impact on the behavior of active firms has to be carefully evaluated from an empirical point of view.

9 CONCLUSIONS

In this chapter we have reviewed the different branches of the IO literature that analyzes free entry equilibria and the endogenous determination of market structure. A recurrent theme is the assumption of symmetric firms, which in a long-run perspective can be justified when the friction of access to technology and the features of demand allow all firms to refer to a common set of best practice techniques and to exploit the possibility of (horizontal) product differentiation. In this perspective, a very rich class of oligopoly models is characterized by significantly similar comparative statics properties of the market prices, quantities and profits when the number of active firms increases. Two limiting cases emerge, perfectly competitive and the monopolistic competitive outcomes, when the number of firms increases indefinitely. The monotone comparative statics tools allow the identification of the general conditions behind these results. Long-run market structures under free entry are determined by a small set of elements referring to technology (economies of scale) and preferences (market size), with an additional ingredient related to strategies and the intensity of price competition.

[42] See Bailey and Panzar (1981) and Fawcett and Farris (1989).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Hence, the general result of free entry equilibria provides a solid theoretical foundation to the traditional approach of industrial economics based on the structure–conduct–performance paradigm.

The normative properties of free entry equilibria show that in a homogeneous product setting the business-stealing effect is the key element that creates a wedge between the private incentives and the social planner, determining an excessive number of firms. When product differentiation is introduced, however, an opposite externality leading to underprovision of varieties is also introduced, since the private incentives to enter do not include the benefits of an increased number of substitute products on consumers.

While symmetric market games are a useful reference for the long-run evolution of markets, asymmetric settings may be relevant both in the long run, when frictions persist, and as a starting point from which to study the evolution of market structure under free entry. It is important to notice that some form of symmetry is also maintained in this framework, which exploits the aggregative nature of many oligopoly models, by assuming that the (relatively inefficient) marginal entrants are all alike. The zero-profit condition on the marginal entrants, together with the aggregative nature of the market games, then generates unconventional long-run effects when a shock hits the active firms. Indeed, in the new free entry equilibria the total output remains unchanged, while its composition varies, with the change in output of the firms affected by the shock absorbed by an opposite variation in the number of marginal entrants. With these results, a hands-off policy is implied.

Endogenous sunk costs related to market strategies provide a different pattern of adjustment characterized by persistent concentration even in very large markets, in contrast with the tendency to fragmentation when sunk costs are exogenous. Finally, we review the attempt to establish efficient entry equilibria even in markets characterized by huge economies of scale and structural concentration, including natural monopolies, by assuming frictionless entry and giving a role to potential competition. The contestable markets paradigm refreshes the features of Bertrand competition in a free entry set-up, in contrast with the Cournotian paradigm of the exogenous sunk costs approach. Once again, symmetry plays a role, since the effectiveness of potential competition in disciplining dominant firms rests on the assumption that the entrants can perfectly replace the incumbent during their temporary raid in the market.

REFERENCES

Amir R. (2003), Supermodularity and Complementarity in Economics: An Elementary Survey, mimeo.

Amir R. and Lambson V. (2000), On the Effects of Entry in Cournot Markets, *Review of Economic Studies*, 67: 235–54.

Amir R., De Castro L. and Koutsougeras L. (2014), Free Entry versus Socially Optimal Entry, *Journal of Economic Theory*, 154: 112–25.

Anderson S., De Palma A. and Thisse J. (1992), *Discrete Choice Theory of Product Differentiation*, Cambridge, MA: MIT Press.

Anderson S., Erkal N. and Piccinin D. (2015), Aggregative Oligopoly Games with Entry, mimeo.

Bailey E. and Panzar J. (1981), The Contestability of Airline Markets During the Transition to Deregulation, *Law and Contemporary Problems*, 49: 125–45.

Bain J.S. (1956), *Barriers to New Competition*, Cambridge, MA: Harvard University Press.

Baumol W., Panzar J. and Willig R. (1982), *Contestable Markets and the Theory of Industry Structure*, San Diego, CA: Harcourt Brace Jovanovich.

Bulow J., Geanakoplos J. and Klemperer P. (1985), Multimarket Oligopoly: Strategic Substitutes and Complements, *Journal of Political Economy*, 93: 488–511.

Cournot A. (1838), *Recherches sur les Principes Mathématiques de la Théorie des Richesses*, Paris: Chez L. Hachette.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

d'Aspremont C., Gabszewicz J. and Thisse J. (1979), On Hotelling "Stability in Competition", *Econometrica*, 47: 1145–50.
Dixit A. and Stiglitz J. (1977), Monopolistic Competition and Optimum Product Diversity, *American Economic Review*, 67: 297–308.
Etro F. (2006), Aggressive Leaders, *Rand Journal of Economics*, 37: 146–54.
Fawcett S. and Farris M. (1989), Contestable Markets and Airline Adaptability under Deregulation, *Transportation Journal*, 29: 12–24.
Friedman J. (1971), A Non-cooperative Equilibrium for Supergames, *Review of Economic Studies* 38 (1): 1–12.
Fudenberg D. and Maskin E. (1986), The Folk Theorem in Repeated Games with Discounting or with Incomplete Information, *Econometrica* 54 (3): 533–54.
Fudenberg D. and Tirole J. (1984), The Fat Cat Effect, the Puppy Dog Ploy and the Lean and Hungry Look, *American Economic Review*, 74: 361–66.
Häckner J. (2000), A Note on Price and Quantity Competition in Differentiated Oligopolies, *Journal of Economic Theory*, 93: 233–39.
Hotelling H. (1929), Stability in Competition, *Economic Journal*, 39: 41–57.
MacManus M. (1962), Number and Size in Cournot Equilibrium, *Yorkshire Bulletin of Economic and Social Research,* 14: 14–22.
MacManus M. (1964), Equilibrium, Number and Size in a Cournot Oligopoly, *Yorkshire Bulletin of Economic and Social Research,* 16: 68–75.
Mankiw G. and Whinston M. (1986), Free Entry and Social Efficiency, *Rand Journal of Economics*, 17: 448–58.
Manski C. (1977), The Structure of Random Utility Models, *Theory and Decisions*, 8: 229–54.
Milgrom P. and Roberts J. (1990), The Economics of Modern Manufacturing: Technology, Strategy and Organization, *American Economic Review*, 80: 511–28.
Milgrom P. and Roberts J. (1994), Comparing Equilibria, *American Economic Review,* 84: 441–59.
Milgrom P. and Shannon C. (1994), Monotone Comparative Statics, *Econometrica,* 62: 157–80.
Motta M. and Polo M. (1997), Concentration and Public Policy in the Broadcasting Industry, *Economic Policy*, 25: 295–334.
Motta M. and Polo M. (2003), Beyond the Spectrum Constraint: Concentration and Entry in the Broadcasting Industry, in Baldassarri M. and Lambertini L. (eds.), *Antitrust, Regulation and Competition*, Basingstoke, UK: Palgrave Macmillan.
Novshek W. (1985), On the Existence of Cournot Equilibrium, *Review of Economic Studies,* 52: 85–98.
Perry M. (1984), Scale Economies, Imperfect Competition and Public Policy, *Journal of Industrial Economics*, 32: 313–30.
Quandt R. (1956), A Probabilistic Theory of Consumer Behavior, *Quarterly Journal of Economics*, 70: 507–36.
Roberts J. and Sonnenschein H. (1976), On the Existence of Cournot Equilibrium Without Concave Profit Functions, *Journal of Economic Theory*, 13: 112–17.
Salant S., Switzer S. and Reynolds R. (1983), Losses from Horizontal Mergers: The Effects of an Exogenous Change in Industry Structure on Cournot-Nash Equilibrium, *Quarterly Journal of Economics*, 98: 185–99.
Salop S. (1979), Monopolistic Competition with Outside Good, *Bell Journal of Economics*, 10: 141–56.
Scherer F.M. (1980), *Industrial Market Structure and Economic Performance*, Chicago, IL: Rand McNally.
Schmalensee R. (1989), Inter-industry Differences of Structure and Performance, in Schmalensee R. and Willig R. (eds.), *Handbook of Industrial Organization, Vol. 2*, 951–1009, Amsterdam: North Holland.
Schmalensee R. (1992), Sunk Costs and Market Structure: A Review Article, *Journal of Industrial Economics*, 40: 125–34.
Selten R. (1970), *Preispolitik der Mehrproduktunternehmung in der statischen Theorie*, Berlin: Springer-Verlag.
Shaked A. and Sutton J. (1983), Natural Oligopolies, *Econometrica*, 51: 1469–83.
Shubik M. and Levitan R. (1980), *Market Structure and Behavior*, Cambridge, MA: Harvard University Press.
Singh N. and Vives X. (1984), Price and Quantity Competition in a Differentiated Duopoly, *Rand Journal of Economics*, 15: 546–54.
Spence M. (1976), Product Selection, Fixed Costs and Monopolistic Competition, *Review of Economic Studies*, 43: 217–35.
Sutton J. (1991), *Sunk Costs and Market Structure*, Cambridge, MA: MIT Press.
Sutton J. (1998), *Technology and Market Structure*, Cambridge, MA: MIT Press.
Sutton J. (2007), Market Structure: Theory and Evidence, in Amstrong M. and Porter R. (eds.), *Handbook of Industrial Organization, Vol. 3*, Amsterdam: North Holland, pp. 2303–64.
Vives X. (1990), Nash Equilibrium with Strategic Complementarities, *Journal of Mathematical Economics,* 19: 305–21.
Vives X. (1999), *Oligopoly Pricing*, Cambridge, MA: MIT Press.
Von Weiszäcker C. (1980), A Welfare Analysis of Barriers to Entry, *Bell Journal of Economics*, 11: 399–420.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

12. Evolutionary oligopoly games with heterogeneous adaptive players

Gian Italo Bischi, Fabio Lamantia and Davide Radi

1 INTRODUCTION

Since 1838, the year of publication of the seminal duopoly model (Cournot, 1938), Cournot oligopoly games have become one of the standard and most widely employed models of imperfect competition. Although this contribution had been proposed long before game-theoretic methods were formally developed, it has become central in the literature on industrial organization and, in general, on game theory. The equilibrium concept proposed by Cournot is regarded as the first example of Nash equilibrium, a concept elaborated by Nash in 1950 for general N-person games.[1]

According to its definition, when all agents are in such an equilibrium, no unilateral deviation by a single agent is profitable and, therefore, no competitor has an incentive to depart from that state. However, the possibility of attaining a Nash equilibrium, even in a theoretical setting, is not a weak outcome, as its immediate realization involves the postulation that every player is endowed with the highest degree of rationality and information. In oligopoly games, this means that each firm possesses cognitive and computational skills to understand exactly the microeconomic structure of the game: the market in which it operates (i.e. the exact specification of the demand for the produced goods) and the production technologies adopted by the various firms (represented by their cost functions); moreover, it must be assumed that firms have rational expectations, which, in a deterministic setting, means perfect foresight into the competitors' future production. Such assumptions are often considered too demanding when compared with real-world cases, where firms usually have partial knowledge of the complex and uncertain environment in which they operate, and are not endowed with such high levels of rationality, information and computational skills.

So, in the economic literature several heuristics, characterized by bounded rationality and/or limited information, have been proposed as a proxy for more realistic "behavioral rules" followed by firms to make their production decisions. This point of view is consistent with recent trends in economic research, where a shift in the modeling has taken place from the standard paradigm of the rational and representative agent (who is endowed with unlimited computational ability and perfect information) to alternative perspectives, which allow for assumptions such as bounded rationality, agents' heterogeneity, social interaction and learning. As a result, agents' behavior is governed by simpler "rules of thumb", "trial and error" or even "imitation" mechanisms. Under such weaker and more realistic assumptions, firms are unlikely to simultaneously select a Nash equilibrium output level; consequently, some firms might be able to increase their profits by changing their output level in the next

[1] In the context of oligopoly models where firms compete in output, it is sometimes called Cournot–Nash equilibrium. However, we will follow the convention of most game theorists and simply call it Nash equilibrium.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

time period. This implies that, under such behavioral rules, firms will change their outputs step by step. In other words, such heuristics often mean that decisions are repeated over time. This approach may seem, at first sight, a quite unsatisfactory and dismissive (in the sense of understating and reductive) representation of how economic agents face decisions. However, the significance of this perspective becomes interesting and meaningful when decisions are repeatedly made over time. In some cases, the repetition of boundedly rational decisions, which we denote by the general term "adaptive", may represent a much more realistic (and even more efficient) behavior than a rigid optimizing attitude, which follows by postulating rational choices. Indeed, the latter approach may become quite unreliable (even misleading) under incomplete information on the environment where the economic agents operate or on other agents' degree of rationality, or under intrinsic uncertainty on the evolution of the system. The latter occurs, for instance, when the dynamics of the system are governed by nonlinear laws that allow for chaotic behavior, with the associated phenomenon of sensitivity to arbitrarily small perturbations, which is a quite common occurrence in economics and social sciences. Instead, adaptive agents are allowed to adjust their repeated actions on the basis of the information collected as the system evolves, so that they can adapt to events or even "learn" by repeatedly comparing expected and observed results of their decisions. As a consequence, the assumption on firms' dynamic adjustments may give rise to long-run evolutions of their productions that are more complex than simple convergence to a Nash equilibrium; furthermore, different heuristics may entail different degrees of stability as well as different long-run dynamic scenarios. This naturally leads to the question of whether and how such firms will coordinate on playing a Nash equilibrium, a question that has not been unambiguously resolved yet. More importantly, what kind of long-run pattern prevails when convergence to a Nash equilibrium does not occur?

Clearly, given an oligopoly game, a unique way of being rational exists. However, when the rationality assumption is abandoned, several different heuristics for mimicking bounded rationality can be postulated. Indeed, during the years, many adaptive behavioral rules have been proposed in the specialized literature; extensive overviews and references are provided in Bischi et al. (2010), to which we refer the interested reader. From an historical point of view, the first model aimed at describing the evolution of production plans with boundedly rational firms was based on best reply (BR) with naive expectations and was introduced by Cournot himself in 1938. This adaptive behavior has been considered and employed by several authors under different assumptions on the microeconomic structure of the game. Among the most relevant contributions, it is worth remembering Theocharis (1960), Fisher (1961), Rand (1978), Puu (1991) and Kopel (1996). Another adaptive behavioral rule, which will be extensively considered in this chapter, is the so-called local monopolistic approximation (LMA), proposed for the first time by Silvestre in 1977 with the name of "strong monopolistic equilibrium". Then, it has been employed more recently in Tuinstra (2004), Bischi, Naimzada, and Sbragia, (2007) and Cavalli, Naimzada, and Tramontana (2015), see also Naimzada and Sbragia (2006). Under the LMA heuristic, the oligopolists do not know the form of the market demand function and conjecture that it is linear; in addition, they assume that competitors' productions do not influence the expected price. Thus, firms proceed by estimating locally the slope of the demand curve and the current market state in terms of total output and price. Interestingly, Bischi et al. (2007) show that oligopoly dynamics with LMA behavior converge to a Nash equilibrium even in cases where the classic best reply dynamics fail to converge to it. Another heuristic considered in the following is gradient dynamics (GD); see e.g. Kenneth and Hurwicz

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

(1960), Varian (1992), Corchon and Mas-Colell (1996), and Bischi and Naimzada (1999). In gradient-like adjustments, players are not required to have complete knowledge of the demand and cost functions, nor do they choose their strategy by solving an optimization problem; indeed, they just employ a local estimate of their marginal profit. Each firm's output level is then updated by following the direction of increment of its profit, which is indeed regulated by the gradient of the profit. The reactivity of this adjustment is governed by a parameter, which can be set in different ways, as specified in the quoted literature.

The main feature that unites GD and LMA heuristics is that firms do not need "global" information on the demand function in order to adjust their production; however, they need some "local" information about it. Thus, one can assume that the adoption of either one of these two heuristics requires a lower amount of information than, for example, that required with best reply adjustments, see, e.g. Droste, Hommes, and Tuinstra (2002), Hommes, Ochea, and Tuinstra (2011) and Bischi, Lamantia, and Radi (2015), where different behavioral rules are compared with one rule employing global knowledge of the demand function. It is useful to observe that all these dynamic adjustment mechanisms share the property that any Nash equilibrium is a fixed point of the corresponding dynamical system. Thus, although these heuristics are driven by local (or myopic) decision rules from boundedly rational and heterogeneous agents, they may lead the system to long-run convergence to the same equilibrium forecasted (and reached in one shot) under the assumptions of full rationality and complete information of all economic agents.

In this chapter, after a brief review of the related literature, we describe a general framework for dealing with evolutionary oligopoly models with different behavioral rules. Here we show some general properties of such models by considering a switching mechanism based on *replicator dynamics*, the most common and probably simplest model proposed in the literature on evolutionary games to mimic selection pressure in favor of groups obtaining the highest payoffs (see Taylor and Jonker, 1978 and also Vega-Redondo, 1996, Hofbauer and Sigmund, 1998, and Weibull, 1997). Following Bischi, Lamantia, and Radi (2015) and Cerboni Baiardi, Lamantia, and Radi (2015), we consider an improvement of the standard replicator dynamics in discrete time that was first proposed in Cabrales and Sobel (1992) (see also Hofbauer and Sigmund, 2003 and Hofbauer and Weibull, 1996) and takes a monotone transformation of a discounted average of past profits as a fitness measure.[2]

Then we apply this general evolutionary framework to some examples, in which the population of firms can choose between two different heuristics, so that the long-run spread of either behavioral rule in the firms' population can be seen as a proxy for the long-run profitability of the two heuristics. When only two heuristics are considered, the population of firms can be split at each time period into two groups, according to the different rules they employ. Then, a switching mechanism based on last-period profits can be modeled by a three-dimensional dynamical system,[3] characterized by invariant planes where all players follow the same "pure" behavioral rule. The dynamics on these invariant planes are

[2] Although many papers are based on the replicator dynamics, this is not the only evolutionary selection mechanism, as many others have been proposed in the literature – see e.g. Brock and Hommes, 1997, 1998 and Droste et al., 2002 – for many different switching mechanisms and their applications.

[3] Here we consider only models where last-period profits influence firms' decisions, i.e. we only describe models without memory. By augmenting the dimension of the system, it is possible to deal with these extensions, as shown in Bischi, Lamantia, and Radi (2015). For discussions on this point, we refer the interested reader also to Hommes et al. (2012).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

governed by two-dimensional restrictions of the model that assume the form of triangular maps, i.e. unidirectionally coupled dynamical systems or systems with a skew-product structure (see Bischi, Lamantia, and Radi, 2015 and Cerboni Baiardi et al., 2015). Thus, many properties of the attractors that characterize the long-run evolution in the case of homogeneous firms (all firms following the same heuristic) can often be separately studied even by analytical methods. For the model with heterogeneous firms, several kinds of long-run evolutions can be detected, characterized either by the coexistence in the long run of both the considered heuristics, i.e. cyclic or chaotic attractors involving both kinds of behavioral rules, or possibly by on–off intermittency phenomena between monomorphic and polymorphic states in the oligopoly market, with either all firms following the same behavioral rule or not. In other words, interesting dynamic scenarios are observed when asymptotic dynamics do not converge to a Nash equilibrium. Moreover, we detected coexistence of different attracting sets that characterize the possible long-run distribution of the behavioral rules, so that the initial conditions of the game do play a crucial role for the resulting long-run behaviors.

In any case, if the Nash equilibrium of the game is stable under any considered behavioral rule, then the asymptotic behavior of the model is very simple. In general, the cheapest (in terms of information costs) behavioral rule will prevail and all agents will end up producing the Nash equilibrium quantity. In this case, the standard results from oligopoly theory are retrieved by the evolutionary model. However, intriguing questions – from an economic as well as mathematical point of view – arise when the Nash equilibrium is unstable at least under one behavioral rule. In these cases, the principal points of investigation are the following:

- What is the most likely short-run and asymptotic production of the industry when firms employ different behavioral rules and fail to converge to a Nash equilibrium?
- What is the long-term distribution of the behavioral rules in the population of firms? Can behavioral heterogeneity arise endogenously as a result of interaction among identical firms switching between different heuristics?
- What is the effect of different information costs of behavioral rules on the asymptotic dynamics of the system (or, equivalently, of the presence of a bias in the population towards a particular behavioral rule)?

In this chapter, we briefly address these questions for two important examples, namely the case of competition between LMA and BR rules, as in Bischi, Lamantia, and Radi (2015), and the case of competition between GD and LMA rules, as in Cerboni Baiardi et al. (2015).

The plan of the chapter is the following. In Section 2 we briefly survey some relevant contributions on dynamic oligopoly games, with particular reference to the literature on evolutionary models. Then in Section 3 a general evolutionary model is described together with its properties. In Section 4 an oligopolistic market characterized by isoelastic demand is considered; in subsection 4.1 a particular couple of heuristics is considered, namely BR and LMA, whereas subsection 4.2 focuses on the comparison between LMA and GD heuristics. Section 5 concludes.

2 RELATED LITERATURE

As we remarked before, different dynamic processes exist that may converge or not to the same Nash equilibrium. This occurrence opened up the road to the modeling of oligopoly

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

games with heterogenous firms, which decide their outputs over time by adopting different behavioral rules. Several works deal with oligopoly models with two or more firms following heterogeneous behavioral rules; see, just to cite a few, Leonard and Nishimura (1999), Den Haan (2001), Agiza and Elsadany (2003, 2004), Angelini, Dieci, and Nardini (2009), Tramontana (2010), Dubiel-Teleszynski (2011), Cavalli, Naimzada, and Tramontana (2015), Anufriev, Kopányi, and Tuinstra (2013), and Bischi et al. (2007). These papers consider duopoly and triopoly models with firms adopting different kinds of adaptive adjustments, involving different degrees of rationality and information, whence different costs. For these models, one important research question concerns the effect of such heterogeneities on the stability of Nash equilibria in the space of parameters, compared with the stability of the same equilibria when all firms adopt the same behavioral rule.

Recently, the more general case with a population of N firms has been considered, where agents are subdivided into complementary fractions (or classes) of adopters of the same heuristic. Here, one research question regards the stability of the Nash equilibria as the number N of firms increases, a classical problem in the literature on oligopoly games that dates back to the works of Theocharis (1960), Fisher (1961), Hahn (1962) and McManus and Quandt (1961), see, in particular, Hommes et al. (2011) on this point. Another important issue for these models is to investigate whether a Nash equilibrium is stable under a given subdivision into fractions of adopters of the various behavioral rules. This question can be ascertained by taking the shares of adopters of the different behavioral rules as bifurcation parameters, see e.g. Bischi, Lamantia, and Radi (2015), Cerboni Baiardi et al. (2015) and Cavalli, Naimzada, and Pireddu (2015). Indeed, if "better" rules are available for the firms, it seems reasonable that their management would adopt these rules to improve the overall "performance" of the firm. This argument suggests endogenizing the dynamics of the fraction of firms playing the different behavioral rules, according to a switching mechanism governed by profit-driven evolutionary pressure. Following the spirit of evolutionary games, such a switching mechanism determines at each time step how the shares of adopters of the different strategies are updated. These models are based on the distinct principle that the fraction of agents playing a better-performing strategy will increase in the next period at the expense of the fractions of worse-performing strategies. One way to model this in discrete time is to assume that at the end of each period each agent compares her payoff with the average payoff of the population of agents; then she decides to change her strategy if she assesses that adopting a different rule might provide her with higher average gains. An interpretation of this mechanism is that each agent observes the (current period) performance, for instance the payoff, of a randomly chosen agent among those that have used a different behavioral rule, and then she decides to imitate the sampled agent if the latter's gain is higher.

The general approach of evolutionary pressure on the basis of past firms' performance has been extensively employed for understanding the dynamic choices of agents with differential information, different expectation formation mechanisms or different objective functions. Several authors have studied models where agents are endowed with different kinds of expectations, such as rational expectations vs naive or adaptive expectations, with an index of performance defined in terms of discrepancy between expected and realized values, or profits gained (see e.g. Brock and Hommes, 1997, 1998, and Hommes, 2013), or market competition (see e.g. Chiarella and He, 2001 and Chiarella, Dieci, and Gardini, 2002 for applications to financial markets with heterogeneous agents, such as chartists and fundamentalists, adopting different strategies in forecasting price trends). Droste et al. (2002) provide an interesting

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

example with a Cournot duopoly where ex ante identical firms can employ different behavioral rules to set the quantities to produce. The economic structure of the underlying game is particularly simple, with homogeneous goods, linear demand and quadratic production costs. The fitness of each behavioral rule at each time period is assessed by considering average payoffs obtained by pairs of firms that are randomly matched to play the game. An evolutionary mechanism based on average profits regulates the distribution of the various rules over time. Droste et al. concentrate on the comparison between a "best reply" rule and a (costly, since sophisticated) "Nash" rule and show that endogenous fluctuations and complicated dynamics may arise, mainly due to the dominance of best reply behavior in a neighborhood of the Nash equilibrium because of information costs. Also Hommes et al. (2011) present a similar setup with linear demand and linear production costs, where firms switch between different expectation rules concerning aggregate output of their rivals but with random matching of N firms at a time. On the basis of past performance, these firms decide to switch between costly rational and cheap boundedly rational expectations on aggregate output of their competitors. Hommes et al. (2011) find that the classic Theocharis result in 1960 on the instability of the Nash equilibrium as the number of firms increases is also confirmed qualitatively under evolutionary competition between heterogeneous (costly and costless) expectation heuristics. Recent literature closely related to these works includes Bischi, Lamantia, and Radi (2015) and Cerboni Baiardi et al. (2015), which are extensively discussed in the second part of this chapter.

Other contributions deal with evolutionary selection to explain whether behavioral heterogeneity, such as the evolution of preferences, can arise as a result of this dynamic process. One of the first papers in the oligopoly literature dealing with the issue is Schaffer (1989), which shows that profit-maximizing firms are not necessarily the best survivors provided that firms have market power. Similarly, Heifetz, Shannon, and Spiegel (2007) establish that evolutionary motives do not always justify payoff maximization under strategic interaction among the players. In fact, the population will not converge to payoff-maximizing behavior under any payoff-monotone selection dynamics. Relatedly, several authors have studied the emergence of so-called "*spiteful behavior*" in evolutionary oligopolies to explain evolutionary dominance of Walrasian behavior over Cournotian behavior, as the former allows the reduction of losses with respect to other strategies see, for example, Vega-Redondo (1997), Vriend (2000), and Vallée and Yildizoglu (2009). In this vein, Radi (2017) analyzes an oligopoly game with a population of firms subdivided into a fraction following the best reply rule with naive expectations and complementary fraction following a Walrasian rule, i.e. each firm acts as naïve price taker, so that the resulting dynamics have a Walrasian equilibrium output as the unique steady state. Rhode and Stegeman (2001) propose a differentiated duopoly to argue that if firms' choices follow an evolutionary process, then the long-run outcome is not a Nash equilibrium and evolutionary pressure tweaks the objective towards revenue maximization. Adopting an indirect evolutionary approach, Königstein and Müller (2001) ascertain that it is worthwhile for firms to include a share of consumer welfare in their objective function. Kopel, Lamantia, and Szidarovsky (2014) consider a similar evolutionary mechanism to describe an oligopoly market where a fraction of firms just behave as profit maximizers and the complementary fraction exhibit social responsibility by maximizing the sum of their profits plus a share of the consumer surplus. De Giovanni and Lamantia (2016a, 2016b) study several versions of an evolutionary oligopoly with control delegation to managers and different amounts of

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

information to the firms about how managers are compensated under linear and nonlinear market demand. Other generalizations are given by multipopulation models, e.g. models with m different populations of firms (e.g. m nations or m industrial districts) with different numbers of individuals, each with a given set of strategies (or behavioral rules) available. For example, Bischi, Dawid, and Kopel (2003a, 2003b) deal with two-population models (e.g. two industrial districts), where agents of each population have two available strategies (invest in the industry or in financial markets), with a form of "switching by imitation of the more profitable strategy". They investigate the global dynamics of the system with intra-group and inter-group spillovers and demonstrate the kind of complex dynamics that may occur through interaction. Other examples of switching by imitation can be found in Hofbauer and Sigmund (2003), Weibull (1997) and Bischi, Lamantia, and Sbragia (2009b).

Interesting extensions of these oligopoly models are provided within the literature on evolutionary resource exploitation models. In environmental economics, evolutionary models have been employed to study cooperative vs noncooperative behaviors in the exploitation of a natural resource (e.g. a fishery; see Bischi, Lamantia, and Sbragia, 2009a), the establishment of protected areas (see Bischi et al., 2009b and also Bischi and Lamantia, 2007 on this point) and related issues. Bischi, Lamantia, and Radi (2013a, 2013b) analyze evolutionary oligopoly models where the behavioral rule to adopt consists of deciding the species to harvest; in addition, several different multispecies interactions are assumed in these papers. Lamantia and Radi (2015) describe a fishery model where two different harvesting technologies can be employed: a standard one and an environmentally friendly one (less intensive). Thus, fishermen may decide to employ a less efficient but more "environmentally friendly" fishing technology if the loss in efficiency is counterbalanced by a higher price that consumers might be willing to pay for the "green" product. In this model, the harvesting technologies to adopt constitute the behavioral rules at disposal of the oligopolists. The cases of continuous time and the hybrid system (continuous time for the resource and discrete time for decisions on technology switching) are analyzed in Lamantia and Radi (2015), whereas Bischi, Cerboni Baiardi, and Radi (2015) study the same model with discrete-time scale.

Next, we specify a general evolutionary setting in which we define and compare the dynamic properties of two specific pairs of behavioral rules, namely BR vs LMA and LMA vs GD.

3 THE GENERAL MODEL

Let us consider an oligopoly market with N ex ante identical firms that produce homogeneous goods. We assume that the set of strategies (output levels) is a nonempty compact and convex set of $\mathbb{R}^N$ and each firm's profit is concave in its own strategy.[4] These assumptions guarantee that a Nash equilibrium exists; see Rosen (1965). Firms can conceive different behavioral rules for setting their next-period productions. For the sake of the argument, let us consider that only two different behavioral rules are available and denote by $x_i(t)$ the production at

[4] In particular, the case with isoelastic demand and linear costs, which is developed in the next sections, satisfies these assumptions.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

time t by a generic firm adopting rule $i = 1, 2$.[5] All firms employ the same technology and bear the same production cost $C(x)$.

Behavioral rule i entails a fixed "information" cost $K_i \geq 0$. At time t, the first behavioral rule is distributed with frequency $r(t) \in [0, 1]$ among the firms and, obviously, the second rule with complementary frequency $1 - r(t)$.

A heuristic, or adaptive behavioral rule, can be defined as a rule that specifies next-period production $x_i(t+1)$ as a function of the current quantities[6] $x_j(t)$, $i, j = 1, 2$, as well as of the frequency $r(t)$, i.e.[7]

$$x_i(t+1) = H_i(x_1(t), x_2(t), r(t)); \quad i = 1, 2 \tag{12.1}$$

In the following, we consider behavioral rules that are *stationary* at any symmetric Nash equilibrium of the underlying game. More precisely, we assume that if the industry is at a Nash equilibrium, each behavioral rule prescribes that the industry stays at that equilibrium, regardless of the distribution $r(t)$. By assumption, Nash equilibria of the game always exist. In particular, since agents are homogeneous, a symmetric Nash equilibrium exists, which is characterized by the same production by all agents, i.e. $x_1^* = x_2^* = x^*$. The stationary property at a symmetric equilibrium can be thus expressed as

$$x^* = H_i\left(x^*, x^*, r\right); \quad i = 1, 2 \tag{12.2}$$

Stated differently, a symmetric Nash equilibrium is a fixed point of the two-dimensional map (12.1) for any $r \in [0, 1]$. Now consider the profit obtained by employing behavioral rule i:

$$\pi_i(t) = \left[p(t) - C(t)\right] x_i(t) - K_i \tag{12.3}$$

and assume that firms can observe this quantity and switch, from period to period, to the more profitable behavioral rule, thus modifying the next-period distribution of behavioral rules. In the following, we adopt a model that we call *exponential replicator*, which was first proposed by Cabrales and Sobel (1992) (see also Hofbauer and Sigmund, 2003, Hofbauer and Weibull, 1996 and Kopel et al., 2014 for an application in oligopoly theory), which assumes the form

$$r(t+1) = \frac{r(t)\, e^{\beta \pi_1(t)}}{r(t)\, e^{\beta \pi_1(t)} + [1 - r(t)]\, e^{\beta \pi_2(t)}} = \frac{r(t)}{r(t) + [1 - r(t)]\, e^{\beta[\pi_2(t) - \pi_1(t)]}} \tag{12.4}$$

In (12.4) $\beta \geq 0$ is the *intensity of choice*, which measures how sensitive the players are at selecting behavioral rules with the best relative performances. The minimum value $\beta = 0$

[5] In the general setup, it is easy to generalize to M different behavioral rules. However, we present the idea with only two different rules, as in the second part of this chapter such an example is developed. Another example with two rules is proposed in Bischi, Lamantia, and Radi (2015).

[6] In general, a behavioral rule can also incorporate older information through a "memory" term; see Bischi, Lamantia, and Radi (2015) for details. However, we do not consider this issue in this chapter.

[7] Notice that in principle a behavioral rule H_i should have $N + 1$ arguments, i.e. the quantities by the N oligopolists and the fraction r. However, if agents that employ the same behavioral rule set the same quantities then each behavioral rule only depends on quantities of representative agents of the different groups. This reflects in the notation, with a slight abuse of it.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

corresponds to the case with fixed fractions, being $r(t+1) = r(t) = r$. The other extreme case, $\beta = \infty$, corresponds to a situation where all firms immediately switch to the behavioral rule showing a (even negligible) better performance, i.e. $r(t) \to 1$ if $\pi_1(t) > \pi_2(t)$ and $r(t) \to 0$ if $\pi_1(t) < \pi_2(t)$.

Coupling the behavioral rules in (12.1), which specify quantity dynamics, with the evolutionary dynamics in (12.4), a three-dimensional map T is defined in the phase space $(x, y, r) \in A \subseteq \mathbb{R}^2_+ \times [0, 1]$, where $\mathbb{R}_+ = [0, +\infty)$:

$$T: \begin{cases} x(t+1) = H_1(x(t), y(t), r(t)) \\ y(t+1) = H_2(x(t), y(t), r(t)) \\ r(t+1) = \dfrac{r(t)}{r(t) + [1 - r(t)]\, e^{\beta[\pi_2(t) - \pi_1(t)]}} \end{cases} \tag{12.5}$$

The choice of a specific evolutionary model, namely the exponential replicator, originates from its interesting properties. First of all, the strictly monotone transformation $\pi_i(t) \to e^{\beta\pi_i(t)}$ guarantees that the fractions obtained through (12.4) are always contained in the interval $[0, 1]$ even when $\pi_i(t) < 0$. Another important property of (12.4) concerns the role of invariant planes $r = 0$ and $r = 1$, where only one pure strategy is employed (H_2 or H_1 respectively). On these planes, the dynamics are governed by the two-dimensional restrictions of T on them. From an economic point of view, this fact has the obvious interpretation that absent behaviors remain absent. However, the introduction of a mutation in agents' behavior may spread over the population or may be reabsorbed. This phenomenon can be ascertained through the study of transverse stability of the attractors on the invariant planes. In general, an attractor on one of these two-dimensional restrictions of the phase space may be *transversely stable*, so that it attracts trajectories starting outside the restriction, i.e. from $r(0) \in (0, 1)$; in this case, the attractor on the restriction is also an attractor of the three-dimensional map T.

From (12.2), it follows that if productions are at a symmetric Nash equilibrium level x^*, then point $E_* = (x^*, x^*, r^*)$ is an equilibrium (or fixed point) of the three-dimensional map T when $r^* = 0$, $r^* = 1$ or $r^* \in (0, 1)$ such that $\pi_1 = \pi_2$. By abuse of notation, equilibrium $(x^*, x^*, 0)$ is also denoted as E_0 and $(x^*, x^*, 1)$ is also denoted as E_1. In addition, when the Nash equilibrium is stable with respect to the quantity adjustments (12.1) only then it is easy to characterize the asymptotic behavior of the evolutionary map T. In fact, regardless of the specific behavioral rules considered, the following stability properties hold:

Proposition 1 *Consider the dynamical system T defined in* (12.5) *with $\beta > 0$ and assume that the symmetric Nash equilibrium of quantity dynamics under constant r is locally asymptotically stable $\forall r \in [0, 1]$, ($\forall r \in [0, 1]$, (x^*, x^*) is a locally asymptotically stable fixed point of the two-dimensional map* (12.1) *with $r(t) = r$). The following hold:*

- *If $K_1 = K_2$, then a continuum of equilibrium points E_* exists along the segment $E = (x^*, x^*, r)$, with $r \in [0, 1]$. Each fixed point filling this invariant segment is stable.*
- *If $K_1 \neq K_2$, then the segment $E = \left\{(x^*, x^*, r) \in \mathbb{R}^2_+ \times [0, 1]\right\}$, is invariant for T and only the two extreme points of the segment E are equilibria, namely*

$$E_0 = \left(x^*, x^*, 0\right) \quad \text{and} \quad E_1 = \left(x^*, x^*, 1\right)$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

in which all agents adopt the same behavioral rule, which is H_2 or H_1 respectively. When $K_1 < K_2$ [$K_1 > K_2$] equilibrium E_1 [E_0] is locally asymptotically stable, whereas E_0 [E_1] is unstable.

Proof Assume that firms of either type produce the Nash equilibrium quantity x^*. Then, for any $r \in [0, 1]$, the difference in their profits is given by the information costs, $\pi_1 - \pi_2 = K_2 - K_1$ (see (12.3)). Therefore, if fixed information costs are equal, i.e. $K_1 = K_2$, the replicator equation in (12.4) reduces to $r(t+1) = r(t)$ so that any point of the form $E = (x^*, x^*, r)$, $r \in [0, 1]$, is a fixed point for map T. Instead, if $K_1 \neq K_2$, then at any point of E, it is $\pi_2 \neq \pi_1$, so that the stationary condition can be satisfied only at the boundary points $E_0 = (x^*, x^*, 0)$, and $E_1 = (x^*, x^*, 1)$, with all agents employing the same behavioral rule. Stability analysis of equilibria can be studied through the Jacobian matrix, which assumes the following form:

$$J(x^*, y^*, r) = \begin{pmatrix} \dfrac{\partial H_1(x^*, y^*, r)}{\partial x_1} & \dfrac{\partial H_1(x^*, y^*, r)}{\partial x_2} & 0 \\ \dfrac{\partial H_2(x^*, y^*, r)}{\partial x_1} & \dfrac{\partial H_2(x^*, y^*, r)}{\partial x_2} & 0 \\ \dfrac{\partial r}{\partial x_1} & \dfrac{\partial r}{\partial x_2} & \dfrac{e^{\beta(\pi_1+\pi_2)}}{\left((r-1)e^{\beta\pi_2} - re^{\beta\pi_1}\right)^2} \end{pmatrix}$$

From (12.2), the entries J_{13} and J_{23} of the Jacobian matrix are equal to zero, and the characteristic equation becomes $P(z) = \left(\frac{e^{\beta(\pi_1+\pi_2)}}{\left((r-1)e^{\beta\pi_2} - re^{\beta\pi_1}\right)^2} - z\right) P_2(z)$, where $P_2(z)$ is the characteristic equation of the two-dimensional model (12.1), whose roots are in modulus less than 1 by the assumption of stability of the Nash equilibrium with respect to quantity dynamics. Thus, when $\pi_1 = \pi_2$, which occurs when $K_1 = K_2$, $z_3 = 1$ is an eigenvalue of $J(x^*, y^*, r)$, and any point of the form $E = (x^*, x^*, r)$, with $r \in [0, 1]$, is a stable equilibrium. If $K_1 < K_2$, it is $T(E) \subset E$ by the assumption that (12.1) is stationary at a symmetric Nash equilibrium. Thus, E is invariant for T. At $r = 0$, the third eigenvalue is $z_3 = \frac{e^{\beta(\pi_1+\pi_2)}}{\left((r-1)e^{\beta\pi_2} - re^{\beta\pi_1}\right)^2} = e^{\beta(K_2-K_1)} \in (1, +\infty)$, whereas at $r = 1$ the third eigenvalue is $z_3 = e^{\beta(K_1-K_2)} \in (0, 1)$, thus proving the statement. The eigenvector associated with the third eigenvalue is clearly $(0, 0, 1)$. The case $K_2 < K_1$ is analogous and left to the reader. ■

The conditions under which the symmetric Nash equilibrium is stable can be violated and some attractors can be created through bifurcations, as stated in the following corollary:

Corollary *Consider the dynamical system T defined in* (12.5) *with $\beta > 0$. When the conditions for local asymptotic stability of the Nash equilibrium (x^*, x^*) are broken in* (12.1) *with $r(t) = r^*$, the following cases occur:*

- *If $K_1 = K_2$, an attractor can appear in $\mathbb{R}_+ \times \mathbb{R}_+ \times [0, 1]$ (or in $\mathbb{R}_+ \times \mathbb{R}_+ \times (0, 1)$ if $r^* \notin \{0, 1\}$) through a bifurcation of codimension 1, 2 or 3.*
- *If $K_1 < K_2$ and $r^* = 1$ [$K_1 > K_2$ and $r^* = 0$], then equilibrium E_1 [E_0] undergoes a bifurcation and an attractor may appear in the invariant subspace $r = 1$ [$r = 0$], whereas E_0 [E_1] remains unstable.*

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Proof The first part of the corollary follows the observation that two of the roots of the characteristic polynomial associated with the Jacobian matrix of dynamical system T (defined in (12.5) with $\beta > 0$) computed at the fixed point $E_* = (x^*, x^*, r^*)$, coincide with the two roots of the characteristic polynomial of the Jacobian matrix of system (12.1) computed at the fixed point (x^*, x^*) with $r(t) = r^*$. The second part of the corollary follows by noting that the dynamics on the invariant plane $r = 1$ (or $r = 0$) of the dynamical system T is equal to the one of system (12.1) with $r(t) = 1$ (or $r(t) = 0$). ■

On the other hand, if the Nash equilibrium is unstable for the quantity adjustments (12.1) and more complex attractors exist, then the asymptotic behavior of the model becomes more complicated but also more interesting both from a mathematical as well as economic point of view. In the rest of the chapter, we develop this model for two different pairs of behavioral rules, namely best reply (BR) vs local monopolistic approximation (LMA) as a first case and gradient dynamics (GD) vs LMA as a second, and we investigate the main dynamic properties of the two systems.

4 AN EVOLUTIONARY OLIGOPOLY CHARACTERIZED BY ISOELASTIC DEMAND

The type of behavioral rule adopted by firms depends on the market structure of the oligopoly. For a quantity-setting oligopoly, the inverse market demand and the production cost function are essential ingredients with which to identify the game for which the behavioral rules are employed.

Concerning the inverse market demand, let us assume that firms produce homogeneous goods (commodities) and their entire output is sold in the market at a price that is determined according to an isoelastic demand function, with constant elasticity equal to one. This means that the inverse demand function (or price function) is given by

$$p = f(Q) = \frac{1}{Q} \tag{12.6}$$

where $Q = \sum_{i=1}^{N} q_i > 0$ is the total industry output and p is the selling price.[8] This particular demand function is widely employed in the literature; see, e.g., Puu (1991), Bischi et al. (2010), Tramontana, Gardini, and Puu (2010), Agliari, Gardini, and Puu (2006), Lambertini (2010) and Lamantia (2011).[9] Concerning the production costs, they are assumed to be a linear function of the output of each firm:

$$C_i(q_i) = c \cdot q_i \tag{12.7}$$

[8] The reader will note that for $Q = 0$ this (inverse) demand function is undefined. Here we avoid discussing such technical aspects, for which the interested reader can consult Cerboni Baiardi et al. (2015). In the following, we focus on dynamics such that $Q > 0$ and refer to the case $Q = 0$ as the *infeasibility of the oligopoly*.

[9] In particular, isoelastic demand is obtained when a representative consumer maximizes a log-linear (or Cobb-Douglas) utility function; see Lambertini (2010) for details.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

where q_i is the quantity produced by firm i, $i = 1, \ldots, N$, and $c > 0$ denotes the marginal cost (here assumed the same for all the firms). It follows that the profit function of the generic firm i is given by:

$$\pi_i(q_i, Q_{-i}) = f(Q_{-i} + q_i)\, q_i - C_i(q_i), \; i = 1, \ldots, N$$

where $Q_{-i} = Q - q_i$. Let us indicate by $Q^e_{-i}(t+1)$ firm i's prediction at time t for time $t+1$ production of the rest of the industry and let us assume that the actual production $Q_{-i}(t)$ is employed as a proxy for $Q^e_{-i}(t+1)$. In other words, firms believe that competitors' output will not change in the next period (naive expectations). Moreover, we assume that firms may gain knowledge of the form of the price function $p = f(Q)$, $\forall Q > 0$, or at least they can acquire such information by paying a cost. Given this wealth of knowledge, firms decide their own next-period level of production in accordance with the rational principle of profit maximization. In the attempt to reach such a target, each firm has the possibility of employing different levels of knowledge or types of behavior, from which we obtain different behavioral rules.

Consistent with the rational principle of increasing own profits and coherently with the level of knowledge assumed, a variety of behavioral rules are available. Let us focus on three of them, which are also among the most well known in literature. The first one is the best reply (BR) rule with naive expectation (see Bischi et al., 2007), according to which:

$$\begin{aligned} q_i(t+1) &= \max\left\{0, \arg\max_{q_i} \pi_i\left(q_i, Q^e_{-i}(t+1)\right)\right\} = \\ &= \max\left\{0, \arg\max_{q_i} \left[f(Q_{-i}(t) + q_i)\, q_i - C_i(q_i)\right]\right\} \end{aligned}$$

This behavioral rule requires global knowledge of the inverse demand function, i.e. the maximum level of market cognizance that we can assume.

Another behavioral rule is the so-called local monopolistic approximation (LMA) rule (see Bischi et al., 2007), according to which:

$$q_i(t+1) = \max\left\{0, \arg\max_{q_i} \left[\left(f(Q(t)) + \frac{\partial f(Q(t))}{\partial q_i}(q_i - q_i(t))\right) q_i - C_i(q_i)\right]\right\}$$

This behavioral rule does not require any type of expectation about the production of the rest of the industry, as a firm employing LMA does not take into account the negative strategic effect of competitors' outputs when attempting to maximize its own profit. All in all, the LMA rule requires only local knowledge of the demand function, and thus, it is fully consistent with the maximum wealth of knowledge that we have assumed.

A third behavioral rule is the so-called gadient dynamics (GD) rule (see Bischi and Naimzada, 1999), according to which:

$$q_i(t+1) = \max\left\{0, q_i(t)\left(1 + \frac{\partial \pi_i(q_i(t), Q_{-i}(t))}{\partial q_i}\right)\right\}$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Following the latter behavioral rule, a firm adjusts its output according to its marginal profit, i.e. the relative change in the quantity produced at time $t + 1$ is equal to the marginal profit experienced by the firm at time t.[10] Thus, the ability to compute or estimate own marginal profits is the only ability required to a firm to employ this behavioral rule. Among the three, this is the only behavioral rule that does not require the solution to an optimization problem.

It follows that the three behavioral rules can be ranked on the basis of the knowledge that they require. The BR and the GD rules are, respectively, the more and the less advanced and sophisticated of the three. The LMA rule represents a middle ground among the three. This could lead one to think that if firms can freely decide which behavioral rule to adopt and their cost is the same, they will choose the BR rule; equivalently, the BR rule would seem to be the behavioral rule that achieves higher profits. However, this would not be the case if firms base their decisions on relative performance indices. For example, a strategy that is the best in terms of absolute performances is not automatically evolutionarily desirable. Evolutionary game theory provides an explanation for this occurrence in terms of spiteful behaviors. A behavioral rule (or, generally speaking, a strategy) can be evolutionarily stable because in adopting that behavioral rule a firm damages the rivals more than it damages itself. Then, to study the evolution of firms' preferences on the behavioral rules and the consequent industrial outputs and profits, one must study the dynamics of a related evolutionary oligopoly model. In an attempt to do so, in the following we analyze and discuss two different cases, with competition between two behavioral rules at a time. In the first example we allow firms to choose between BR rule and LMA rule, while in the second case firms can choose between GD rule and LMA rule. The investigation reveals interesting phenomena, such as *evolutionary stable heterogeneity* and complex dynamics that coincide with performances (in terms of average profits of the industry) higher than the ones observed at the Nash equilibrium.

4.1 An Example with Best Reply and Local Monopolistic Approximation Behavioral Rules

Here, following Bischi, Lamantia, and Radi (2015), the general model introduced in Section 3 is applied to describe the time evolution of a population of firms that can choose between the BR rule and the LMA rule in order to update their output decisions. A firm that adopts the BR rule is called BR firm while a firm that adopts the LMA rule is called LMA firm.

At a given time t, the fraction of BR firms is denoted by $r(t) \in [0, 1]$.[11] Then, $1 - r(t)$ is the fraction of LMA firms. The outputs at time t of a representative BR firm and a representative LMA firm are denoted, respectively, by $x(t)$ and $y(t)$. Taking

$$(N - 1)\left[r(t)x(t) + (1 - r(t))\, y(t)\right] \tag{12.8}$$

[10] In general, this rule is given in the form

$$q_i\,(t+1) = \max\left\{0, q_i\,(t)\left(1 + \xi_i \frac{\partial \pi_i\,(q_i\,(t)\,, Q_{-i}\,(t))}{\partial q_i}\right)\right\}$$

where $\xi_i \geq 0$ can be interpreted as a speed of adjustment (see Bischi and Naimzada, 1999 and Bischi et al., 2010). Here, for the sake of brevity, we assume that $\xi_i = 1$ for all $i = 1, \ldots, N$.

[11] It is worth pointing out that the use of the replicator dynamics, in whatever form, implies a random matching, see, e.g., Plank (1997). Then, $r\,(t)$ must be interpreted as the probability to meet a BR firm. Nevertheless, with an abuse of language we call $r\,(t)$ the fraction of BR firms.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

as a proxy for $Q_{-i}(t)$ and denoting by H_{BR} the heuristic that identifies the BR rule and by H_{LMA} the heuristic that identifies the LMA rule, these can be defined by the following functions:

$$H_{BR}(x,y,r) = \max\left\{0, \sqrt{\frac{(N-1)(rx+(1-r)y)}{c}} - (N-1)(rx+(1-r)y)\right\}$$

$$H_{LMA}(x,y,r) = \max\left\{0, \frac{y+N\left[rx+(1-r)y\right]\left(1-cN\left[rx+(1-r)y\right]\right)}{2}\right\}$$

where the first one specifies the output of a BR firm while the second one that of an LMA firm. Substituting H_1 with H_{BR} and H_2 with H_{LMA} in model (12.5), the following three-dimensional map $T : \mathbb{R}^2_+ \times [0,1] \longrightarrow \mathbb{R}^2_+ \times [0,1]$ is obtained:

$$T : \begin{cases} x(t+1) = H_{BR}(x(t), y(t), r(t)) \\ y(t+1) = H_{LMA}(x(t), y(t), r(t)) \\ r(t+1) = \dfrac{r(t)e^{\beta\pi_{BR}(t)}}{r(t)e^{\beta\pi_{BR}(t)} + (1-r(t))e^{\beta\pi_{LMA}(t)}} \end{cases} \tag{12.9}$$

where $\pi_i(t)$, $i \in \{BR, LMA\}$, are, respectively:

$$\pi_{BR}(t) = p(t)x(t) - (cx(t) + K_{BR}) = \left(\frac{N-1}{NQ_{-1}(t)} - c\right)x(t) - K_{BR} \tag{12.10}$$

$$\pi_{LMA}(t) = p(t)y(t) - (cy(t) + K_L) = \left(\frac{N-1}{NQ_{-1}(t)} - c\right)y(t) - K_{LMA}$$

and K_{BR} and K_{LMA} are the respective information costs defined in Section 3.

Given the market structure of the oligopoly previously described, the unique Nash quantity of the game is

$$q^{NE} = \frac{N-1}{cN^2} > 0 \tag{12.11}$$

(see Bischi et al., 2010 for details). Moreover, it is easy to show that:

$$q^{NE} = H_{BR}\left(q^{NE}, q^{NE}, r\right) = H_{LMA}\left(q^{NE}, q^{NE}, r\right) \ \forall r \in [0,1]$$

It follows that the fixed points of the model (12.9) are of the form

$$E_* = \left(q^{NE}, q^{NE}, r^*\right) \text{ with } r^* \in [0,1],$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

which we call Nash equilibria independently of the value of r^*. Stability properties of these Nash equilibria are specified in the following proposition:

Proposition 2 *Consider the dynamical system T in (12.9):*

- *If $K_{BR} = K_{LMA}$, then a continuum of equilibrium points E_* exists along the segment $E = \left(q^{NE}, q^{NE}, r\right)$, with $r \in [0, 1]$. A fixed point in E has an associated eigenvalue equal to 1, one positive eigenvalue less than 1, and a negative eigenvalue greater than -1 for $N < N^+ (r^*)$ where*

$$N^+ \left(r^*\right) = \frac{10 + 2r^*}{2 + r^*}$$

- *If $K_{BR} \neq K_{LMA}$, then only the two extremum points of the segment E are equilibria, namely*

$$E_0 = \left(q^{NE}, q^{NE}, 0\right) \quad \text{and} \quad E_1 = \left(q^{NE}, q^{NE}, 1\right)$$

  in which all agents adopt the same strategy, which is LMA or BR respectively:

  – *If $K_{LMA} < K_{BR}$, then equilibrium E_0 is stable for $N < 5$, and it loses stability through a bifurcation of eigenvalue -1 at $N = 5$. E_0 is unstable for $N > 5$. E_1 is always unstable.*
  – *If $K_{LMA} > K_{BR}$, then equilibrium E_1 is stable for $N < 4$, and it loses stability through a bifurcation of eigenvalue -1 at $N = 4$. E_1 is unstable for $N > 4$. E_0 is always unstable.*

Proof The local stability properties of the Nash equilibria are determined through the study of the associated Jacobian matrix. For $K_{BR} = K_{LMA}$ we have

$$J\left(E_*\right) = \begin{bmatrix} \dfrac{(2-N)\,r^*}{2} & \dfrac{(2-N)\,(1-r^*)}{2} & 0 \\ \dfrac{(2-N)\,r^*}{2} & \dfrac{1}{2} + \dfrac{(2-N)\,(1-r^*)}{2} & 0 \\ J_{31} & J_{32} & 1 \end{bmatrix}$$

Hence, the characteristic equation is $(1-\lambda)\,P(\lambda) = (1-\lambda)\left(\lambda^2 - \frac{3-N}{2}\lambda + \frac{2-N}{4}r^*\right) = 0$. From which the three eigenvalues

$$\lambda_1 + \lambda_2 = \frac{3-N}{2} \leq \frac{1}{2},\, \lambda_1\lambda_2 = \frac{2-N}{4}r^* \leq 0,\, \lambda_3 = 1$$

It follows that one eigenvalue, say λ_1, is positive and another one, say λ_2, is negative. Since $P(1) > 0\ \forall r^* \in (0, 1)$, it follows that $0 \leq \lambda_1 < 1$. By imposing $P(-1) > 0$, we obtain the

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

condition $N < N^{+}(r^{*})$, which guarantees that $-1 < \lambda_2 < 0$. For $K_{BR} \neq K_{LMA}$, by similar calculations we obtain the eigenvalues associated with E_0:

$$\lambda_1 = 0, \lambda_2 = \frac{3-N}{2}, \lambda_3 = e^{\beta(K_{LMA}-K_{BR})}$$

By imposing $\lambda_i \in (-1, 1)$, $i = 1, 2, 3$, the stability conditions for E_1 follow, i.e. $K_{LMA} < K_{BR}$ and $N < 5$. For $K_{LMA} < K_{BR}$ and $N = 5$, one eigenvalue associated with E_0 is equal to -1 while the other two belong to the interval $(-1, 1)$. Thus, E_0 undergoes a bifurcation of eigenvalue -1. For $N > 5$, one eigenvalue associated with E_0 is always smaller than -1. Hence, E_0 is unstable. In the same way, we obtain the eigenvalues associated with E_1:

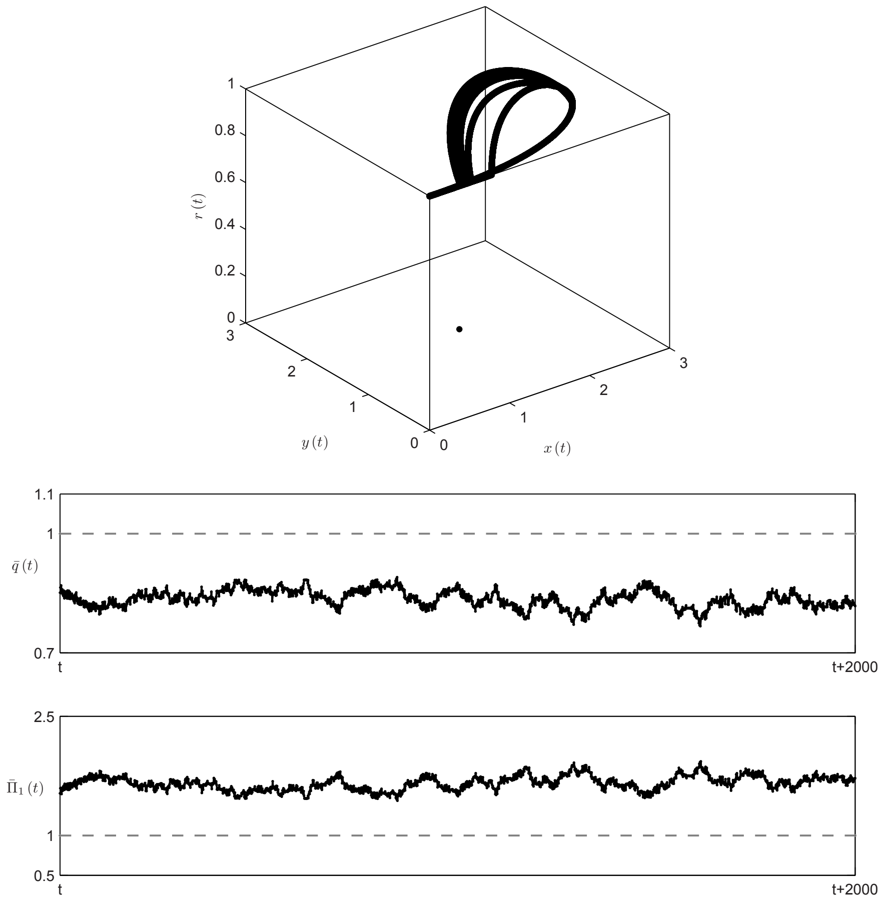
$$\lambda_1 = \frac{2-N}{2}, \lambda_2 = \frac{1}{2}, \lambda_3 = e^{\beta(K_{BR}-K_{LMA})}$$

Again, by imposing $\lambda_i \in (-1, 1)$, $i = 1, 2, 3$, the stability conditions for E_1 follow, i.e. $K_{LMA} > K_{BR}$ and $N < 4$. Moreover, for $K_{LMA} > K_{BR}$ and $N = 4$, one eigenvalue associated with E_1 is equal to -1 while the other two belong to the interval $(-1, 1)$. Thus, E_1 undergoes a bifurcation of eigenvalue -1. For $N > 4$, one eigenvalue associated with E_1 is always smaller than -1. Hence, E_1 is unstable. ■

The proposition underlines the crucial role played by $K_{BR} - K_{LMA}$. When firms have a natural propensity to play the LMA rule ($K_{BR} - K_{LMA} > 0$), or equivalently the LMA rule is cheaper than the BR rule, at the Nash equilibrium all firms will adopt the LMA rule. The opposite occurs for $K_{BR} - K_{LMA} < 0$. Another interesting aspect that arises from the analysis concerns the stability of the Nash equilibrium, which loses stability as the number of firms involved in the oligopoly increases. Nevertheless, the exact number of firms at which the Nash equilibrium loses stability depends on the type of behavioral rule the firms decide to adopt. When firms have a natural propensity to play the LMA rule, the Nash equilibrium loses stability when the number of firms is greater than five. When firms have a natural propensity to play the BR rule, the Nash equilibrium loses stability already with four firms. This observation is relevant to clarify the different complexity of the dynamics of the two behavioral rules, as explained in the example that follows.

Let us start by assuming that firms have a natural propensity to adopt the LMA rule ($K_{LMA} < K_{BR}$). By Proposition 2 and numerical simulations, E_0 is the unique attractor and E_1 is transversely unstable when $N = 3$. Instead, when $N = 4$, E_1 undergoes a period-doubling bifurcation through which a two-cycle appears in the subspace $r = 1$. Through the usual period-doubling cascade an attractor (either periodic or chaotic) appears in the subspace $r = 1$. For $N < 5$, this attractor coexists with the stable Nash equilibrium E_0, then the model is characterized by path dependence. Starting with a large fraction of LMA firms, the system converges to E_0; otherwise, in the long run all firms will adopt the BR rule and the system will be characterized by either periodic or aperiodic oscillations in the level of the outputs. A detailed analysis of the basins of these two attractors can be found in Bischi, Lamantia, and Radi (2015). Despite the oscillations of the output dynamics, adopting the BR rule is the choice implied by evolutionary pressure. The average level of profits are substantially higher than the ones experienced at the Nash equilibrium. Then, the output oscillations represent a profitable, and so desirable, alternative to the Nash equilibrium. One typical example of such

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham



Note: Parameters: $N = 5$, $c = 0.1$, $\beta = 1$, $K_{GD} - K_{LMA} = 0.00001$. First line, dynamics of the model in the phase space (attractor on the plane $r = 1$ and the Nash equilibrium E_0). Second line, time series, in black $\overline{q}(t) = MA\left(r(t)\,x(t) + (1 - r(t))\,y(t)\,/q^{NE}\right)$, where q^{NE} is the level of production at the Nash equilibrium (depicted by a gray and dashed line). Third line, $\overline{\Pi}(t) = MA\left((r\Pi_{GD} + (1 - r)\,\Pi_{LMA})\,/\,\Pi^{NE}\right)$, where Π^{NE} is the level of profit at the Nash equilibrium (depicted by a gray and dashed line) and MA stands for moving average with 80 lags. Initial conditions: $x(0) = q^{NE} + 0.002$, $y(0) = q^{NE} - 0.002$, $r(0) = 0.95$.

Figure 12.1 Dynamics of model 12.9

behavior is depicted in Figure 12.1 where $N = 5$. For $N \geq 5$, all the Nash equilibria, i.e., E_0 and E_1, are unstable.

4.2 An Example with Gradient Dynamics and Local Monopolistic Approximation Behavioral Rules

Here, following Cerboni Baiardi et al. (2015), the general model described in Section 3 is applied to study the time evolution of a population of firms that can choose between GD rule and LMA rule in order to update their output decisions. A firm that adopts the GD rule is

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

called GD firm while a firm that adopts the LMA rule is called LMA firm. At a given time t the fraction of GD firms is denoted by $r(t) \in [0,1]$. Then, $1-r(t)$ is the fraction of LMA firms. The output at time t of a representative GD firm and a representative LMA firm are denoted, respectively, by $x(t)$ and $y(t)$. Again, using a proxy such as (12.8) for $Q_{-i}(t)$ and denoting by H_{GD} and by H_{LMA} the heuristics that identify the GD and the LMA rule respectively, they can be written through the functions

$$H_{GD}(x,y,r) = \max\left\{0, x + x\left(\frac{(N-1)\left[rx+(1-r)y\right]}{(x((N-1)r+1)+y(N-1)(1-r))^2} - c\right)\right\}$$

$$H_{LMA}(x,y,r) = \max\left\{0, \frac{y + N\left[rx+(1-r)y\right]\left(1 - cN\left[rx+(1-r)y\right]\right)}{2}\right\}$$

where the first one specifies the output of a GD firm while the second one that of an LMA firm. Following the steps of the previous case, we can substitute in model (12.5) H_1 with H_{GD} and H_2 with H_{LMA}, thus obtaining the following three-dimensional map $T : A \longrightarrow A$ (with $A \subseteq \mathbb{R}^2_+ \times [0,1]$):[12]

$$T : \begin{cases} x(t+1) = H_{GD}(x(t), y(t), r(t)) \\ y(t+1) = H_{LMA}(x(t), y(t), r(t)) \\ r(t+1) = \dfrac{r(t)e^{\beta\pi_{GD}(t)}}{r(t)e^{\beta\pi_{GD}(t)} + (1-r(t))e^{\beta\pi_{LMA}(t)}} \end{cases} \tag{12.12}$$

where $\pi_i(t)$, $i \in \{GD, LMA\}$, are, respectively:

$$\pi_{GD}(t) = p(t)x(t) - (cx(t) + K_G) = \left(\frac{N-1}{NQ_{-1}(t)} - c\right)x(t) - K_{GD} \tag{12.13}$$

$$\pi_{LMA}(t) = p(t)y(t) - (cy(t) + K_L) = \left(\frac{N-1}{NQ_{-1}(t)} - c\right)y(t) - K_{LMA}$$

and K_{GD} and K_{LMA} are the information costs of the two behavioral rules. Clearly, the unique Nash quantity of the game is again (12.11) and it is easy to show that:

$$q^{NE} = H_{GD}\left(q^{NE}, q^{NE}, r\right) = H_{LMA}\left(q^{NE}, q^{NE}, r\right) \; \forall r \in [0,1]$$

It follows that, as for model (12.9), the equilibria of the model (12.12) are

$$E_* = \left(q^{NE}, q^{NE}, r^*\right) \text{ with } r^* \in [0,1]$$

[12] Note that the region $\mathbb{R}^2_+ \backslash A$ where map T is not defined, represents cases of "*infeasibility of the oligopoly*". It is worth pointing out that the infeasibility of the oligopoly is to impute that the demand function is not defined for $Q(t) = 0$. See the previous discussion and Cerboni Baiardi et al. (2015) and references therein for details.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

which, again, we call Nash equilibria independently of the value r^*. Stability properties of these Nash equilibria are specified in the following proposition:

Proposition 3 *Consider the dynamical system T in (12.12):*

- *If* $K_{GD} = K_{LMA}$*, then a continuum of equilibrium points* E_* *exists along the segment* $E = \left(q^{NE}, q^{NE}, r\right)$*, with* $r \in [0,1]$. *One fixed point in E undergoes a codimension-two bifurcation changing* $K_{GD} - K_{LMA}$ *with an associated eigenvalue equal to 1 and another one equal to −1, with the possible creation of a stable two-cycle* $\left\{\left(\underline{x}, \underline{y}, \underline{r}\right), (\overline{x}, \overline{y}, \overline{r})\right\}$. *The point in E that undergoes the codimension-two bifurcation is the one such that* $\left(q^{NE}, q^{NE}, r^*\right) = \left(\underline{x}, \underline{y}, \underline{r}\right) = (\overline{x}, \overline{y}, \overline{r})$ and $c = c^+(N, r^*)$, *where*

$$c^+\left(N, r^*\right) = \frac{2N\left(3 + (2-N)\left(1-r^*\right)\right)}{8 - 2(N-1) - 5(2-N)r^*}$$

- *If* $K_{GD} \neq K_{LMA}$*, then only the two extremum points of the segment E are equilibria, namely*

$$E_0 = \left(q^{NE}, q^{NE}, 0\right) \quad \textit{and} \quad E_1 = \left(q^{NE}, q^{NE}, 1\right)$$

in which all agents adopt the same strategy, which is LMA or GD respectively:

  - *If* $K_{LMA} < K_{GD}$*, then equilibrium* E_0 *is stable for* $2 \leq N < 5$ *and* $c < N$*, and it loses stability through a bifurcation of eigenvalue* -1 *at* $N = 5$ *given* $c < N$. E_0 *is unstable for* $N > 5$. E_1 *is always unstable.*
  - *If* $K_{LMA} > K_{GD}$*, then equilibrium* E_1 *is stable for* $c < 2$ *and undergoes a bifurcation of eigenvalue* -1 *at* $c = 2$. E_1 *is unstable for* $c > 2$. E_0 *is always unstable.*

Proof The local stability properties of the Nash equilibria are determined through the study of the associated Jacobian matrix. For $K_{GD} = K_{LMA}$ we have

$$J\left(E_*\right) = \begin{bmatrix} 1 + c\frac{(2-N)r^*-2}{N} & \frac{c(2-N)\left(1-r^*\right)}{N} & 0 \\ \frac{(2-N)r^*}{2} & \frac{1}{2} + \frac{(2-N)\left(1-r^*\right)}{2} & 0 \\ J_{31} & J_{32} & 1 \end{bmatrix}$$

Hence, the characteristic equation is $(1-\lambda)P(\lambda) = (1-\lambda)\left(\lambda^2 - Tr\left(J\left(E_*\right)\right)\lambda + \det\left(J\left(E_*\right)\right)\right) = 0$, where

$$\det\left(J\left(E_*\right)\right) = \frac{(2-N)\left(1-r^*\right)+1}{2} + c\frac{(2-N)r^* - 2 - 2(2-N)\left(1-r^*\right)}{2N}$$

$$Tr\left(J\left(E_*\right)\right) = \frac{(2-N)\left(1-r^*\right)+3}{2} + c\frac{(2-N)r^*-2}{N}$$

Thus, one eigenvalue associated with the equilibrium is equal to 1, say $\lambda_3 = 1$. Moreover, to have the other two eigenvalues inside the unit circle (in the complex plane) it is required that

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

the following system of inequalities (known as Schur or Jury's conditions, see, e.g. Medio and Lines, 2001) are satisfied:

$$\begin{cases} P(1) > 0 \\ P(-1) > 0 \\ 1 - \det(J(E_*)) > 0 \end{cases} \tag{12.14}$$

Condition $1 - \det(J(E_*)) > 0$ is always satisfied. Thus, E_* cannot lose stability through a Neimark-Sacker bifurcation. Condition $P(1) > 0$ is always satisfied as well, being:

$$P(1) = c\frac{(N-2)r^* + (N-2)(1-r^*) + 2}{N} > 0$$

From $P(-1) = 0$, we have $c = c^+(N, r^*)$, where

$$c^+\left(N, r^*\right) = \frac{2N(3 + (2-N)(1-r^*))}{8 - 2(N-1) - 5(2-N)r^*}$$

Then, the equilibrium E_* has an associated eigenvalue equal to -1, one equal to 1 and another that takes a value between -1 and 1. These conditions are necessary for a degenerate codimension-two bifurcation through which a two-cycle is created. The sufficient condition can be verified by standard calculations. For $K_{GD} \neq K_{LMA}$, by similar calculations, we obtain the eigenvalues associated with E_0:

$$\lambda_1 = \frac{N-2c}{N}, \lambda_2 = \frac{3-N}{2}, \lambda_3 = e^{\beta(K_{LMA}-K_{GD})}$$

By imposing $\lambda_i \in (-1, 1)$, $i = 1, 2, 3$, the stability conditions for E_0 follow, i.e. $K_{LMA} < K_{GD}$ and $2 \leq N < 5$ and $c < N$. For $K_{LMA} < K_{GD}$, $c < N$ and $N = 5$, one eigenvalue associated with E_0 is equal to -1 while the other two belong to the interval $(-1, 1)$. Thus, E_0 loses stability through a bifurcation of eigenvalue -1. For $N > 5$, one eigenvalue associated with E_0 is always smaller than -1. Hence, E_0 is unstable. In the same way, we obtain the eigenvalues associated with E_1:

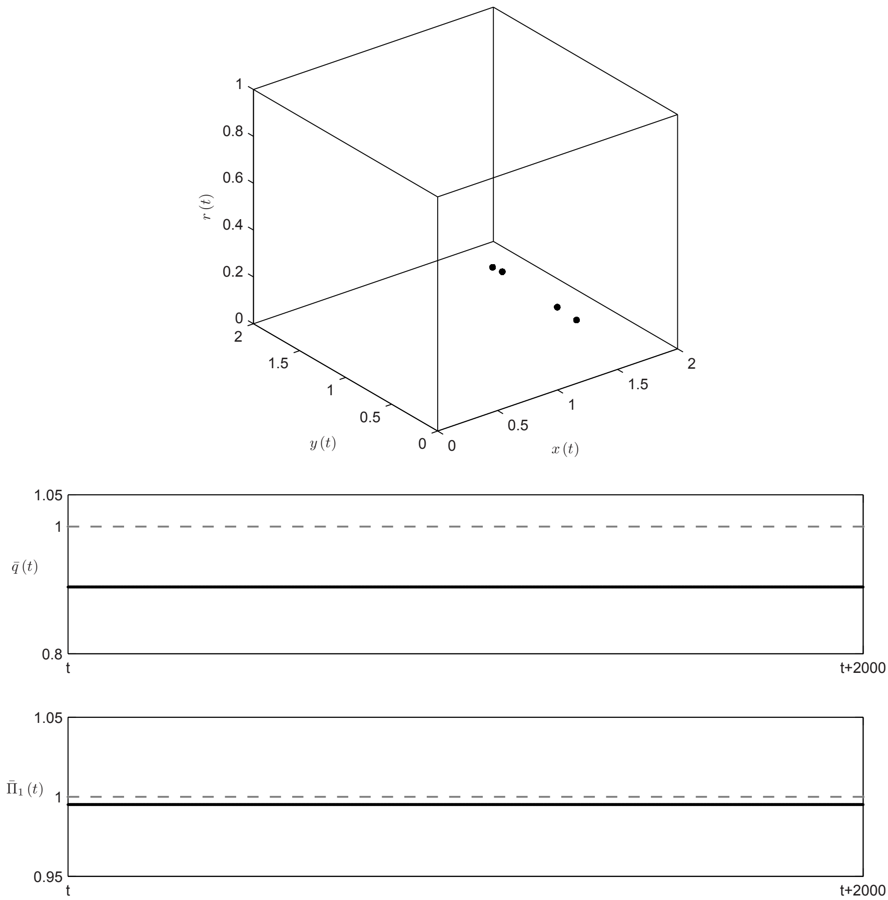
$$\lambda_1 = 1 - c, \lambda_2 = \frac{1}{2}, \lambda_3 = e^{\beta(K_{GD}-K_{LMA})}$$

Again, by imposing $\lambda_i \in (-1, 1)$, $i = 1, 2, 3$, the stability conditions for E_1 follow, i.e. $K_{LMA} > K_{GD}$ and $c < 2$. For $K_{LMA} > K_{GD}$ and $c = 2$, one eigenvalue associated with E_1 is equal to -1 while the other two belong to the interval $(-1, 1)$. Thus, E_1 undergoes a bifurcation of eigenvalue -1. For $c > 2$, one eigenvalue associated with E_1 is always smaller than -1. Hence, E_1 is unstable. ■

The proposition underlines the crucial role played by the difference in information costs $K_{LMA} - K_{GD}$. When firms have a natural propensity to play the LMA rule, that is, the LMA rule is cheaper than the GD rule ($K_{LMA} - K_{GD} < 0$), at the Nash equilibrium all firms will adopt the LMA rule. The opposite occurs for $K_{LMA} - K_{GD} > 0$. Nevertheless, the Nash equilibrium can lose stability, for example when the number of firms N is larger than 5, and in this case

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

two interesting scenarios emerge. The first is characterized by homogeneity in the choice of the behavioral rule to adopt. All firms will be LMA for $K_{LMA} - K_{GD} < 0$ while all firms will be GD for $K_{LMA} - K_{GD} > 0$, with periodic or aperiodic dynamics of the output. This occurs when the difference between K_{LMA} and K_{GD} is relatively large. For example, let us assume that firms have a propensity to play the LMA rule, i.e. $K_{LMA} - K_{GD} < 0$. Moreover, let us assume that the marginal production cost is $c = 0.1$. This implies that the condition $c < N$ is always guaranteed. So, the equilibrium E_0 is stable for $N < 5$ and it loses stability for $N > 5$. Then, let us consider the case with $N = 6$. If firms have a "strong" propensity to adopt the LMA rule, say $K_{LMA} - K_{GD} = -0.1$, the output dynamics is periodic and all the firms adopt the LMA rule as shown by numerical simulations; see Figure 12.2. In this case, the loss of stability of the Nash equilibrium is related to a reduction of both the level of production and of the profit of a single firm; see again Figure 12.2.



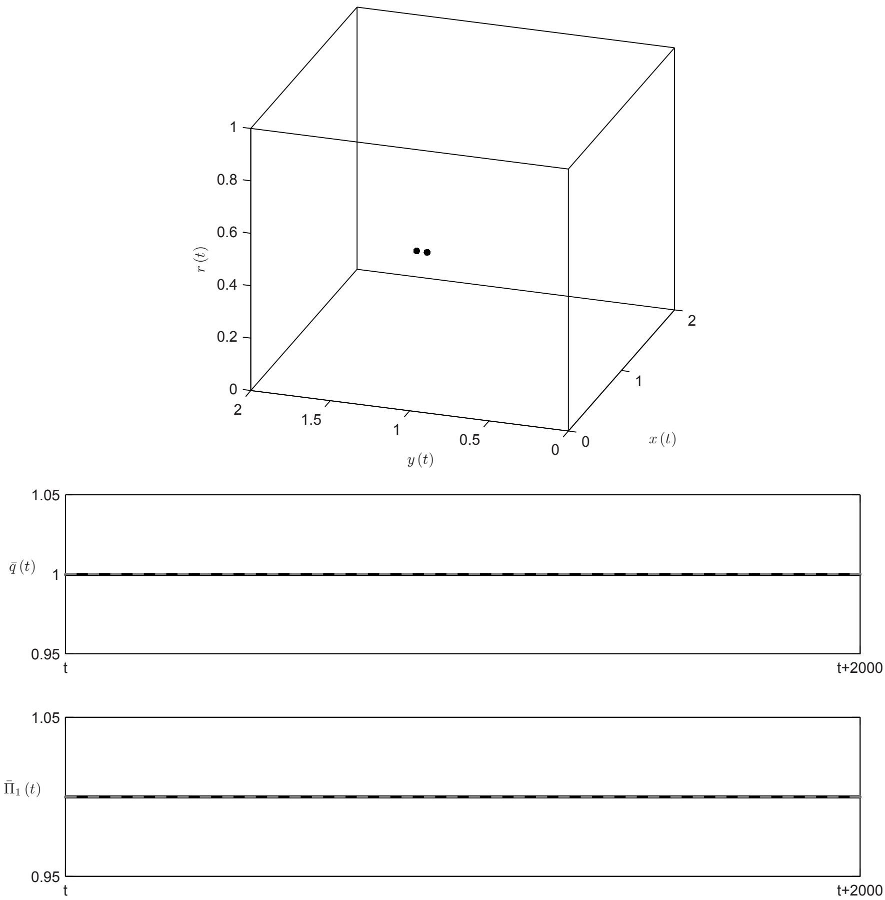
Note: Parameters: $N = 6$, $c = 0.1$, $\beta = 1$, $K_{GD} - K_{LMA} = 0.1$. Initial conditions: $x(0) = q^{NE} + 0.002$, $y(0) = q^{NE} - 0.002$, $r(0) = 0.1$. The meaning of the pictures is the same as in Figure 12.1.

Figure 12.2 Dynamics of model 12.12

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

The scenario is different when the propensity to play one behavioral rule over the other is small. Let us still consider a natural propensity to play the LMA rule, but smaller than before, say $K_{LMA} - K_{GD} = -0.0001$. Numerical simulations show that trajectories converge to a two-cycle and firms are heterogeneous; see Figure 12.3. Over and over again, some firms will decide the next-period output according to the LMA rule and some others according to the GD rule. It is worth noting that despite the heterogeneity and the oscillations in the fraction of adopters of the two behavioral rules, the average level of production is close to that experienced at the Nash equilibrium, and so are the average profits; see again Figure 12.3.

Increasing firms' propensity to chase the best relative-performing rule, i.e. increasing the intensity-of-choice parameter β in (12.4), heterogeneity in the choice of the behavioral rules persists. In addition, the dynamics of the model get more complicated. Numerical simulations

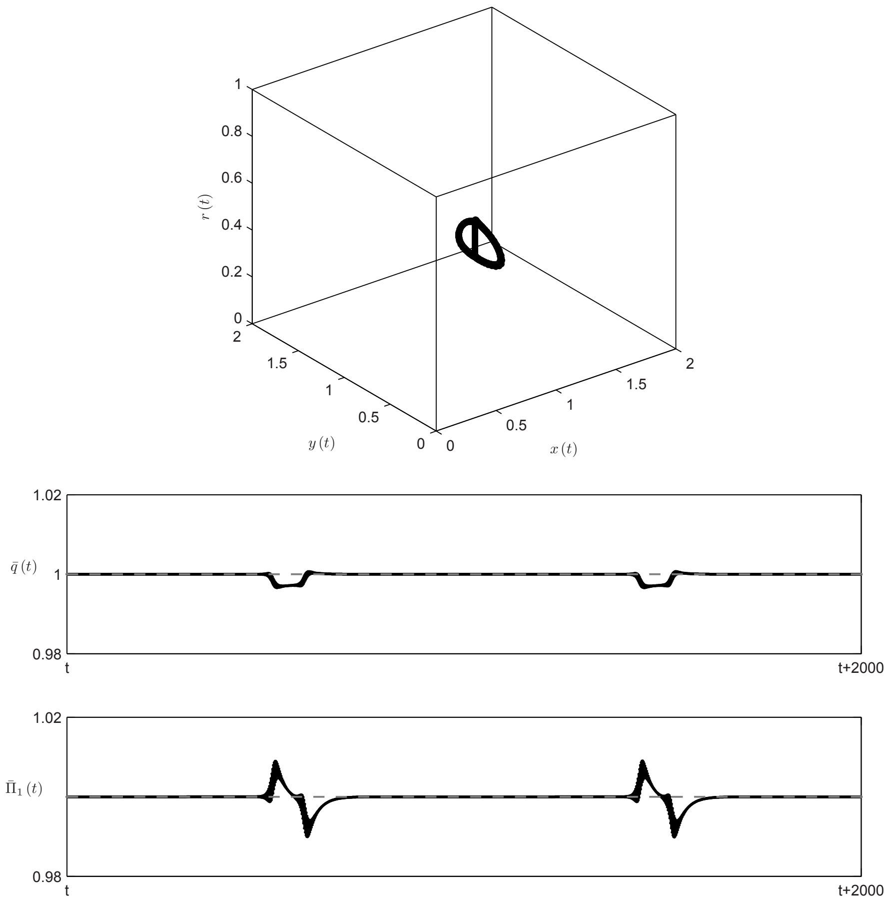


Note: Parameters: $N = 6$, $c = 0.1$, $\beta = 1$, $K_{GD} - K_{LMA} = 0.0001$. The initial conditions and the meaning of the pictures are the same as in Figure 12.2.

Figure 12.3 Dynamics of model 12.12

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

show that the two-cycle undergoes a Neimark-Sacker bifurcation and two stable invariant curves appear. Increasing β further, the two invariant curves have a contact and a strange attractor appears; see Figure 12.4. Even more interesting is the level of production and the profits that firms experience in these scenarios. For example, when $\beta = 10$ we observe that the average level of profit of a single firm (computed as simple moving average with 80 lags of the average profit of the industry) is higher than the profit at Nash equilibrium (at least at regular time windows); see again Figure 12.4. This situation is even more marked when the marginal cost and the number of firms increases. These examples testify how complicated dynamics and *evolutionary stable heterogeneity* can improve the overall performance of an industry. The interested reader can find more details about the dynamics of this evolutionary oligopoly model in Cerboni Baiardi et al. (2015).



Note: Parameters: $N = 6$, $c = 0.1$, $\beta = 10$, $K_{GD} - K_{LMA} = 0.0001$. The initial conditions and the meaning of the pictures are the same as in Figure 12.2.

Figure 12.4 Dynamics of model 12.12

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

5 CONCLUSIONS

In this chapter, we have presented an overview of oligopoly models with boundedly rational firms that choose the behavioral rule to adopt on the basis of relative past performances, i.e. through evolutionary processes. In this regard, we have described an evolutionary framework to model oligopolistic markets where a population of firms can employ different heuristics for deciding their next-period production plans. In this general evolutionary oligopoly game, the time evolution of the fractions of firms adopting the different adaptive behavioral rules is simulated by a profit-driven switching mechanism based on the well-known replicator dynamics. The only assumption made for the behavioral rules is that the dynamical system defined through the heuristics should admit the symmetric Nash equilibrium as steady state. Despite the generality of the model, some analytical results can be provided regarding its dynamics and the stability of its equilibria.

Nevertheless, to undertake a deep analytical and numerical investigation of the dynamics of the model we need to specify the type of heuristics that the firms can choose. Thus, after having underlined some dynamical properties of the general model, we have considered two different cases in detail. In each of these two cases two particular heuristics have been considered. In the first, we chose the best reply (BR) dynamics with naive expectations vs the local monopolistic approximation (LMA). In the second, the gradient dynamics (GD) vs – again – the LMA. All these three behavioral rules model boundedly rational adaptive adjustments that admit the same Nash equilibrium as the fixed point. In other words, all the adaptive heuristics considered, even if governed by local (or myopic) decision rules of boundedly rational and heterogeneous agents, may converge in the long run to a rational equilibrium, i.e. the same equilibrium forecast (and reached in one shot) under the assumption of full rationality and complete information of all economic agents. When a Nash equilibrium is reached, then the outcome of the long-run share of behavioral rules within the population of firms is simple: the cheapest behavioral rule is preferred over the other(s), so that the poorly performing ones disappear in the long run, thus giving rise to a homogeneous oligopoly, with all firms following the same heuristic. Despite their simple properties, when convergence to the Nash equilibrium is not achieved, these models reveal quite rich dynamic behaviors even from the point of view of the time evolution of the distribution of the behavioral rules. In fact, when convergence to a Nash equilibrium does not occur, attractors located on the invariant planes, where all agents adopt the same behavioral rule, can coexist, and also attractors in the interior of the three-dimensional phase space can appear. The latter attractors represent cases where heterogeneous heuristics coexist within the population of firms, thus providing an interesting example of *evolutionary stable heterogeneity*, i.e. polymorphic population states are convenient over time. These attractors are characterized by very complicated dynamics that are not discussed in detail in this chapter. For the sake of completeness, and with the hope of attracting the attention of the reader, we conclude with a brief description of these attractors and related complex phenomena. They are generally characterized by cyclic, quasi-periodic or chaotic dynamics, each with its own basin of attraction. So, two kinds of complexity can be observed: one related to the kind of attractors, the other related to the topological structure of the basins' boundaries. Moreover, even more complex situations can be obtained for the attractors embedded inside the invariant planes where homogeneous behavior take place. In fact, these chaotic two-dimensional invariant sets inside a three-dimensional phase space can be transversely stable on average, thus giving rise to weaker attractors in the Milnor (1985)

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

sense and, consequently, to on–off intermittency phenomena or riddled basins. The interested reader is referred to Milnor (1985), Alexander et al. (1992), Bischi, Gardini, and Stefanini (1998), Bischi, Cerboni Baiardi, and Radi (2015). The study of transverse stability of these "pure strategy" attractors provides useful information about the fate of small mutations, i.e. if the introduction of a different heuristic followed by just one firm (or a few firms) in an oligopoly market, will give rise to a spread of it inside the population or will die out spontaneously; see e.g. Cerboni Baiardi et al. (2015) and Bischi, Lamantia, and Radi (2015).

REFERENCES

Agiza, H.N. and A.A. Elsadany. Nonlinear dynamics in the Cournot duopoly game with heterogeneous players. *Physica A*, 320:512–524, 2003.

Agiza, H.N. and A.A. Elsadany. Chaotic dynamics in nonlinear duopoly game with heterogeneous players. *Applied Mathematics and Computation*, 149(3):843–860, 2004.

Agliari, A., L. Gardini, and T. Puu. Global bifurcations in duopoly when the Cournot point is destabilized via a subcritical Neimark bifurcation. *International Game Theory Review*, 8(1):1–20, 2006.

Alexander, J.C., J.A. Yorke, Z. You, and I. Kan. Riddled basins. *International Journal of Bifurcation and Chaos*, 2 (4):795–813, 1992.

Angelini, N., R. Dieci, and F. Nardini. Bifurcation analysis of a dynamic duopoly model with heterogeneous costs and behavioural rules. *Mathematics and Computers in Simulation*, 79:3179–3196, 2009.

Anufriev, M., D. Kopányi, and J. Tuinstra. Learning cycles in Bertrand competition with differentiated commodities and competing learning rules. *Journal of Economic Dynamics & Control*, 37(12):2562–2581, 2013.

Bischi, G.I. and F. Lamantia. Harvesting dynamics with protected and unprotected areas. *Journal of Economic Behavior and Organization*, 62:348–370, 2007.

Bischi, G.I. and A. Naimzada. Global analysis of a dynamic duopoly game with bounded rationality. In *Annals of the International Society of Dynamic Games, Vol. 5,* 361–385. Basel: Birkhauser, 1999.

Bischi, G.I., L. Cerboni Baiardi, and D. Radi. On a discrete-time model with replicator dynamics in renewable resource exploitation. *Journal of Difference Equations and Applications*, 21(10):954–973, 2015.

Bischi, G.I., C. Chiarella, M. Kopel, and F. Szidarovszky. *Nonlinear Oligopolies: Stability and Bifurcations*. Berlin: Springer-Verlag, 2010.

Bischi, G.I., H. Dawid, and M. Kopel. Gaining the competitive edge using internal and external spillovers: A dynamic analysis. *Journal of Economic Dynamic & Control*, 27:2171–2193, 2003a.

Bischi, G.I., H. Dawid, and M. Kopel. Spillover effects and the evolution of firm clusters. *Journal of Economic Behavior and Organization*, 50:47–75, 2003b.

Bischi, G.I., L. Gardini, and L. Stefanini. Synchronization, intermittency and critical curves in a duopoly game. *Mathematics and Computers in Simulation*, 44(6):559–585, 1998.

Bischi, G.I., F. Lamantia, and D. Radi. Multi-species exploitation with evolutionary switching of harvesting strategies. *Natural Resource Modeling*, 26(4):546–571, 2013a.

Bischi, G.I., F. Lamantia, and D. Radi. A prey-predator model with endogenous harvesting strategy switching. *Applied Mathematics and Computation*, 219(20):10123–10142, 2013b.

Bischi, G.I., F. Lamantia, and D. Radi. An evolutionary Cournot model with limited market knowledge. *Journal of Economic Behavior & Organization*, 116:219–238, 2015.

Bischi, G.I., F. Lamantia, and L. Sbragia. Competition and cooperation in natural resources exploitation: An evolutionary game approach. In: C. Carraro and V. Fragnelli (eds.), *Game Practice and the Environment*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing. 2009a.

Bischi, G.I., F. Lamantia, and L. Sbragia. Strategic interaction and imitation dynamics in patch differentiated exploitation of fisheries. *Ecological Complexity*, 6:353–362, 2009b.

Bischi, G.I., A.K. Naimzada, and L. Sbragia. Oligopoly games with local monopolistic approximation. *Journal of Economic Behavior & Organization*, 62(3):371–388, 2007.

Brock, W.A. and C.H. Hommes. A rational route to randomness. *Econometrica: Journal of the Econometric Society*, 65(5):1059–1095, 1997.

Brock, W.A. and C.H. Hommes. Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics and Control*, 22:1235–1274, 1998.

Cabrales, A. and J. Sobel. On the limit points of discrete selection dynamics. *Journal of Economic Theory*, 57(2): 407–419, 1992.

Cavalli, F., A. Naimzada, and M. Pireddu. Effect of size, composition, and evolutionary pressure in heterogeneous Cournot oligopolies with best response decisional mechanisms. *Discrete Dynamics in Nature and Society*, 2015.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Cavalli, F., A. Naimzada, and F. Tramontana. Nonlinear dynamics and global analysis of a heterogeneous Cournot duopoly with a local monopolistic approach versus a gradient rule with endogenous reactivity. *Communications in Nonlinear Science and Numerical Simulation*, 23(1–3):245–262, 2015.
Cerboni Baiardi, L., F. Lamantia, and D. Radi. Evolutionary competition between boundedly rational behavioral rules in oligopoly games. *Chaos, Solitons & Fractals*, 79:204–225, 2015.
Chiarella, C. and X. He. Asset price and wealth dynamics under heterogeneous expectations. *Quantitative Finance*, 1(5):509–526, 2001.
Chiarella, C., R. Dieci, and L. Gardini. Speculative behaviour and complex asset price dynamics: A global analysis. *Journal of Economic Behavior & Organization*, 49(2):173–197, 2002.
Corchon, L.C. and A. Mas-Colell. On the stability of best reply and gradient systems with applications to imperfectly competitive models. *Economics Letters*, 51(1):59–65, 1996.
Cournot, A. *Recherches sur les Principes Mathematiques de la Theorie des Richesses*. Paris: Chez L. Hachette, 1938.
De Giovanni, D. and F. Lamantia. Control delegation, information and beliefs in evolutionary oligopolies. *Journal of Evolutionary Economics*, 26(5): 1089–1116, 2016a.
De Giovanni, D. and F. Lamantia. Evolutionary dynamics of a duopoly game with strategic delegation and isoelastic demand. *Journal of Evolutionary Economics*, 2016b. Available at https://link.springer.com/article/10.1007/s00191-017-0501-0.
Den Haan, W.J. The importance of the number of different agents in a heterogeneous asset-pricing model. *Journal of Economic Dynamic & Control*, 25:721–746, 2001.
Droste, E., C.H. Hommes, and J. Tuinstra. Endogenous fluctuations under evolutionary pressure in Cournot competition. *Games and Economic Behavior*, 40(2):232–269, 2002.
Dubiel-Teleszynski, T. Nonlinear dynamics in a heterogeneous duopoly game with adjusting players and diseconomies of scale. *Communications in Nonlinear Science and Numerical Simulation*, 16(1):296–308, 2011.
Fisher, F.M. The stability of the Cournot oligopoly solution: The effect of speeds of adjustment and increasing marginal costs. *The Review of Economic Studies*, 28(2):125–135, 1961.
Hahn, F.H. The stability of the Cournot oligopoly solution. *The Review of Economic Studies*, 29(4):329–331, 1962.
Heifetz, A., C. Shannon, and Y. Spiegel. What to maximize if you must. *Journal of Economic Theory*, 133(1):31–57, 2007.
Hofbauer, J. and K. Sigmund. *Evolutionary Games and Population Dynamics*. Cambridge, UK: Cambridge University Press. 1998.
Hofbauer, J. and K. Sigmund. Evolutionary game dynamics. *Bulletin (New Series) of the American Mathematical Society*, 40(4):479–519, 2003.
Hofbauer, J. and J.W. Weibull. Evolutionary selection against dominated strategies. *Journal of Economic Theory*, 71(2):558–573, 1996.
Hommes, C.H. *Behavioral Rationality and Heterogeneous Expectations in Complex Economic Systems*. Cambridge, UK: Cambridge University Press, 2013.
Hommes, C.H., T. Kiseleva, Y. Kuznetsov, and M. Verbic. Is more memory in evoluionary selection (de)stabilizing? *Macroeconomic Dynamics*, 16:335–357, 2012.
Hommes, C.H., M.I. Ochea, and J. Tuinstra. On the stability of the Cournot equilibrium: An evolutionary approach. Working paper, 2011.
Kenneth, J.A. and L. Hurwicz. Stability of the gradient process in *n*-person games. *Journal of the Society for Industrial and Applied Mathematics*, 8(2):280–294, 1960.
Königstein, M. and W. Müller. Why firms should care for customers. *Economics Letters*, 72(1):47–52, 2001.
Kopel, M. Simple and complex adjustment dynamics in Cournot duopoly models. *Chaos, Solitons & Fractals*, 7: 2031–2048, 1996.
Kopel, M., F. Lamantia, and F. Szidarovszky. Evolutionary competition in a mixed market with socially concerned firms. *Journal of Economic Dynamics & Control*, 48:394–409, 2014.
Lamantia, F. A nonlinear duopoly with efficient production-capacity levels. *Computational Economics*, 38(3): 295–309, 2011.
Lamantia, F. and D. Radi. Exploitation of renewable resources with differentiated technologies: An evolutionary analysis. *Mathematics and Computers in Simulation*, 108:155–174, 2015.
Lambertini, L. Oligopoly with hyperbolic demand: A differential game approach. *Journal of Optimization Theory and Applications*, 145(1):108–119, 2010.
Leonard, D. and K. Nishimura. Nonlinear dynamics in the Cournot model without full information. *Annals of Operations Research*, 89: 163–173, 1999.
McManus, M. and R.E. Quandt. Comments on the stability of the Cournot oligipoly model. *The Review of Economic Studies*, 28(2):136–139, 1961.
Medio, A. and M. Lines. *Nonlinear Dynamics*. Cambridge, UK: Cambridge University Press, 2001.
Milnor, J. On the concept of attractor: Correction and remarks. *Communications in Mathematical Physics*, 102(3): 517–519, 1985.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Naimzada, A.K. and L. Sbragia. Oligopoly games with nonlinear demand and cost functions: Two boundedly rational adjustment processes. *Chaos, Solitons & Fractals*, 29(3):707–722, 2006.

Nash, J.F. Equilibrium points in N-person games. *Proceedings of the National Academy of Sciences*, 36(1):48–49, 1950.

Plank, M. Some qualitative differences between the replicator dynamics of two player and *n* player games. *Nonlinear Analysis: Theory, Methods & Applications*, 30(3):1411–1417, 1997.

Puu, T. Chaos in duopoly pricing. *Chaos, Solitons & Fractals*, 1(6):573–581, 1991.

Radi, D. Walrasian versus Cournot behavior in an oligopoly of boundedly rational firms. *R&R in Journal of Evolutionary Economics*, 27(5):933–961, 2017.

Rand, D. Exotic phenomena in games and duopoly models. *Journal of Mathematical Economics*, 5:173–184, 1978.

Rhode, P. and M. Stegeman. Non-Nash equilibria of Darwinian dynamics with applications to duopoly. *International Journal of Industrial Organization*, 19:3–4, 2001.

Rosen, J.B. Existence and uniqueness of equilibrium points for concave N-person games. *Econometrica*, 33(3): 520–534, 1965.

Schaffer, M.E. Are profit-maximisers the best survivors?: A Darwinian model of economic natural selection. *Journal of Economic Behavior & Organization*, 12:29–45, 1989.

Silvestre, J. A model of general equilibrium with monopolistic behavior. *Journal of Economic Theory*, 16:425–442, 1977.

Taylor, P. and L. Jonker. Evolutionarily stable strategies and game dynamics. *Mathematical Biosciences*, 40(1): 145–156, 1978.

Theocharis, R.D. On the stability of the Cournot solution on the oligopoly problem. *Review of Economic Studies*, 27 (2):133–134, 1960.

Tramontana, F. Heterogeneous duopoly with isoelastic demand function. *Economic Modelling*, 27(1):350–357, 2010.

Tramontana, F., L. Gardini, and T. Puu. Global bifurcations in a piecewise-smooth Cournot duopoly game. *Chaos, Solitons & Fractals*, 43(1–12):15–24, 2010.

Tuinstra, J. A price adjustment process in a model of monopolistic competition. *International Game Theory Review*, 6(3):417–442, 2004.

Vallée, T. and M. Yildizoglu. Convergence in the finite Cournot oligipoly with social and individual learning. *Journal of Economic Behavior & Organization*, 72(2):670–690, 2009.

Varian, H.L. *Microeconomics Analysis*. New York: W.W. Norton & Company, Inc., 1992.

Vega-Redondo, F. *Evolution, Games, and Economic Behaviour*. Oxford: Oxford University Press, 1996.

Vega-Redondo, F. The evolution of Walrasian behavior. *Econometrica*, 65(2):375–384, 1997.

Vriend, N.J. An illustration of the essential difference between individual and social learning, and its consequences for computational analyses. *Journal of Economic Dynamics & Control*, 24(1):1–19, 2000.

Weibull, J. *Evolutionary Game Theory*. Cambridge, MA: MIT Press, 1997.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

PART III

GAMES OF COLLUSION IN INDUSTRIAL ORGANIZATION

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

13. Coalitions and networks in oligopolies

Francis Bloch

1 INTRODUCTION

This chapter discusses the formation of coalitions and networks in oligopolies. It weaves together a literature in game theory on cooperation and a literature in industrial organization on the formation of groups of oligopolistic firms. The literature on coalitions in oligopolies started with Stigler's discussion of cartel instability in the 1950s. It was particularly active in the late 1980s and the 1990s, spurred on by new regulations on cooperative research, new merger policies and the emergence of "co-opetition", to use a phrase coined by Brandenburger and Nalebuff (2011), as a new form of interaction where oligopolistic firms cooperate on some dimensions and compete on others. At the same time, a number of empirical studies on collaborative projects established the prevalence of new forms of cooperation across oligopolistic firms.

From its inception, the literature on cooperation in oligopolies has made use of solution concepts derived from game theory. The initial solution concepts of cartel stability were borrowed from the literature on cooperative games, focusing on notions of internal and external stability dating back from von Neumann and Morgenstern (1944). Gradually, these cooperative solution concepts gave way to equilibrium outcomes of non-cooperative games – either simultaneous or sequential – providing strategic foundations for cooperation. More recently, the emphasis has been placed on bilateral rather than multilateral cooperation, leading to the development of new models of network formation, which replace models of coalition formation.

Our discussion of cooperation in oligopolies starts with a brief presentation of the game-theoretic models used to predict the formation of coalitions and networks. We then consider two different forms of cooperation. We start by analyzing collusion, discussing the formation of cartels and horizontal mergers in oligopolies. The last part of the chapter is devoted to the analysis of strategic alliances, which encompass both research joint ventures and information exchange platforms.

The literature on cooperation in oligopolies covers a large fraction of theoretical research in industrial organization, and it is, of course, impossible to cover it all in one chapter. We have chosen to be very selective, restricting attention to theoretical models that aim at predicting whether cooperation will arise, and what sizes of groups and architecture of networks are likely to emerge among oligopolistic firms. Due to lack of space, we cannot cover the empirical literature on collusion and collaboration, the theoretical literature on supergames and tacit collusion, or the more recent theoretical literature on the interaction between antitrust policies and cooperation decisions.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

2 MODELS OF COALITION AND NETWORK FORMATION

In this section, we review models of coalition and network formation developed in game theory that have been applied to industrial organization. The models fall into three categories: simultaneous models of coalition formation that are played in one shot and where all players announce the coalitions they want to form; dynamic games of coalition formation, where coalitions are formed sequentially through extensive form games; and models of network formation where pairs of players form bilateral links. We start with some simple notations. Let $N = \{1, 2., , , n\}$ be a set of players with typical element i and $2^N \setminus \emptyset$ the set of all non-empty coalitions of players with typical element C. A partition π is a collection of coalitions that are disjoint, non-empty and cover the entire set N. We denote by Π the set of all possible partitions. We also consider the set G of all undirected networks over n players. An undirected network can be identified with a symmetric $n \times n$ matrix of $0, 1$ with $g_{ij} = 1$ if and only if there is a link between i and j. This formulation shows that there are $2^{\frac{n(n-1)}{2}}$ networks on the set of players. We let $ij \in g$ denote the fact that players i and j are directly linked in network g.

2.1 Open Membership Games

In the open membership game, members cannot prevent other players from joining a coalition. In the simplest model considered by d'Aspremont et al. (1983), all players select a strategy in $S = \{0, 1\}$. All players who announce 1 belong to the coalition, and all players who announce 0 remain independent. Hence $C = \{i | s_i = 1\}$ and the partition formed is $\pi = \{C, \{j\}_{j \notin C}\}$. Following d'Aspremont et al. (1983), we say that a coalition is *internally stable* if no player wants to leave the coalition and *externally stable* if no player wants to join the coalition. A coalition that is internally and externally stable is called *stable*. Alternatively, we can define a stable coalition as the Nash equilibrium outcome of the open membership game.

The cartel formation game of d'Aspremont et al. (1983) only allows for one coalition to form. A natural extension of this game is the address game discussed by Yi (1997). Players choose addresses in the state $S = \{0, a_1, .., a_n\}$. All players who choose the same address a_j form the coalition C_j. All players who choose 0 remain independent. The partition formed is thus $\pi = \{C_1, ..C_J, \{k\}_{k \notin C_j \forall j}\}$.

2.2 Exclusive Membership Games

In exclusive membership games, players announce the coalition they want to form and can thus prevent the entry of other members into the coalition. The earliest game of coalition formation was proposed by von Neumann and Morgenstern (1944), (pp. 243–244). Each player i announces a coalition C_i to which she wants to belong. The outcome function assigns to any vector of announcements $C_1, \ldots, C_n$, a coalition structure π as follows: $C \neq \{i\} \in \pi$ if and only if, for all agents $i \in C$, $C_i = S$. A singleton i belongs to the coalition structure π if and only if either $C_i = \{i\}$ or $C_i = C$ and there exist $j \in C$ such that $C_j \neq C$. In this procedure, a coalition is formed if and only if all its members unanimously agree to form the coalition.

This procedure was rediscovered by Hart and Kurz (1983), who labeled it "model γ". They contrast it with another procedure, labeled "model δ", where unanimity is not required for a coalition to form. In the δ procedure, the outcome function assigns to any vector of

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

announcements $C_1, \ldots, C_n$, a coalition structure π where: $C \in \pi$ if and only if $C_i = C_j \supseteq C$ for all $i, j \in C$. In other words, coalitions are formed by any subset C of players who coordinate and announce the same coalition C_i. In this procedure, the announcement serves to coordinate the actions of the players, and indicates what is the largest coalition that players are willing to form.

2.3 Bidding Game

The bidding game of coalition formation was proposed by Kamien and Zang (1990). Every agent i submits a vector of bids, b^i_j over all agents j in N. A bid b^i_j for $i \neq j$ is interpreted as the amount of money that agent i is willing to pay to acquire the resources of agent j. The bid b^i_i is interpreted as the asking price at which agent i is willing to sell her resources. Given a matrix $B = \left[b^j_i\right]$ of non-negative bids, one can assign the resources of every agent i either to another agent j or to agent i herself, if she remains a singleton. Formally, let

$$S(i) = \left\{j \in N, j \neq i, b^j_i \geq b^k_i \forall k \neq j\right\}$$

denote the set of players other than i such that (i) the bid they offer is no smaller than the bid of any other player and (ii) the bid they offer is higher than the asking price. If $S(i)$ is a singleton, the assignment of the resources of player i to the unique player in $S(j)$ (and hence the formation of a coalition S containing $\{i, j\}$) is immediate. If $S(i)$ is not a singleton, one needs to define an exogenous tie-breaking rule to assign the resources of player i to some member of $S(i)$. As a result of this bidding procedure, resources of some players are bought by other players, resulting both in the formation of a coalition structure π and in transfers across players given by $t^j_i = b^j_i$ and $t^i_j = -b^j_i$ if player j acquires the resources of player i. Pérez-Castrillo (1994) independently proposed a procedure of coalition formation that bears a close resemblance to Kamien and Zang's (1990) bidding game. The main difference is that Pérez-Castrillo introduces competitive outside players (the "coalition developers") who simultaneously bid for the resources of the players.

2.4 Sequential Formation of Coalitions

Sequential games of coalition formation are based on Rubinstein's (1982) model of alternative offers bargaining. As in Rubinstein's (1982) model, the representative model has an infinite horizon, players discount future payoffs, and at each period in time, one of the players (the proposer) makes an offer to other players (the respondents) who must approve or reject the proposal.

Different variants of this scenario have been proposed. Chatterjee et al. (1993) propose a rejector–proposer version. Players are ordered according to an exogenous protocol. At the initial stage, player 1 chooses a coalition C to which she belongs and a vector of payoffs for all members of C, $\mathbf{x}_C$ satisfying $\sum_{i \in C} x_i = v(C)$, where $v(C)$ describes the coalitional surplus of the coalition C. Players in C then respond sequentially to the offer. If all accept the offer, the coalition C is formed, and the payoff vector $\mathbf{x}_C$ is implemented. The first player in $N \backslash C$ is chosen as proposer with no lapse of time. If one of the players in C rejects the offer, one period

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

elapses and the rejector becomes the proposer at the following period. Okada (1996) analyzes a coalitional bargaining game where the proposer is selected at random after every rejection.

In the context of coalition formation, payoffs depend on the entire coalition structure, and underlying gains from cooperation depend on the coalitions formed by other players. In this context, Bloch (1996) proposes a coalitional bargaining game capturing this forward-looking behavior when the division of the surplus across coalition members is fixed. At any stage player i announces a coalition C_i that she wants to form. If all players in C_i agree, the coalition is formed and the next player is chosen to make a proposal. If one of the members of C_i rejects the proposal, she becomes the proposer next period. Consider symmetric games where payoffs only depend on the size distribution of coalitions. In that case, the equilibrium coalition structures of the infinite horizon bargaining game can be computed by using the following finite procedure. Let players be ordered exogenously. The first player announces an integer k_1, corresponding to the size of the coalition she wants to form. Player $k_1 + 1$ then announces the size k_2 of the second coalition formed. The game ends when all players have formed coalitions, i.e. $\sum k_t = n$.

While Bloch (1996) assumes that the division rule of the surplus is fixed, Ray and Vohra (1999) consider a model of coalitional bargaining with externalities, where the division of coalitional surplus is endogenous, and payoffs are represented by an underlying game in partition function form. Ray and Vohra (1999) first establish the existence of stationary equilibria in mixed strategies, where the only source of mixing is the probabilistic choice of a coalition by each proposer. Their main theorem establishes an equivalence between equilibrium outcomes of the game and the result of a recursive algorithm. This algorithm, in four steps, characterizes equilibrium coalition structures for symmetric games. It can easily be implemented on computers and has been successfully applied in Ray and Vohra (2001) to study the provision of pure public goods.

2.5 Successive Formation of Coalitions

In successive games of coalition formation, players meet in pairs and decide whether to merge. If the players agree on a merger, one of the players acquires the resources of the other, and forms a single entity that continues to take part in the process. In successive games of coalition formation, coalitions are thus formed by successive acquisition of the resources of the other players. Gul (1989) proposed the first game of successive formation of coalitions and showed that the equilibrium payoff converges to the Shapley value of the underlying cooperative game. The set of active players in the game varies over time, as the resources of players are acquired by other players. At any period, a pair of active players is selected and one of the players is chosen at random to make a take-it-or-leave-it offer to acquire the resources of the other player. If the offer is accepted, the set of active players is reduced by one (the player whose resources have been acquired) and the process continues. If the offer is rejected, the set of active players does not change, one period of time elapses, and a new pair is chosen.

2.6 Formation of Networks

We now consider the formation of links in networks. Myerson (1991) proposed a game of undirected network formation that is very similar to models γ and δ. Agents simultaneously announce the set of agents with whom they want to form links. Hence, a pure strategy in the

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

game is a subset $C_i \subseteq N \setminus \{i\}$ for every agent i. The formation of a link requires *consent* by both parties. Link ij is formed if and only if $i \in C_j$ and $j \in C_i$. We let $Y_i(g)$ denote the value of player i in network g.

Given the typical indeterminacy of Nash equilibrium in models of undirected networks, it is not surprising that other equilibrium notions have been considered in the literature. These equilibrium refinements allow for some cooperation among players. Because it takes agreement of both players i and j to form the link ij, it is natural to consider coalitions of size two since this is the minimal departure from a purely non-cooperative equilibrium concept. Jackson and Wolinsky (1996) specify a very weak notion of stability for networks:

A network g is *pairwise stable* if for all $i, j \in N$:

(i) $Y_i(g) \geq Y_i(g - ij)$;
(ii) $Y_i(g + ij) > Y_i(g)$ implies that $Y_j(g + ij) < Y_j(g)$.

This concept of stability is very weak because it restricts deviations to change only *one* link at a time – either some agent can delete a link or a pair of agents can add the link between them. This notion of stability is not based on any specific procedure of network formation. A stronger concept of stability based on bilateral deviations uses Myerson's network formation game: a pairwise Nash equilibrium is a Nash equilibrium if it is a Nash equilibrium of the Myerson game that is immune to the formation of a new link by a pair of players. In a pairwise Nash equilibrium, players can delete any subset of links, and pairs of players can coordinate on the formation of a new link.

3 CARTELS AND MERGERS

3.1 Cartel Formation in Cournot Oligopolies

In Cournot oligopolies, the formation of a cartel leads its members to reduce quantities in order to increase the selling price. This provides a public good to firms that do not belong to the cartel – the outsiders, who benefit from the price increase without paying the cost of a limitation in quantities. Hence firms may be reluctant to form or join cartels, resulting in a "puzzle" of cartel formation that was first noted by George Stigler (1950, pp. 25–26) in his discussion of mergers:

> The major difficulty in forming a merger is that it is more profitable to be outside a merger than to be a participant. The outsider sells at the same price but at a much larger output at which marginal cost equals price. Hence the promoter of a merger is likely to receive much encouragement from each firm – almost every encouragement in fact except participation.

The "Stigler effect" can easily be observed in a linear Cournot oligopoly. Let n firms on the market, with zero marginal cost, produce homogeneous products with a linear demand $P = 1 - Q$. The profit of each firm only depends on the number of *active* firms on the market and is given by $R = \frac{1}{(n+1)^2}$. Now suppose that a cartel of size k forms on the market, with the remaining $n - k$ firms remaining independent. The total number of active firms in the market

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

reduces to $n - k + 1$. As cartel members share equally the profit of the cartel, the profit of an insider is $R^i(k) = \frac{1}{k(n-k+2)^2}$, whereas the profit of an outsider is $R^o(k) = \frac{1}{(n-k+2)^2}$.

We immediately observe that the profit of an outsider is always greater than the profit of an insider: $R^o(k) > R^i(k)$ for all k. Following d'Aspremont et al. (1983), a cartel of size k is stable if $R^i(k) \geq R^o(k-1)$ (internal stability) and $R^o(k) \geq R^i(k+1)$ (external stability). These conditions amount to

$$(n-k+3)^2 \geq k(n-k+2)^2,$$
$$(k+1)(n-k+1)^2 \geq (n-k+2)^2$$

It is easy to check that the first inequality (internal stability) cannot be satisfied for $k \geq 2$. Hence, the only stable partition is the partition of singletons, where no cartel is formed and all firms remain independent. This simple computation suggests that free riding in the formation of cartels is so strong as to prevent the formation of cartels or mergers on any market. A closer inspection of the profits of insiders and outsiders, due to Salant, Switzer and Reynolds (1983), shows that R^o is increasing in k but R^i is non-monotonic in k and assumes a U-shape, first decreasing, then increasing in k. The minimal profitable cartel size is defined as the unique value of k for which $R^i(k) = R^i(1)$, namely the solution to the equation

$$(n+1)^2 = k(n-k+2)^2,$$

giving $k^* = \frac{2n+3-\sqrt{4n+5}}{2n}$, or around 80 percent of the size of the market. The observation made by Salant et al. (1983) is thus that mergers must involve a very large fraction of the firms in the industry to become profitable.

The computation of the minimal profitable cartel size also has important implications for the study of exclusive membership and sequential models of coalition formation. In the γ game, when a firm leaves the cartel, the cartel dissolves. Hence any cartel of size $k \geq k^*$ is an equilibrium outcome, because no player wants to deviate from the cartel and obtain $R^o(1)$ instead of $R^i(k)$. By contrast, in the δ model, when a firm leaves the cartel, other cartel members remain together, so that the deviating firm compares $R^i(k)$ with $R^o(k-1)$ and always has an incentive to deviate: no cartel of a size greater than one can be formed at equilibrium. In the model of sequential coalition formation, Bloch (1996) and Ray and Vohra (1999) show that the only subgame equilibrium outcome is for the first firms to remain outsiders, and the last k^* firms to agree to form a coalition. Hence, in equilibrium, the minimum profitable cartel size is formed. Macho-Stadler, Pérez-Castrillo, and Porteiro (2006) use the successive coalition formation model where firms meet bilaterally and decide whether or not to merge. They show that the equilibrium outcome is either that all firms remain singletons or that they merge into a single coalition. Merger to monopoly arises for a specific region of the parameters that is described by a complex recursive formula. Mauleon and Vannetelbosch (2004) explore the formation of cartels in the linear Cournot oligopoly using a farsighted solution concept: the largest consistent set of Chwe (1994). They observe that any coalition structure where the size of the cartel is above the minimal profitable cartel size belongs to the largest consistent set. However, other coalition structures can also be sustained, including some involving the formation of multiple cartels.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Kamien and Zang (1990, 1991) propose a different approach to the study of horizontal mergers. In their acquisition game, firms announce a bidding price for the assets of all other firms and an asking price for their own assets. They observe that no merger will arise at equilibrium. To understand this point, notice that if a firm forms a cartel of size k, it must compensate all $k-1$ firms for their participation in the cartel at a price $\pi^{\circ}(k-1)$. Now clearly, $k\pi^{i}(k) < (k-1)\pi^{o}(k-1)$, so that the cartel cannot profitably acquire $k-1$ other firms at their asking price $\pi^{\circ}(k-1)$. This line of reasoning is reminiscent of a classical argument on the difficulty of successful take-overs in the corporate finance literature, when an investor must acquire shares from different shareholders of the target firm.

The fact that cartels are inherently unstable, and unlikely to emerge as equilibrium outcomes of a game of coalition formation is a puzzle, as cartels and mergers are indeed observed on many markets. The puzzle can be solved by enriching the model in order to give an advantage to cartels over independent firms. In d'Aspremont et al. (1983), Donsimoni (1985) and Donsimoni, Economides and Polemarchakis (1986), Thoron (1998), Schaffer (1995) and Prokop (1999), cartels are dominant firms fixing prices and independent firms form a competitive fringe, responding in quantity to the price of the dominant firm. In all these papers, cartels are formed by a simultaneous open membership game and a non-trivial stable cartel size exists. Diamantoudi (2005) considers a more general farsighted solution concept, where firms anticipate the sequence of moves following deviations. Using an indirect dominance relation, she shows that von Neumann-Morgenstern stable sets always exist in the cartel game, and singles out the smallest stable cartel as the most appealing prediction in the game of cartel formation. In the same spirit, Kuipers and Olaizola (2008) define a different dynamic process of cartel formation where firms move from one cartel structure to another considering myopic improvements, but moves that can be countered immediately are excluded. With this alternative model of transitions, Kuipers and Olaizola (2008) show that stable cartels are a size larger than the minimal profitable cartel size – a conclusion that stands in sharp contrast to Diamantoudi (2005). Konishi and Lin (1999) generalize the analysis of the Stackelberg game where the cartel chooses its quantity first to arbitrary demand and cost functions. They offer a conjecture on the size of the stable cartel and numerically compute it for small values of n. Recently, Zu, Zhang and Wang (2012) have provided an exact formula for the size of the stable cartel in Konishi and Lin's (1999) model (which results in higher cartel sizes than originally conjectured).

Perry and Porter (1985) assume that costs are quadratic and that firms own capital units that can be recombined after a merger. Hence the merged entity can produce more efficiently than outsiders by distributing production in the plants of the constituent firms. In this model of convex costs, the cartel benefits from a cost advantage over the outsiders, and profitable mergers can form. Farrell and Shapiro (1990) move away from the homogeneous linear Cournot oligopoly and assume a general demand and cost structure. They highlight the fact that cartels benefit from synergies among their members so that the marginal cost of a cartel is lower than the marginal cost of independent firms. Hence, in the presence of synergies, concentration and welfare may move in the same direction, casting a new light on the antitrust treatment of horizontal mergers.

Any other strategic advantage given to cartel members over independent firms would help explain the formation of cartels on actual oligopolistic markets. Brown and Chiang (2002) and Banal-Estañol and Ottaviani (2006) suppose that firms face idiosyncratic shocks in demand and costs. When firms are risk averse, a merger allows firms to diversity risk, and the merged

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

entity has a strategic advantage over independent firms, which leads it to increase production. Hence, mergers occur more frequently when firms are risk averse and the environment more variable. Brown and Chiang (2002) study a three-firm environment and a sequential merger process, and observe that since mergers of two firms are more likely to be profitable, merger to monopoly is less likely to emerge with risk-averse firms. Banal-Estañol and Ottaviani (2006) consider a general environment and characterize the sharing contract among participants to the merger, showing that under Cournot competition, firms have an incentive to equalize the number of shares they possess in each other's firm. Davidson and Ferrett (2007) suppose that cartel members can share the benefit of R&D investments and show that this allows for profitable cartels. Nocke (1999) studies the formation of cartels when firms face capacity constraints. Because cartels have larger capacity, they enjoy a strategic advantage when firms are capacity constrained. Large cartels are easier to sustain when demand is high and the capacities of individual firms are low. Espinosa and Macho-Stadler (2003) incorporate moral hazard in the Cournot model, assuming that production is realized by independent teams. The moral hazard problem becomes more stringent when teams are larger (free-riding incentives are higher) so that, at first glance, cartels are less likely to form. However, as intermediate cartels are unlikely to form, large cartels are easier to sustain – firms realize that by leaving the cartel, they will lead to an unstable intermediate cartel that unravels so that in the end all firms become independent. Horn and Persson (2001) study cartel formation through a cooperative game-theoretic solution concept. They assume that mergers are profitable and show that the most concentrated coalition structure will always emerge: firms form monopolies if monopolies are allowed, duopolies if duopolies are allowed but not monopolies, etc.

When firms have heterogeneous costs, the formation of cartels becomes harder, because the gains from cooperation cannot be divided equally among cartel members. Characterizing conditions under which cartels are profitable, and the equilibrium coalition structure becomes a difficult exercise, and results have only been obtained for small numbers of firms or small numbers of types. Donsimoni (1985) studies cartels among heterogeneous firms in the dominant firm–competitive fringe model when firms have different quadratic costs. Barros (1998) and Brown and Chiang (2003) discuss the formation of cartels in the linear Cournot model among three firms with different costs. Faulí-Oller (2000) analyzes a four-firm model where two firms have low costs and two firms have high costs. When costs are privately known, the cartel must in addition elicit information about costs from the cartel members. Cramton and Palfrey (1990) provide a complete analysis of the mechanism design problem faced by a cartel when firms have to reveal their production costs.

3.2 Cartel Formation in Bertrand and Spatial Oligopolies

The intuition underlying the instability of cartels in Cournot oligopolies is related to the fact that quantities are strategic substitutes: a reduction in quantity by cartel members leads outsiders to expand their own quantity, thereby depressing the price and reducing the profit of cartel members. In a Bertrand oligopoly, prices are strategic complements, and the increase in price resulting from collusion among cartel members leads outsiders in turn to increase their prices, resulting in an equilibrium with higher prices and profits for all cartel members. Deneckere and Davidson (1985) were the first to make this observation in a model of symmetric product differentiation, when firms set prices rather than quantities. They note that both cartel members and outsiders benefit from the formation of a merger, even though

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

outsiders benefit more than insiders. They compute the profit functions $R^i(k)$ and $R^o(k)$ and show that they are both strictly increasing in k. However, it remains true that outsiders obtain higher profits than insiders (or more generally members of smaller cartels obtain higher profits than members of larger cartels), so that the formation of mergers is not guaranteed. Deneckere and Davidson (1985) provide a numerical example to show that firms may be unwilling to merge even under Bertrand competition, but note that, when the degree of product differentiation becomes small, and fierce Bertrand competition erodes the firms' profits, merger to monopoly is obtained as the equilibrium outcome of an open membership game of cartel formation.

The formation of cartels in spatial models has been studied both in the circular city and on the line. Following early work by Levy and Reitzes (1992) and Brito (2003) computes the effect of a merger between two consecutive firms. He shows that the profit of insiders always goes up, providing a positive incentive to merger as in the model of symmetric product differentiation of Deneckere and Davidson (1985). When two consecutive firms merge around the circle, the pricing game is no longer symmetric, and firms' equilibrium prices depend on their proximity to the merged entity. Not surprisingly, firms closer to the merged entity are more affected by the merger, and hence raise their prices more and benefit from a larger increase in profits than firms at a higher distance. However, the effect of the merger ripples through the entire circle, and all firms effectively raise their prices and experience an increase in profit. Giraud-Héraud, Hammoudi and Mokrane (2003) use the same model of a circular city but assume that one of the firms sells products at all locations (the multi-product firm). They analyze the incentives of the multi-product firm to merge with some of its independent rivals. One difficulty that they highlight is that merging firms are no longer ex ante symmetric, and the profitability of the merger depends on the post-merger division of the gains from cooperation.

Studies of mergers on the line have also led to significant insights. Braid (1986) studies mergers between two adjacent stores on an infinite line. He shows that when prices are set simultaneously, collusion among stores only has an effect if the two stores are nearest neighbors, and that affects the prices of all other stores on the infinite line. If the merged entity acts as a Stackelberg leader, merger has an effect even when stores are not adjacent. Braid (1999) builds on this model to study mergers between two stores located on a two-dimensional space and computes numerically the effect on equilibrium prices. On the Hotelling line, Rothschild, Heywood and Monaco (2000) analyze a three-firm model, where two firms have the opportunity to merge. The innovation of their paper is that they consider how the possibility of merger affects the firms' location decisions. They thus consider a three-stage model where firms initially choose locations, then two of the three firms decide whether to merge and finally firms compete in prices. In this model, the two merging firms obtain a higher gain than the outsider. Heywood, Monaco and Rothschild (2001) extend the analysis to n firms, and distinguish between corner cases where the merging firms are at the extremity of the Hotelling segment and interior cases. They show that outsiders are always harmed in corner cases but not in interior cases.

In the context of vertical differentiation, the analysis of mergers has so far been restricted to oligopolies with three firms. Norman, Pepall and Richards (2005) analyze a model where the two merging firms sell the goods of lowest qualities. They show that the merged entity will always choose to sell the good of lowest quality and argue that the post-merger equilibrium may lead to higher market prices. Gabszewicz, Marini and Tarola (2015) analyze general

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

mergers among three competing firms and show that the only stable mergers involve the firms producing the bottom two qualities or the firm producing the high quality and the firm producing the low quality.

We conclude by noting that three papers have attempted to characterize equilibrium coalition structures in an abstract context encompassing both the Cournot and Bertrand games. Currarini and Marini (2006) explore the difference between situations where the competitive game among firms has strategic complements (Bertrand) or substitutes (Cournot). They show that non-trivial coalition structures emerge in games with strategic complements and provide conditions under which non-trivial coalition structures also emerge in games with strategic substitutes. Yi (1997) and Finus and Rundshagen (2009) consider a general model with positive externalities that encompasses mergers and cartels. They obtain interesting results comparing the sizes of cartels formed under different processes of coalition formation.

3.3 Dynamic Mergers

Dynamic models of mergers emphasize the changing environment under which firms interact, the interplay between entry, exit and merger decisions, and the role of repeated interactions on the enforcement of collusion. The seminal model proposed by Gowrisankaran (1999), in the spirit of Ericsson and Pakes (1995), analyzes a dynamic model where firms choose to enter, invest, merge and exit at every period. Each firm evaluates the outcome of its decision based on expected discounted profit calculations, and the equilibrium concept is a Markov perfect equilibrium outcome in a complex environment where the state captures all relevant information about the industry. Equilibrium is shown to exist, and can be computed using numerical techniques. Computations show that the possibility of mergers greatly affects the structure of the industry, reducing the number of active firms in equilibrium. Once mergers are introduced, production, prices and profits go up, but consumer surplus decreases. Gowrisankaran and Holmes (2004) analyze a dynamic model with a dominant firm and a competitive fringe, where, as in Perry and Porter (1985), merging firms can reallocate productive capital to reduce production costs. In the dynamic environment, capital is not given but results from firms' investment decisions. The analysis shows that both perfect competition and monopoly are absorbing states. In some situations, fringe firms acquire the capital of the dominant firm; in others, the dominant firm successively acquires all the capital of the dominant firm, resulting in a monopoly.

Pesendorfer (2005) studies a simpler model of mergers and entry and provides an explicit characterization of Markov perfect equilibria. He assumes that, at every period, a single firm has the opportunity to enter, and that firms make offers to merge as in Kamien and Zang (1990). Firms are identical and profits only depend on the number of active firms every period. In this simple setting, conditions are obtained under which no merger ever takes place, and under which mergers result in monopoly. Mergers may occur because firms anticipate that other mergers will follow – this is the pre-emptive role of mergers. Pesendorfer (2005) establishes the existence of merger cycles, under which k mergers happen at some period, followed by $k - 1$ periods with no mergers. Pre-emptive mergers also occur in Fridolfsson and Stennek (2005) who analyze a three-firm model, where following a shock, firms race to merge with another firm. In their model, firms merge not only to increase prices as in the classical framework but also in order to guarantee that they will not be left out of the wave of mergers.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Mergers have also been analyzed in the context of repeated interaction between oligopolistic firms. When firms have different capacity constraints, mergers allow firms to recombine capacities, and change the environment under which collusive agreements can be enforced. Compte, Jenny and Rey (2002), Vasconcelos (2005) and Kuhn and Motta (2001) analyze different models of collusion with mergers. They focus attention on situations where all firms in the industry (the merged firm and the independent firms) collude. Both Compte et al. (2002) and Kuhn and Motta (2001) observe that collusion is easier to sustain when firms have equal capacities, so that any merger of small firms that leads to an equalization of capacities may help collusion. By contrast, mergers involving large firms may make collusion harder to sustain and hence be procompetitive. In addition, by reducing the number of active firms, a merger helps sustain collusion in a repeated interaction. Vasconcelos (2005) generalizes the analysis by allowing merged firms to recombine capital as in Perry and Porter (1985) and allowing for more general punishment schemes in the repeated game. In all the previous papers, collusion involves all the firms in the industry, By contrast, Bos and Harrington (2010) allow for collusion to involve a subset of firms, and propose a model that combines endogenous cartel formation, enforcement through repeated interaction and asymmetric capacities. They show that a cartel is stable if the smallest firm finds it optimal to be in the cartel and the largest firm finds it optimal to be outside the cartel. This characterization yields a formula to compute stable cartels. A merger of firms under partial collusion produce complex effects, as it simultaneously affects firms outside the cartel and the incentives to collude inside the cartel. After a merger, the set of stable cartels may change, and hence post-merger equilibrium prices and quantities may be difficult to analyze. Bos and Harrington (2010) use numerical computations to evaluate the effects of mergers in their model.

3.4 Bidding Rings

Bidding rings are groups of buyers who submit their bids cooperatively in auctions. Graham and Marshall (1987) and Mailath and Zemsky (1991) have analyzed bidding rings in second price private value auctions. Suppose that values are independently distributed according to a common distribution F with density f. In a second-price auction, the optimal bidding strategy is to bid one's valuation and the expected profit is given by

$$R = \int_0^{\infty} \int_0^{z} (z - y)(n - 1)F(y)^{n-2} f(y) f(z) dy dz,$$

where $(n-1)F(y)^{n-2}f(y)$ is the distribution of the highest bid among $n-1$ bidders.

If a bidding ring of size k forms, the distribution of the highest bid among ring members is $kF(y)^{k-1}f(y)$ and of the highest bid among independent bidders $(n-k)F(y)^{n-k-1}f(y)$. Hence the expected profit of a ring member (insider) is

$$R^i(k) = \frac{1}{k} \int_0^{\infty} \int_0^{z} kF(y)^{k-1} f(y)(n-k)F(y)^{n-k-1} f(y) dy dz,$$

whereas the expected profit of an independent bidder (outsider) remains $R^o(k) = R$. In the special case where the distribution of values is uniform on $[0, 1]$, the profits are

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

$R^i(k) = \frac{1}{(n-k+1)(n+1)}$ and $R^o = \frac{1}{n(n+1)}$. We immediately observe that R^o is independent of k, R^i increasing in k and $R^i(k) > R^o$ for all $k > 1$. Hence the only stable bidding ring is the complete bidding ring including all the bidders. This is also the unique equilibrium outcome of the γ and δ games and of the sequential game of coalition formation. In sharp contrast to the oligopoly case, bidding rings in auctions are always profitable to all the bidders.

Mailath and Zemsky (1991) consider a general situation where values are drawn from different distributions and prove a stronger result. They show that the sum of utilities of bidders is increasing and convex in the size of the bidding ring. Hence, in cooperative game-theoretic terms, the coalitional function is convex, so that the core of the game is non-empty. Even when bidders are heterogeneous, there always exists a distribution of the surplus of the bidding ring that will be accepted by all the bidders.

The analysis of bidding rings in first-price auctions is much more complex, as it requires computing the equilibrium payoff of a first-price auction with asymmetric bidders (the bidding ring and independent bidders), a notoriously complex task. MacAfee and MacMillan (1992) compute equilibrium strategies when values are independent and identically distributed according to a binary distribution, $v = 1$ with probability p and $v = 0$ with probability $1 - p$. In that case, the distribution of values of all bidders is a binomial distribution with parameter p, allowing for simple computations of the distributions of order statistics. The expected profit of a member of the bidding ring is $R^i(k) = \frac{1}{k}(1-p)^{n-k}(1-(1-p)^k)$ whereas the expected profit of an independent bidder is $R^o(k) = p(1-p)^{n-k}$. As opposed to the second-price auction, but in line with the oligopoly models, the profit of an insider is always smaller than the profit of an outsider, $\pi^i(k) < \pi^o(k)$. Stable bidding rings exist when $\pi^i(k) > \pi^\circ(k-1)$, a condition that holds for the unique value k^* such that

$$\frac{1-(1-p)^k}{k} \geq p(1-p) \geq \frac{1-(1-p)^{k+1}}{k+1}.$$

MacAfee and MacMillan (1992) show that k^* is always larger than 3, increasing in p and converges to infinity when p converges to 1.

3.5 Collusive Networks

Collusive networks emerge when firms form reciprocal market-sharing agreements whereby they refrain from entering each other's market. This geographical division of markets has been analyzed by Belleflamme and Bloch (2004). Suppose that firm i is based on market i. By forming a link ij with firm j, firm i refrains from entering market i and firm j refrains from entering market j. For any graph g, $n_i(g) = n - d_i(g)$ is the number of active firms on i's market, where $d_i(g)$ denotes the degree of firm i in the collusive network g. Assuming that firms are symmetric, the profit that each firm makes on market i is given by $\pi(n_i(g))$ and the total profit of firm i is

$$R = \pi(n_i(g)) + \sum_{j|ij\in g} \pi(n_j(g)).$$

Belleflamme and Bloch (2004) characterize pairwise-stable collusive networks when profit functions are decreasing and log-convex in the number of active firms. They show that

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

(i) every stable network must include complete components – when firms sign market-sharing agreements, they must sign them with all other firms in their component – (ii) that components must be of different sizes and (iii) that every component must be of size greater than m^* where m^* is the solution to $\frac{\pi(n-m+1)}{\pi(n-m+2)} = 2$.

In a linear Cournot oligopoly where $\pi(n) = \frac{1}{(n+1)^2}$, it is easy to see that $m^* = n$ so that there are only two candidates for collusive networks: the complete network and the empty network. In a second-price auction with uniform distributions, the expected profit of a bidder is $\pi(n) = \frac{1}{n(n+1)}$ and we find that $m^* = n - 1$ so that there exist three possible stable collusive networks: the empty network, the complete network, and an asymmetric configuration where one independent bidder faces a bidding ring of $n - 1$ bidders.

4 ALLIANCES, RJVs AND TRADE ASSOCIATIONS

4.1 Research Joint Ventures and Alliances

Cost-reducing alliances have been extensively studied in the context of research joint ventures (RJVs). The seminal papers by Katz (1986) and d'Aspremont and Jacquemin (1988) considered the incentives of two firms to cooperate in cost-reducing research before competing on the market. This line of research was prompted by a change in the regulatory environment, with programs aimed at stimulating cooperative research among firms both in the USA and in Europe in the mid-1980s – the National Cooperative Research Act of 1984 and the National Cooperative Production Amendments of 1993 in the USA and the block exemption to collusion in R&D of Regulation 418/85 in 1985 in the European Union. The main trade-off embodied in these models compares the direct benefit of the cost reduction experienced by a firm with the indirect cost of facing a competitor who also experiences a cost reduction and thus behaves more aggressively in a Cournot market. d'Aspremont and Jacquemin (1988) discuss how this trade-off is affected by the presence of spillovers, when some part of the research output of one firm is leaked to the other firm. Because there are only two firms involved in the models, the formation of an RJV has no external effect on other firms in the industry. Suzumura (1992) and Kamien, Morton and Zang (1990) consider an industry with an arbitrary number n of firms. Kamien et al. (1992) distinguish between different types of alliances: RJVs where firms share their R&D results but do not coordinate their investments, R&D cartels, where firms coordinate their investments but do not share research outputs and RJV cartels where firms coordinate their investments and share research outputs. Both Suzumura (1992) and Kamien et al. (1992) restrict attention to situations where the formation of an RJV has no external effects by considering alliances covering all the firms in the industry. Kamien and Zang (1993) analyze a model with symmetric alliances. Poyago-Theotoky (1995) considers partial alliances that only cover a fraction of the firms in the industry but supposes that only one alliance is formed. Hinloopen (1997) contrasts cooperative research with R&D subsidies and concludes that research subsidies are a more effective policy tool than allowing firms to cooperate in research.

Bloch (1995) proposes a model to endogenously derive the structure of cost-reducing alliances in oligopolies. Consider a linear Cournot oligopoly with inverse demand $P = 1 - Q$. Let a_k denote the size of the alliance A_k. We suppose that firms have complementary assets in R&D so that the marginal cost of production of a firm is linearly decreasing in the size of

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

the association it belongs to. Formally, if firm i belongs to association $k(i)$, its marginal cost of production is given by $c_i = \lambda - \mu a_{k(i)}$. The equilibrium profit of a firm belonging to an association of size a_i is given by

$$R_i = \left[\frac{1-\lambda}{n+1} + \mu a_i - \frac{\mu \sum_k a_k^2}{n+1} \right]^2 .$$

We observe significant differences between the formation of alliances and cartels. First, the formation of an alliance has a negative externality on the profit of outsiders. An increase in a_k, $k \neq i$ reduces the profit of firm i. Second, in a fixed coalition structure, members of larger alliances have higher profits, as $\sum a_k^2$ is constant, but profit is increasing in a_i. These differences lead to very different predictions on equilibrium coalition structures. For example, in an open membership game, firms always have an incentive to join a larger alliance so that the only equilibrium outcome is for all firms to join in a single RJV. On the other hand, in exclusive membership games, the equilibrium alliance structure will not be the grand coalition. To understand this fact, notice that, when a single firm joins an alliance, benefits are asymmetric. The single firm benefits from a large cost reduction whereas alliance members only experience a small reduction in costs as the size of the alliance only increases by one unit. This implies that, when an alliance is very large, it will be reluctant to admit new members. A careful look at the profit function shows that R_i is increasing in a_i until $a_i = \frac{n}{2}$ and decreasing afterwards. The symmetric association structure with two associations of size $\frac{n}{2}$ cannot be an equilibrium either. In order to increase the cost difference with members of the rival association, any association has an incentive to accept more than $\frac{n}{2}$ members. Anticipating that the remaining players will form an association of size $n - a$, members of the first association optimally choose a coalition size of $a^* = \frac{3n+1}{4}$. In the sequential game of coalition formation, the unique equilibrium association structure thus results in the formation of two associations of unequal sizes, one with $\frac{3n+1}{4}$ members and the other with $\frac{n-1}{4}$ members.

Bloch (1995) discusses the extension of the model to Cournot and Bertrand competition with differentiated products. As the level of product differentiation increases, competition on the market is less fierce and the dominant association becomes larger. Interestingly, the sizes of equilibrium associations are identical under Cournot and Bertrand competition. Yi (1998) and Yi and Shin (2000) generalize the model by studying arbitrary demand and cost functions. They identify conditions on demand and cost functions for which the grand coalition emerges in an open membership game. Greenlee (2005) considers a general linear model with intra-RJV and industry-wide spillovers and characterizes the equilibrium outcomes of the open membership game and the sequential game of coalition formation. He finds that the grand coalition always forms in the open membership game, but that a more fragmented coalition structure with different alliances arises in the sequential game. However, the number of alliances is bounded above by 3 for all n. Numerical computations are used to illustrate the size of alliances as a function of the two spillover parameters. Belleflamme (2000) extends the model to asymmetric firms and shows that when cost reductions are not symmetric, the grand coalition may fail to form in the open membership game. Morasch (2000) considers a related model where heterogenous firms in a strategic alliance propose output-based transfer payments. Under this formulation of profit-sharing contracts, he computes numerically the equilibrium association structures for small values of n.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

4.2 Networks of Collaboration

As an alternative to multilateral alliances, Goyal and Joshi (2003) propose a model of networks of bilateral collaboration among firms. They assume that the marginal cost of production is linearly decreasing in the number of bilateral alliances a firm has formed (rather than the size of the alliance it belongs to). We then have $c_i = \lambda - \mu d_i(g)$, where $d_i(g)$ denotes again the degree of firm i in the graph g. Equilibrium profits are given by

$$R_i = \left[\frac{1-\lambda}{n+1} + \mu d_i(g) - \frac{\mu \sum_j d_j(g)}{n+1}\right]^2.$$

As in the model of alliances, the formation of a link between two firms i and j hurts all other competitors, and for a fixed network structure, firms with a higher degree obtain a higher profit. If the formula for the profit is very similar to the formula in strategic alliances, the analysis of the model of network formation is very different. When two firms sign a bilateral agreement, they enjoy a symmetric reduction in production cost. One computes the marginal effect of an additional agreement on firm i's equilibrium quantity as

$$\Delta q_i = \frac{n\mu}{n+1} > 0.$$

As equilibrium profits are increasing in quantities, all bilateral agreements thus raise the firms' profits so that the only pairwise stable network is the complete network. When firms face a significant fixed cost of link formation, the complete network ceases to be stable. Goyal and Joshi (2003) show that stable networks have a dominant group architecture, with one complete component and singleton firms.

Goyal and Moraga-Gonzales (2001) extend the analysis by supposing that firms endogenously choose their research effort. Research effort will be decreasing in the number of links a firm has formed, and hence the addition of a new bilateral agreement may result in lower R&D on the market. In a linear Cournot market, Goyal and Moraga Gonzales (2001) show that research efforts are maximized when every firm is linked to exactly $\frac{n-1}{2}$ competitors. However, as the marginal benefit of an additional link remains positive, firms have an incentive to form the complete network. Hence, in a model with endogenous research efforts, firms engage in excessive collaborative activities.

4.3 Exchange of Information and Trade Associations

Another important instance of collaboration among firms is the exchange of information. We distinguish between two types of information: common value information (about market demand) and private value information (about idiosyncratic costs). Information exchange has been studied in the context of trade associations – groups covering all firms in the industry. The first strand of papers by Novshek and Sonnenschein (1982), Clarke (1983), Vives (1984) and Gal-Or (1985) consider information sharing about an unknown parameter of demand. Novshek and Sonnenschein (1982) and Vives (1984) focus on a duopoly model. Novshek and Sonnenschein (1982) solve for the partial pooling of information, when each firm chooses to pool some of the signals they receive. Vives (1984) compares the incentive to share information under Cournot and Bertrand, and under substitutes and complements, showing

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

that it is optimal not to pool information in games of strategic substitutes (Cournot with substitutes and Bertrand with complements) but optimal to share information in games of strategic complements (Cournot with complements and Bertrand with substitutes). The same result – that information sharing is never optimal under Cournot with substitutes – is obtained by Clarke (1983) and Gal-Or (1985) in an oligopoly model with n firms, quadratic payoffs and normally distributed signals. Li (1985) extends the model by allowing for more general signal distributions, and considers also information sharing about private cost parameters. He finds that firms never have an incentive to share information about common market demand but always have an incentive to share information about private costs in a Cournot oligopoly. Shapiro (1986) also notes that information sharing about private costs arises as an equilibrium. Gal-Or (1986) analyzes information sharing about market demand and private costs under Cournot and Bertrand and shows that there is a stark distinction between Cournot and Bertrand and common value and private value, with no information sharing emerging as the equilibrium outcome for Cournot under common values and Bertrand under private values, and full information sharing for Bertrand under common values and Cournot under private values. Okuno-Fujiwara, Postlewaite, and Suzumura (1990) offer a general argument to show that unraveling results in all firms revealing their private cost information. Raith (1996) provides a useful guide to the literature and a generalization of all existing models, indicating exactly which conditions are required for information sharing.

Most of the literature considers information sharing with all other firms in the industry. One exception is the paper by Kirby (1988), which allows for information sharing among a subset of firms, and considers the formation of information pools. Building on Clarke's (1983) model, she shows that information pooling among a subset of firms may be an equilibrium behavior for some subset of parameters – in sharp contrast to the case where firms must exchange information with all other firms in the industry, where no information is ever shared. This result suggests that allowing firms to form smaller exclusive trade associations may lead to more information sharing, increasing the profit of firms and the expected consumer surplus. Vives (1990) compares different disclosure rules in trade associations. He allows trade associations to use exclusionary disclosure rules – the aggregate signal on the market is only distributed to a fraction of the firms in the industry – and shows that exclusionary disclosure rules restore the firms' incentives to share information, but does not necessarily lead to all firms joining the trade association. Currarini and Feri (2015) analyze information sharing as bilateral agreements among firms, and characterize the stable networks of information sharing in Cournot oligopolies. They show that in the case of private values, pairwise stable networks are connected components with some isolated firms. In the case of correlated signals, they show that pairs of firms always have an incentive to exchange information so that the empty network is never pairwise stable and the complete network is always stable. Hence, as in the case of strategic alliances, there is a sharp contrast between coalition and network formation in information sharing, and firms will more easily share information about demand when agreements are bilateral than when they are multilateral.

REFERENCES

Banal-Estañol, Albert and Marco Ottaviani. Mergers with product market risk. *Journal of Economics & Management Strategy*, 15(3):577–608, 2006.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Barros, Pedro P. Endogenous mergers and size asymmetry of merger participants. *Economics Letters*, 60(1): 113–119, 1998.

Belleflamme, Paul. Stable coalition structures with open membership and asymmetric firms. *Games and Economic Behavior*, 30(1):1–21, 2000.

Belleflamme, Paul and Francis Bloch. Market sharing agreements and collusive networks. *International Economic Review*, 45(2):387–411, 2004.

Bloch, Francis. Endogenous structures of association in oligopolies. *The RAND Journal of Economics*, 26(3): 537–556, 1995.

Bloch, Francis. Sequential formation of coalitions in games with externalities and fixed payoff division. *Games and Economic Behavior*, 14(1):90–123, 1996.

Bos, Iwan and Joseph E Harrington Jr. Endogenous cartel formation with heterogeneous firms. *The RAND Journal of Economics*, 41(1):92–117, 2010.

Braid, Ralph M. Stackelberg price leadership in spatial competition. *International Journal of Industrial Organization*, 4(4):439–449, 1986.

Braid, Ralph M. The price and profit effects of horizontal mergers in two-dimensional spatial competition. *Economics Letters*, 62(1):113–119, 1999.

Brandenburger, Adam M and Barry J Nalebuff. *Co-opetition*. New York: Crown Business, 2011.

Brito, Duarte. Preemptive mergers under spatial competition. *International Journal of Industrial Organization*, 21(10):1601–1622, 2003.

Brown, Murray and Shin-Hwan Chiang. Unsystematic risk and coalition formation in product markets. *International Journal of Industrial Organization*, 20(3):313–338, 2002.

Brown, Murray and Shin-Hwan Chiang. *Coalitions in Oligopolies: An Introduction to the Sequential Procedure, Volume 259*. Bingley, UK: Emerald, 2003.

Chatterjee, Kalyan, Bhaskar Dutta, Debraj Ray and Kunal Sengupta. A noncooperative theory of coalitional bargaining. *The Review of Economic Studies*, 60(2):463–477, 1993.

Chwe, Michael Suk-Young. Farsighted coalitional stability. *Journal of Economic Theory*, 63(2):299–325, 1994.

Clarke, Richard N. Collusion and the incentives for information sharing. *The Bell Journal of Economics*, 14(2): 383–394, 1983.

Compte, Olivier, Frederic Jenny and Patrick Rey. Capacity constraints, mergers and collusion. *European Economic Review*, 46(1):1–29, 2002.

Cramton, Peter C and Thomas R Palfrey. Cartel enforcement with uncertainty about costs. *International Economic Review*, 31(1):17–47, 1990.

Currarini, Sergio and Francesco Feri. Information sharing networks in linear quadratic games. *International Journal of Game Theory*, 44(3):701–732, 2015.

Currarini, Sergio and Marco A Marini. Coalition formation in games without synergies. *International Game Theory Review*, 8(01):111–126, 2006.

d'Aspremont, Claude and Alexis Jacquemin. Cooperative and noncooperative R&D in duopoly with spillovers. *The American Economic Review*, 78(5):1133–1137, 1988.

d'Aspremont, Claude, Alexis Jacquemin, Jean Jaskold Gabszewicz and John A Weymark. On the stability of collusive price leadership. *Canadian Journal of Economics*, 16(1):17–25, 1983.

Davidson, Carl and Ben Ferrett. Mergers in multidimensional competition. *Economica*, 74(296):695–712, 2007.

Deneckere, Raymond and Carl Davidson. Incentives to form coalitions with Bertrand competition. *The RAND Journal of Economics*, 16(4):473–486, 1985.

Diamantoudi, Effrosyni. Stable cartels revisited. *Economic Theory*, 26(4):907–921, 2005.

Donsimoni, Marie-Paule. Stable heterogeneous cartels. *International Journal of Industrial Organization*, 3(4): 451–467, 1985.

Donsimoni, Marie-Paule, Nicholas S Economides and Herakles M Polemarchakis. Stable cartels. *International Economic Review*, 27(2):317–327, 1986.

Ericson, Richard and Ariel Pakes. Markov-perfect industry dynamics: A framework for empirical work. *The Review of Economic Studies*, 62(1):53–82, 1995.

Espinosa, Maria Paz and Ines Macho-Stadler. Endogenous formation of competing partnerships with moral hazard. *Games and Economic Behavior*, 44(1):183–194, 2003.

Farrell, Joseph and Carl Shapiro. Horizontal mergers: An equilibrium analysis. *The American Economic Review*, 80(1):107–126, 1990.

Faulí-Oller, Ramon. Takeover waves. *Journal of Economics & Management Strategy*, 9(2):189–210, 2000.

Finus, Michael and Bianca Rundshagen. Membership rules and stability of coalition structures in positive externality games. *Social Choice and Welfare*, 32(3):389–406, 2009.

Fridolfsson, Sven-Olof and Johan Stennek. Why mergers reduce profits and raise share prices: A theory of preemptive mergers. *Journal of the European Economic Association*, 3(5):1083–1104, 2005.

Gabszewicz, Jean J, Marco A Marini and Ornella Tarola. Alliance formation in a vertically differentiated market. *CORE Discussion Papers*, 30:2015, 2015.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Gal-Or, Esther. Information sharing in oligopoly. *Econometrica: Journal of the Econometric Society*, 53(2):329–343, 1985.
Gal-Or, Esther. Information transmission: Cournot and Bertrand equilibria. *The Review of Economic Studies*, 53(1):85–92, 1986.
Giraud-Héraud, Eric, Hakim Hammoudi and Mahdi Mokrane. Multiproduct firm behaviour in a differentiated market. *Canadian Journal of Economics/Revue canadienne d'économique*, 36(1):41–61, 2003.
Gowrisankaran, Gautam. A dynamic model of endogenous horizontal mergers. *The RAND Journal of Economics*, 30(1):56–83, 1999.
Gowrisankaran, Gautam and Thomas J Holmes. Mergers and the evolution of industry concentration: Results from the dominant-firm model. *RAND Journal of Economics*, 35(3):561–582, 2004.
Goyal, Sanjeev and Sumit Joshi. Networks of collaboration in oligopoly. *Games and Economic Behavior*, 43(1): 57–85, 2003.
Goyal, Sanjeev and Jose Luis Moraga-Gonzalez. R&D networks. *Rand Journal of Economics*, 32(4):686–707, 2001.
Graham, Daniel A and Robert C Marshall. Collusive bidder behavior at single-object second-price and English auctions. *Journal of Political Economy*, 95(6):1217–1239, 1987.
Greenlee, Patrick. Endogenous formation of competitive research sharing joint ventures. *The Journal of Industrial Economics*, 53(3):355–391, 2005.
Gul, Faruk. Bargaining foundations of Shapley value. *Econometrica: Journal of the Econometric Society*, 57(1): 81–95, 1989.
Hart, Sergiu and Mordecai Kurz. Endogenous formation of coalitions. *Econometrica: Journal of the Econometric Society*, 51(4):1047–1064, 1983.
Heywood, John S, Kristen Monaco and Robert Rothschild. Spatial price discrimination and merger: The n-firm case. *Southern Economic Journal*, 67(3):672–684, 2001.
Hinloopen, Jeroen. Subsidizing cooperative and noncooperative R&D in duopoly with spillovers. *Journal of Economics*, 66(2):151–175, 1997.
Horn, Henrik and Lars Persson. Endogenous mergers in concentrated markets. *International Journal of Industrial Organization*, 19(8):1213–1244, 2001.
Jackson, Matthew O and Asher Wolinsky. A strategic model of social and economic networks. *Journal of Economic Theory*, 71:44–74, 1996.
Kamien, Morton I and Israel Zang. The limits of monopolization through acquisition. *The Quarterly Journal of Economics*, 105(2):465–499, 1990.
Kamien, Morton I and Israel Zang. Competitively cost advantageous mergers and monopolization. *Games and Economic Behavior*, 3(3):323–338, 1991.
Kamien, Morton I and Israel Zang. Competing research joint ventures. *Journal of Economics & Management Strategy*, 2(1):23–40, 1993.
Kamien, Morton I, Eitan Muller and Israel Zang. Research joint ventures and R&D cartels. *The American Economic Review*, 82(5):1293–1306, 1992.
Katz, Michael L. An analysis of cooperative research and development. *The RAND Journal of Economics*, 17(4): 527–543, 1986.
Kirby, Alison J. Trade associations as information exchange mechanisms. *The RAND Journal of Economics*, 17(4):138–146, 1988.
Konishi, Hideo and Ping Lin. Stable cartels with a Cournot fringe in a symmetric oligopoly. *Keio Economic Studies*, 36(2):1–10, 1999.
Kuhn, Kai-Uwe and Massimo Motta. The economics of joint dominance. Technical report, mimeo, 2001.
Kuipers, Jeroen and Norma Olaizola. A dynamic approach to cartel formation. *International Journal of Game Theory*, 37(3):397–408, 2008.
Levy, David T and James D Reitzes. Anticompetitive effects of mergers in markets with localized competition. *Journal of Law, Economics, & Organization*, 8(2):427–440, 1992.
Li, Lode. Cournot oligopoly with information sharing. *The RAND Journal of Economics*, 16(4):521–536, 1985.
Macho-Stadler, Inés, David Pérez-Castrillo and Nicolás Porteiro. Sequential formation of coalitions through bilateral agreements in a Cournot setting. *International Journal of Game Theory*, 34(2):207–228, 2006.
Mailath, George J and Peter Zemsky. Collusion in second price auctions with heterogeneous bidders. *Games and Economic Behavior*, 3(4):467–486, 1991.
Mauleon, Ana and Vincent Vannetelbosch. Farsightedness and cautiousness in coalition formation games with positive spillovers. *Theory and Decision*, 56(3):291–324, 2004.
McAfee, R Preston and John McMillan. Bidding rings. *The American Economic Review*, 82(3):579–599, 1992.
Morasch, Karl. Strategic alliances as Stackelberg cartels: Concept and equilibrium alliance structure. *International Journal of Industrial Organization*, 18(2):257–282, 2000.
Myerson, Roger B. *Game Theory: Analysis of Conflict*. Cambridge, MA: Harvard University Press, 1991.
Nocke, Volker. Cartel stability under capacity constraints: The traditional view restored. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1158300, 1999.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Norman, George, Lynne Pepall and Daniel Richards. Product differentiation, cost-reducing mergers, and consumer welfare. *Canadian Journal of Economics/Revue canadienne d'économique*, 38(4):1204–1223, 2005.

Novshek, William and Hugo Sonnenschein. Fulfilled expectations Cournot duopoly with information acquisition and release. *The Bell Journal of Economics*, 13:214–218, 1982.

Okada, Akira. A noncooperative coalitional bargaining game with random proposers. *Games and Economic Behavior*, 16(1):97–108, 1996.

Okuno-Fujiwara, Masahiro, Andrew Postlewaite and Kotaro Suzumura. Strategic information revelation. *The Review of Economic Studies*, 57(1):25–47, 1990.

Pérez-Castrillo, J David. Cooperative outcomes through noncooperative games. *Games and Economic Behavior*, 7(3):428–440, 1994.

Perry, Martin K and Robert H Porter. Oligopoly and the incentive for horizontal merger. *The American Economic Review*, 75(1):219–227, 1985.

Pesendorfer, Martin. Mergers under entry. *RAND Journal of Economics*, 36(3):661–679, 2005.

Poyago-Theotoky, Joanna. Equilibrium and optimal size of a research joint venture in an oligopoly with spillovers. *The Journal of Industrial Economics*, 43(2):209–226, 1995.

Prokop, Jacek. Process of dominant-cartel formation. *International Journal of Industrial Organization*, 17(2): 241–257, 1999.

Raith, Michael. A general model of information sharing in oligopoly. *Journal of Economic Theory*, 71:260–288, 1996.

Ray, Debraj and Rajiv Vohra. A theory of endogenous coalition structures. *Games and Economic Behavior*, 26(2):286–336, 1999.

Ray, Debraj and Rajiv Vohra. Coalitional power and public goods. *Journal of Political Economy*, 109(6):1355–1385, 2001.

Rothschild, Robert, John S Heywood and Kristen Monaco. Spatial price discrimination and the merger paradox. *Regional Science and Urban Economics*, 30(5):491–506, 2000.

Rubinstein, Ariel. Perfect equilibrium in a bargaining model. *Econometrica: Journal of the Econometric Society*, 50(1):97–109, 1982.

Salant, Stephen W, Sheldon Switzer and Robert J Reynolds. Losses from horizontal merger: The effects of an exogenous change in industry structure on Cournot-Nash equilibrium. *The Quarterly Journal of Economics*, 98(2):185–199, 1983.

Shaffer, Sherrill. Stable cartels with a Cournot fringe. *Southern Economic Journal*, 61:744–754, 1995.

Shapiro, Carl. Exchange of cost information in oligopoly. *The Review of Economic Studies*, 53(3):433–446, 1986.

Stigler, George J. Monopoly and oligopoly by merger. *The American Economic Review*, 40:23–34, 1950.

Suzumura, Kotaro. Cooperative and noncooperative R&D in an oligopoly with spillovers. *The American Economic Review*, 78(5):1307–1320, 1992.

Thoron, Sylvie. Formation of a coalition-proof stable cartel. *Canadian Journal of Economics*, 31(1):63–76, 1998.

Vasconcelos, Helder. Tacit collusion, cost asymmetries, and mergers. *RAND Journal of Economics*, 36(1):39–62, 2005.

Vives, Xavier. Duopoly information equilibrium: Cournot and Bertrand. *Journal of Economic Theory*, 34(1):71–94, 1984.

Vives, Xavier. Trade association disclosure rules, incentives to share information, and welfare. *The RAND Journal of Economics*, 21(3):409–430, 1990.

von Neumann, John and Oskar Morgenstern. *Theory of Games and Economic Behavior*, *Volume 60*. Princeton, NJ: Princeton University Press, 1944.

Yi, Sang-Seung. Stable coalition structures with externalities. *Games and Economic Behavior*, 20(2):201–237, 1997.

Yi, Sang-Seung. Endogenous formation of joint ventures with efficiency gains. *The RAND Journal of Economics*, 29(3):610–631, 1998.

Yi, Sang-Seung and Hyukseung Shin. Endogenous formation of research coalitions with spillovers. *International Journal of Industrial Organization*, 18(2):229–256, 2000.

Zu, Lei, Jin Zhang and Shouyang Wang. The size of stable cartels: An analytical approach. *International Journal of Industrial Organization*, 30(2):217–222, 2012.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

14. TU oligopoly games and industrial cooperation

*Jingang Zhao**

1 INTRODUCTION

This chapter surveys existing results and lists nine future areas in TU oligopoly games and industrial cooperation or, precisely, cooperative oligopoly games with transferable utilities (TUs). Such results and future research are both empirically important and theoretically interesting.

On the empirical side, the model of TU oligopoly games is the proper tool with which to study industrial cooperation, ranging from early divisions of labor to modern merger contracts; its applications help one understand the structural changes in industries and the effects of regulatory policies.[1] For example, empty-core theory has provided an understanding about the US consolidation movement of the late nineteenth century, which actually originated the field of industrial organization (McWilliams and Keith 1994).

Core theory allows one to estimate the merging costs or the transaction costs of horizontal mergers (Zhao 2009a). Reductions in merging costs provide a new explanation for the two greatest merger waves around the turns of the twentieth and the twenty-first centuries (Zhao 2009b). Similar cost reductions by European Union Directive Solvency II (enacted on January 1, 2016) will likely drive more mergers and acquisitions in the EU insurance industry (Stoyanova and Gruendl 2014).

On the theory side, the results are advances in the refinements and applications of the core, which is the most important solution in cooperative game theory. They are developed around the stability of a monopoly merger contract. It first converts the oligopoly to a TU coalitional game or a partition function game and then characterizes the core. The main task is to identify conditions on the parameters in an oligopoly for a non-empty core. One sufficient condition for a non-empty core is convexity or supermodularity, whose existence is known only in some linear oligopolies.

There is no need to emphasize the importance of core theory in industrial organization, because non-empty core and profitability are the two preconditions for each horizontal merger. It should be pointed out that oligopoly games or cooperative oligopoly games with non-transferable utilities (NTUs) are not surveyed here. The model of NTU games is the tool with which to study collusion such as illegal cartel agreements,[2] which are not in the mainstream of industrial organization.

The rest of this survey is organized as follows: Section 2 reviews three game models and defines their core solutions and some refinements that are relevant in oligopolies. Section 3

* I am indebted to Don Gilchrist, Aymeric Lardon, Marco Marini, Holger Meinhardt, Martin Shubik, Donald Smythe, Giorgos Stamatopoulos, Xinghe Wang, Jongsay Yong and Jianbo Zhang for valuable comments and suggestions. All errors, of course, are my own.

[1] See Daughety and Reinganum (Chapter 9, Volume II of this *Handbook*) for a survey.

[2] See Bloch (Chapter 13 in this *Handbook*), Marini (2009) and Marini (Chapter 3, Volume II of this *Handbook*) for surveys on collusion studies.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

first reviews ten oligopoly models and their equilibrium expressions; it then reviews the existence and refinements of the core; and finally it lists seven extensions. Section 4 reviews the results on non-monopoly partitions. Section 5 reviews empirical studies of the core. Section 6 concludes with a brief discussion about future research.

2 THE CORE AS A SOLUTION IN THREE GAME MODELS

This section reviews the core as a cooperative solution in three forms of games or three models: coalitional TU games or simply coalitional games (also called games in characteristic form or games in coalitional form; von Neumann and Morgenstern 1944), normal form games (also called strategic games; Nash 1950), and partition function games (Thrall and Lucas 1963). These games are defined below.

Let $N = \{1, \ldots, n\}$ be the set of players or firms. Each subset $S \subseteq N$ is called an alliance or a coalition or a merger. Each partition $\Delta = \{S_1, S_2, \ldots, S_h\}$ of N is called a coalition structure or market structure, representing a set of h simultaneous mergers in which each merger S_j has $k_j = |S_j|$ members (so $\Sigma_{j=1}^{h} k_j = n$).[3]

A coalitional game (von Neumann and Morgenstern 1944) is a set function given by

$$\Gamma_c = \{N, v(\cdot)\}, \tag{14.1}$$

specifying a non-negative joint payoff or profit $v(S)$ for each coalition $S \subseteq N$. The central question here is how to split the grand coalition's payoff $v(N)$ among the n players; this implicitly assumes that the grand coalition's payoff (such as monopoly profit) is optimal or maximal among all coalitions and all partitions of N.

A normal form game (Nash 1950) is given by

$$\Gamma = \{N, X_i, u_i\}, \tag{14.2}$$

specifying a choice set $X_i \subset R^{k_i}$ in k_i-dimensional Euclidian space and a payoff function $u_i(x)$, $x = (x_1, \ldots, x_n) \in X = \Pi_{j=1}^{n} X_j$, for each player $i \in N$. Game (14.2) is called a normal form TU game if all u_i are transferable, such as dollars. The central question here is what is the solution or a list of choices $x = (x_1, \ldots, x_n)$ that rational players will choose.

A partition function game (Thrall and Lucas 1963) is given by

$$\Gamma_p = \{N, \phi(\cdot)\}, \tag{14.3}$$

specifying a vector of joint payoffs $\phi(\Delta) = \{\phi_S = \phi_S(\Delta) | S \in \Delta\}$ for each partition Δ and each of its coalitions $S \in \Delta$. One of the central questions here is the same as that in a coalitional game: how to split the grand coalition's payoff $v(N) = \phi_N$, assuming that $v(N)$ is the maximum among all partitions.

---

[3] A partition $\Delta = \{S_1, S_2, \ldots, S_h\}$ satisfies: $S_j \neq \varnothing, \cup S_j = N$, and $S_i \cap S_j = \varnothing$, all $i \neq j$. This also represents an h-firm n-product multi-product oligopoly, in which each firm S_j (or simply j) produces $k_j = |S_j|$ products. As shown in Zhao (2012), the equilibrium expressions for such multi-product oligopoly are identical to that for the postmerger equilibria. See Faulí-Oller and Sandonís (Chapter 2, Volume II of this *Handbook*) for a survey on other studies on mergers.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

The following four assumptions (Zhao 2018) are implicitly assumed to support a variety of solutions for the games in (14.1–14.3):

A1 (Assumption 1) *Players are able to take collective actions.*
A2 *Players are unable to take any form of coordinated or collective actions.*
A3 *Players are able to costlessly negotiate and enforce a joint action.*
A4 *Given a partition* $\Delta = \{S_1, S_2, \ldots, S_h\}$, *A*3 *holds for each* $S \in \Delta$, *and A*2 *holds for each* $T \notin \Delta$ *such that there are* $i \neq j$, $T \cap S_i \neq \varnothing$ *and* $T \cap S_j \neq \varnothing$.

$A1$ applies in most situations, $A2$ characterizes the original Prisoner's Dilemma game in which the two players have no access to any form of communication or coordination or agreements. If they could coordinate their choices or make deals by using a joint counsel, the nature of the game will become that under $A1$ or $A3$ and thus invalidate the Nash equilibrium predicted by $A2$.

Note also that players under $A4$ can negotiate and enforce a joint action if they belong to the same coalition, but they can not take collective or coordinated actions if they are from two or more different coalitions. Thus, $A2$ is a special case of $A4$ for the finest partition or premerger structure $\Delta_0 = \{\{1\}, \ldots, \{n\}\}$, and $A3$ is another special case for the coarsest partition or monopoly structure $\Delta_m = \{N\}$.

Because these assumptions determine or limit a player's rationality, they are the foundations of game theory upon which various solutions or theories are built. For example, $A2$, $A3$ and $A4$ are the foundations of non-cooperative solution or Nash equilibrium (Nash 1950), cooperative solutions (Shapley 1955, von Neumann and Morgenstern 1944), and hybrid solutions (Zhao 1992), respectively.

Care needs to be taken when applying these assumptions in a particular game. For example, the actions for a coalition S under both $A1$ and $A2$ in a normal form game (14.2) are the vectors of their choices given by $x_S = \{x_j \,|\, j \in S\} \in X_S = \Pi_{j \in S} X_j$; but their actions in a coalitional game (14.1) are the splits of their joint payoff given by $\theta_S = \{\theta_j \,|\, j \in S\}$, which satisfies $\sum_{j \in S} \theta_j = v(S)$, and $\theta_j \geq 0$, all j.

2.1 The Core in Coalitional Games

Given a coalitional game (14.1), a split of $v(N)$ is a payoff vector $\theta = (\theta_1, \ldots, \theta_n) \in R^n_+$ such that $\Sigma^n_{j=1} \theta_j = v(N)$, with θ_j as player j's payoff, all j. A split θ is rational for a coalition $S \subseteq N$ (or undominated or unblocked by S) if $\Sigma_{j \in S} \theta_j \geq v(S)$, and θ is in the core (or a core vector) if it is rational for all $S \subseteq N$. This was first defined by Shapley (1955) as given here:[4]

Definition 1 (Shapley 1955) *The core of game* (14.1) *is the set of the splits of* $v(N)$ *that are rational for all proper coalitions. Precisely, this is given by*

$$Core(\Gamma_c) = \{\theta \in R^n_+ | \Sigma^n_{j=1} \theta_j = v(N), \text{ and } \Sigma_{j \in S} \theta_j \geq v(S), \text{ all } S \neq N\}. \quad (14.4)$$

4 Shapley first coined the term *core solution* during 1952–53 in one of his conversations with Shubik, and Gillies first used the term *core* during the same period, referring to some intersections of the stable sets. Gillies (1959) had been mistakenly cited in most previous studies as the first paper that defined the core. See Zhao (2016) for the history of the core.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Lemma 1 below summarizes two complete arguments for a non-empty core:

Lemma 1 *Given* (14.4)*, the following three arguments are equivalent: (i) Core*$(\Gamma_c) \neq \varnothing$; *(ii) the game is balanced* (Bondareva 1962*,* Shapley 1967); *and (iii) the grand coalition's payoff is above the minimum no-blocking payoff* (Zhao 2001b).

Specifically, argument *(ii)* holds if $\Sigma_{T \in B} w_T v(T) \leq v(N)$ holds for each balanced collection of coalitions $B = \{T_1, \ldots, T_k\}$ with a balancing vector $w = \{w_T | T \in B\}$:[5] and argument *(iii)* holds if $v(N) \geq mnbp$ holds, where *mnbp* is the game's *minimum no-blocking payoff* given *by*

$$mnbp = Min\{\Sigma_{j \in N} \theta_j | \theta \in R^n_+, \text{ and } \Sigma_{j \in S} \theta_j \geq v(S), all\ S \neq N\}. \quad (14.5)$$

The above *mnbp* method for core existence has an intuitive interpretation and it enables one to estimate the transaction costs of horizontal mergers (Zhao 2009a).

2.2 The α-core and β-core in Normal Form TU Games

Given a coalition S in the normal form TU game (14.2), recall that its choice vector is given by $x_S = \{x_j | j \in S\} \in X_S = \Pi_{j \in S} X_j$. Let $x_{-S} = \{x_j | j \notin S\} \in X_{-S} = \Pi_{j \notin S} X_j$ be the outsiders' choice vector, and rearrange $x = (x_1, \ldots, x_n) \in X = \Pi^n_{j=1} X_j$ as $x = (x_S, x_{-S})$, so $u_i(x) = u_i(x_S, x_{-S})$ for all i. Then, the coalition's joint payoffs under the α- and β-beliefs are defined by

$$v_\alpha(S) = Max_{x_s \in X_S} Min\{\Sigma_{j \in S} u_j(x_S, x_{-S}) | x_{-S} \in X_{-S}\} \text{ and} \quad (14.6)$$

$$v_\beta(S) = Min_{x_{-s} \in X_{-S}} Max\{\Sigma_{j \in S} u_j(x_S, x_{-S}) | x_S \in X_S\}, \quad (14.7)$$

respectively. $v_\alpha(S)$ is often called the guaranteed or worst payoff, because S can guarantee a joint payoff no less than $v_\alpha(S)$ by choosing some $\overline{x}_S$ (i.e., $\Sigma_{j \in S} u_j(\overline{x}_S, x_{-S}) \geq v_\alpha(S)$ for all x_{-S}). On the other hand, S can not be prevented from receiving at least $v_\beta(S)$, as they have a best response function

$$x^*_S = x_S(x_{-S}) = ArgMax\{\Sigma_{j \in S} u_j(x_S, x_{-S}) | x_S \in X_S\}^6 \quad (14.8)$$

such that $\Sigma_{j \in S} u_j(x_S(x_{-S}), x_{-S}) \geq v_\beta(S)$ for each x_{-S}.

By $Max\{\Sigma_{j \in N} u_j(x) | x \in X\} = v(N) = v_\alpha(N) = v_\beta(N)$, the grand coalition's payoff is the same under both beliefs. An updated version of the α- and β-cores in Aumann (1959) are given here:

[5] A collection of coalitions $B = \{T_1, \ldots, T_k\}$ is balanced if it has a balancing vector w, or a positive weight $w_T > 0$ for each $T \in B$, such that for each player $i \in N$, $\Sigma_{T \in B(i)} w_T = 1$ holds, where $B(i) = \{T \in B | i \in T\}$ is the subcollection of coalitions to which player i belongs.

[6] ArgMax denotes the set of maximal solutions for each maximization problem; precisely, given $Max\{f(x)| x \in X\}$, one has $ArgMax\{f(x)|x \in X\} = \{y \in X | f(y) \geq f(x), \text{all } x \in X\}$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Definition 2 (Aumann 1959) *Given a normal form TU game* (14.2) *and its coalitional payoffs* $v_\alpha(S)$ *and* $v_\beta(S)$ *in* (14.6–14.7), *its* α- *and* β-*coalitional games are*

$$\Gamma_\alpha = \{N, v_\alpha(\cdot)\} \text{ and } \Gamma_\beta = \{N, v_\beta(\cdot)\}, \text{ and} \tag{14.9}$$

the cores of above Γ_α *and* Γ_β *are called the* α-*core and the* β-*core, respectively.*

As shown in Zhao (1999a, p. 156), an empty α-core means that for each θ satisfying $\theta \geq 0$ and $\Sigma\theta_j = v(N)$, there exists S and $x_S \in X_S$ such that $\Sigma_{j\in S} u_j(x_S, x_{-S}) > \Sigma_{j\in S}\theta_j$ for all x_{-S}, and an empty β-core means the existence of S with a reaction function $x_S^* = x_S(x_{-S})$ in (14.8) such that $\Sigma_{j\in S} u_j(x_S(x_{-S}), x_{-S}) > \Sigma_{j\in S}\theta_j$ for all x_{-S}. Thus, an empty α-core implies an empty β-core, so a non-empty β-core implies a non-empty α-core, or $Core(\Gamma_\beta) \subseteq Core(\Gamma_\alpha)$ holds. This can also be understood by the following interpretation due to Jianbo Zhang of the University of Kansas (Zhang 2016, personal communication).

For the α-core, imagine that the outsiders have a spy in S and thus know each action taken by S. Consequently, all actions taken by S are doomed to be disastrous, and the best S could do is damage control or choose the best of the worst given by $v_\alpha(S)$. On the other hand, one imagines, for the β-core, that S have a spy in $N\backslash T$ and know each action taken by the outsiders. In this case, each of the outsiders' actions will lead to the best outcome for S, and the worst harm that the outsiders could do to S is given by $v_\beta(S)$. Having a spy is better than being spied on, so one has $v_\beta(S) \geq v_\alpha(S)$ and thus $Core(\Gamma_\beta) \subseteq Core(\Gamma_\alpha)$.

The general existence of NTU α-core was established by Scarf (1971). He showed that the normal form game (14.2) has a non-empty NTU α-core if (a) all choice sets are compact and convex, and (b) all payoff functions are continuous and quasi-concave. This has been extended to a non-empty TU α-core by adding the assumption of weak separability (Zhao 1999c), and a non-empty TU β-core by adding the assumption of strong separability (Zhao 1999a). These two extensions[7] are relevant in oligopoly models, which are summarized here:

Lemma 2 *Let* $C_\alpha = Core(\Gamma_\alpha)$ *and* $C_\beta = Core(\Gamma_\beta)$ *be the TU* α- *and* β-*cores in* (14.2). *Then, (i)* $C_\alpha \neq \varnothing$ *if (a) all* X_i *are compact and convex, (b) all* $u_i(x)$ *are continuous and quasi-concave, and (c) weak separability holds* (Zhao 1999c); *(ii)* $C_\beta \neq \varnothing$ *if (a) all* X_i *are compact and convex, (b) all* $u_i(x)$ *are continuous and quasi-concave, and (c) strong separability holds* (Zhao 1999a).

Roughly speaking, strong (weak) separability requires that the outsiders' choices that minimize the insiders' joint payoff in (14.7) (in (14.6)) also minimize each insider's individual payoff in a relevant range (at a relevant point). The precise statements of these two conditions are not reviewed here because they both automatically hold in oligopoly models. Readers are referred to Zhao (1999a), Zhao (1999c), and Meinhardt (2002, pp. 69–88) for details and numerical examples.

[7] See Allen (2006), Kajii (1992), Uyanık (2015), Wilson (1978) and Yannelis (2005) for other extensions.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

2.3 The γ-core, δ-core and their Variations in Partition Function Games

Given a partition $\Delta = \{S_1, S_2, \ldots, S_h\}$ in the partition function game (14.3), consider the deviation by, or formation of, a new coalition $S = \{i_1, \ldots, i_k\} \notin \Delta$. Let the set of those partitions of which S is a member be denoted by

$$\Pi(S) = \{\Delta' \in \Pi | \Delta' = \{S, T_1, \ldots, T_m\}\}, \tag{14.10}$$

where Π is the set of all partitions of N. Before deviating, the insiders or players in S are assumed to have hypothesized or believed a reasonable reaction to their deviation by the outsiders in $N \backslash S = \{j | j \notin S\}$.

Six possible reactions based on six kinds of beliefs are known and are reviewed here. These beliefs lead to six core solutions for the game (14.3): the γ-core and δ-core in Hart and Kurz (1983), α^*-core in Zhao (1996, 2013), *e*-core in Yong (2004), *j*-core in Lekeas (2013), and *f*-core in Lekeas and Stamatopoulos (2014). Note that this list excludes those core refinements (such as the *lf*-core of Currarini and Marini, 2003 reviewed at end of subsection 3.3) in a normal form game that are not defined for partition function games.[8]

For simplicity, all definitions here focus on the coarsest or monopoly partition $\Delta_m = \{N\}$, which is extended to a general or non-monopoly partition in section 4:

1. The breakup belief or γ-belief (Hart and Kurz 1983): insiders believe that the $(n-k)$ outsiders in $N \backslash S$ will break up into singletons or the new partition is $\Delta_\gamma = \Delta_\gamma(S, \Delta_m) = \{S, \{j_1\}, \ldots, \{j_{n-k}\}\} \in \Pi(S)$, so the insiders' payoff and the γ-coalitional game are

$$v_\gamma(S) = \phi_S(\Delta_\gamma), \text{ all } S; \text{ and } \Gamma_\gamma = \{N, v_\gamma(\cdot)\}. \tag{14.11}$$

2. The loyal belief or δ-belief (Hart and Kurz 1983): insiders believe that outsiders are loyal to each other and stay in the coalition $N \backslash S$, so the new partition is $\Delta_\delta = \Delta_\delta(S, \Delta_m) = \{S, N \backslash S\} \in \Pi(S)$, and their payoff and the δ-coalitional game are

$$v_\delta(S) = \phi_S(\Delta_\delta), \text{ all } S; \text{ and } \Gamma_\delta = \{N, v_\delta(\cdot)\}. \tag{14.12}$$

3. The cautious belief or α^*-belief (Zhao 1996, 2013): insiders are cautious about their smallest payoff at the worst partition: $\Delta_{\alpha*} \equiv \Delta_{\alpha*}(S) = \{S, T_1^{\alpha*}, \ldots, T_{m(\alpha*)}^{\alpha*}\} \in \Pi(S)$, or they believe that the outsiders partition themselves to minimize the insiders' joint payoff, so the insiders' payoff and the α^*-coalitional game are

$$v_{\alpha*}(S) = \phi_S(\Delta_{\alpha*}), \text{ all } S; \text{ and } \Gamma_{\alpha*} = \{N, v_{\alpha*}(\cdot)\}, \tag{14.13}$$

where $\Delta_{\alpha*} = \Delta_{\alpha*}(S)$ is the solution of $Min\{\phi_S = \phi_S(\Delta') | \Delta' \in \Pi(S)\}$. Note that this cautious or worst partition $\Delta_{\alpha*}$ is independent of all current partitions Δ.

[8] It also excludes related studies such as the core in partition function games from a common pool resource (Funaki and Yamato 1999) and the axiomatization of such cores in partition function games (Bloch and Van den Nouweland 2014).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

4. The efficient belief or e-belief (Yong 2004): insiders believe that the outsiders choose an efficient partition (or optimal partition in TU games) for themselves among all partitions of $N\backslash S$, so the insiders' payoff and the e-coalitional game are

$$v_e(S) = \phi_S(\Delta_e), \text{ all } S; \text{ and } \Gamma_e = \{N, v_e(\cdot)\}, \tag{14.14}$$

where $\Delta_e = \Delta_e(S) = \{S, T_1^e, \ldots, T_{m(e)}^e\}$ solves $Max\{\Sigma_{T\in\Delta'\backslash S}\phi_T(\Delta')|\Delta' \in \Pi(S)\}$. This efficient partition Δ_e is also independent of all current partitions Δ.

The next two beliefs assume that the payoffs for each $\Delta = \{S_1, S_2, \ldots, S_h\}$ are determined by the number and sizes of its coalitions or precisely by h and $s_i = |S_i|$, $i = 1, \ldots, h$. Such property holds when players are symmetric within each coalition.

5. j-belief (Lekeas 2013): let $s = |S|$ be the cardinality or the number of insiders for each $S \neq N$, then a j-belief is an integer-to-integer function $j(s)$, $1 \leq j(s) \leq n-s$, for $s = 1, \ldots, n-1$, defining the belief for all coalitions with s members that outsiders are divided into $j(s)$ coalitions and the worst of such j-partitions will be formed, so the insiders' payoff and the j-coalitional game are

$$v_j(S) = \phi_S(\Delta^*_{j(s)}), \text{ all } S; \text{ and } \Gamma_j = \{N, v_j(\cdot)\}, \tag{14.15}$$

where $\Delta^*_{j(s)} = \{S, T_1^*, \ldots, T^*_{j(s)}\}$ solves $Min\{\phi_S = \phi_S(\Delta_{j(s)})|\Delta_{j(s)} \in \Delta_j(S)\}$, with $\Delta_j(S) = \{\Delta|\Delta = \{S, T_1, \ldots, T_{j(s)}\} \in \Pi(S)\}$ as the set of all j-partitions or all partitions in which the outsiders are divided into $j(s)$ coalitions.[9]

The next belief further assumes that the payoff $\phi_S(\Delta)$ of each $S \in \Delta = \{S_1, S_2, \ldots, S_h\}$ is determined by h or the number of coalitions in Δ. This property holds in standard symmetric homogeneous Cournot model with linear cost.

6. The probability belief or f-belief (Lekeas and Stamatopoulos 2014): a probability belief is an integer-to-probability vector function $f(s)$ (i.e., $f(s) \in R_+^{n-s}$, $\Sigma_{j=1}^{n-s} f_j(s) = 1$) for $s = 1, \ldots, n-1$, defining the belief for all coalitions with s members that outsiders are randomly partitioned into j-coalitions with a probability $f_j(s)$, $j = 1, \ldots, n-s$, so the insiders' payoff and the f-coalitional game are

$$v_f(S) = \Sigma_{j=1}^{n-s} f_j(s)\phi_S(\Delta_j), \text{ all } S; \text{ and } \Gamma_f = \{N, v_f(\cdot)\}, \tag{14.16}$$

where Δ_j is any $\Delta = \{S, T_1, \ldots, T_j\} \in \Pi(S)$, all of which yield the same ϕ_S for S.

Definition 3 *The γ-, δ-, α^*-, e-, j- and f-cores of the game* (14.3) *are, respectively, the core of the above coalitional games Γ_γ, Γ_δ, $\Gamma_{\alpha*}$, Γ_e, Γ_j and Γ_f in* (14.11–14.16), *which are precisely defined by $C_\gamma = Core(\Gamma_\gamma)$, $C_\delta = Core(\Gamma_\delta)$, $C_{\alpha*} = Core(\Gamma_{\alpha*})$, $C_e = Core(\Gamma_e)$, $C_j = Core(\Gamma_j)$, and $C_f = Core(\Gamma_f)$.*

[9] One future research topic is to combine the efficient- and j-beliefs: replace $\Delta^*_{j(S)}$ in (14.15) with $\Delta^e_{j(S)}$, or the outsiders' efficient partition among all $\Delta_{j(S)}$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Note that for the constant j-belief such that $j(s) \equiv 1$, all s, j-belief is the same as the δ-belief, so $v_j(S) = v_\delta(S)$ and $C_j = C_\delta$. For the special j-belief such that $j(s) = n - s$, all s, j-belief is the same as γ-belief, so $v_j(S) = v_\gamma(S)$ and $C_j = C_\gamma$.

Subsection 3.3 reviews the existence results for the above cores in a set of games (14.3) that are derived from oligopoly models. However, no similar results are known in a general normal form game (14.2). The relationships among the refinements (such as which is the strongest) are also unknown, with the only exception of the obvious relation that the α^*-core is the largest (i.e., $C_k \subseteq C_{\alpha*}$ holds for $k = \gamma, \delta, e, j$ and f).

3 CORE AND ITS REFINEMENTS AS CANDIDATES OF MONOPOLY SOLUTIONS

This section first reviews ten oligopoly models and the equilibrium expressions, to help readers to extend the known results in standard Cournot models to the other nine models in future research. It then reviews the known core results, including its existence as a precondition for the involved horizontal merger, its convexity and its empirical studies. It next reviews the results on core refinements, and at the end it lists seven large areas of future research in the core and its refinements in more advanced oligopoly models.

3.1 The Equilibrium Expressions in Ten Oligopoly Models

This subsection reviews ten oligopoly models and the involved equilibria,[10] which can be obtained using the inverse matrix A^{-1} in equation (16) in Zhao (2012). Because most previous studies have focused on a symmetric linear Cournot oligopoly (Cournot 1838), which is a special case of model 9 in (14.26), there is a long way to go in extending the known results to the more general cases of model 9 and then to the other nine even more general models.

A linear multi-product oligopoly with n differentiated goods is defined by three parts: (1) n cost functions $C_k(q_k) = c_k q_k$, $k \in N = \{1, \ldots, n\}$; (2) a set of multi-product firms $H = \{1, \ldots, h\}$ or a partition $\Delta = \{S_1, S_2, \ldots, S_h\}$ of N, with each firm $i \in H$ producing $n_i = |S_i|$ products in $S_i \in \Delta$ ($1 \leq n_i \leq n$, $\Sigma_{j=1}^{h} n_j = n$); and (3) a demand (in price-setting) or inverse demand (in quantity-setting) function for each of the n products, whose definitions are given below.

Let $p = (p_1, \ldots, p_n)^\top = (p_S, p_{-S}) = \{p_S |\ S \in \Delta\} = \{p_{S_j} |\ j \in H\}$ be the vector of prices, with p_k as the price of each good $k \in N$, $p_S = \{p_k |\ k \in S\}$ as the price vector of each firm $S \in \Delta$, and $p_{-S} = \{p_k |\ k \in N \backslash S\}$ as the price vector of all other firms. Similarly, $q = (q_1, \ldots, q_n)^\top = (q_S, q_{-S}) = \{q_S |\ S \in \Delta\}$ denotes the vector of products. In a price-setting oligopoly, or simply Bertrand oligopoly, or more accurately Edgeworth-Bertrand oligopoly,[11] the demand for each good $k \in S_i$ produced by each firm $S_i \in \Delta$ are

$$q_k(p) = q_k(p_{S_i}, p_{-S_i}) = V - \gamma_{kk} p_k + \gamma_i \Sigma_{m \in S_i \backslash \{k\}} p_m + \Sigma_{j \in H \backslash \{i\}} \gamma_{ij} \Sigma_{m \in S_j} p_m, \tag{14.17}$$

[10] See Amir (Chapter 3 in this *Handbook*) for a survey on related oligopoly equilibria.

[11] This title is suggested in Shubik (1980), because it was Edgeworth (1881) who originated the price-setting idea in Bertrand (1883).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

where $V > 0$ is demand size; $\gamma_{kk} > 0$, $\gamma_i, \gamma_{ij} = \gamma_{ji} \in (0, 1], k \in N$ and $i \neq j \in H$ are the substitution parameters.[12] Now, the profit for each firm $S \in \Delta$ is $\pi_S(p) = \pi_S(p_S, p_{-S}) = \Sigma_{k \in S}(p_k - c_k)q_k(p)$.

In a quantity-setting or Cournot oligopoly, the inverse demands for the products of each firm $S_i \in \Delta$ are

$$p_k(q) = \widehat{V} - \widehat{\gamma}_{kk}q_k - \widehat{\gamma}_i \Sigma_{m \in S_i \setminus \{k\}} q_m - \Sigma_{j \in H \setminus \{i\}} \widehat{\gamma}_{ij} \Sigma_{m \in S_j} q_m, \text{ all } k \in S_i, \tag{14.18}$$

where $\widehat{\gamma}_{kk}, \widehat{\gamma}_i, \widehat{\gamma}_{ij} > 0$, $k \in N$ and $i \neq j \in H$ are the parameters, and $\pi_S(q) = \pi_S(q_S, q_{-S}) = \Sigma_{k \in S}(p_k(q) - c_k)q_k$ is the profit of each $S \in \Delta$.

Strategic behavior assumes that each firm chooses a *best response*, or that it takes other firms' choices as given and chooses its choices to maximize its profit. In price setting, a strategic equilibrium or non-cooperative solution or Bertrand-Nash equilibrium (Bertrand 1883, Nash 1950) is a price vector $p^* = \{p_S^* | \ S \in \Delta\}$ such that each p_S^* solves *Max* $\{\pi_S(p_S, p^*_{-S}) | \ p_S \geq 0\}$, which is (under usual conditions) the solution of the following h sets of first-order conditions:

$$\partial \pi_{S_i}(p)/\partial p_k = 0, \text{ all } k \in S_i \text{ and for each } S_i \in \Delta, \text{or } Bp = d, \tag{14.19}$$

where $B = B_{n \times n}$ is partitioned into h^2 submatrices.[13]

In quantity setting, a strategic equilibrium or Cournot-Nash equilibrium (Cournot 1838, Nash 1950) is an output vector $q^* = \{q_S^* | \ S \in \Delta\}$ such that each q_S^* solves $Max\{\pi_S(q_S, q^*_{-S}) | q_S \geq 0\}$, or the solution of these first-order conditions:

$$\partial \pi_S(q_S, q_{-S})/\partial q_k = 0, \text{ all } k \in S \text{ and each } S \in \Delta, \text{or } \overline{B}q = \overline{d}, \tag{14.20}$$

where $\overline{B}$ has the same structure of B in (14.19).

Keep in mind that $A2$ (i.e., players are unable to take any form of coordinated or collective actions) is implicitly assumed behind the multi-product equilibria in (14.19–14.20) or the single-product equilibria in (14.21–14.22). If (14.19–14.20) are treated as the postmerger equilibria discussed below, then $A4$ (i.e., firms within each merger or players in each coalition $S \in \Delta$ are able to costlessly negotiate and enforce a joint action but players in all other coalitions $T \notin \Delta$ are unable to take any form of coordinated or collective actions) or its variations are implicitly assumed.

It is convenient to call the Bertrand or Cournot equilibrium in single-product oligopolies (i.e., $h = n$ in (14.17) and (14.18)) a *premerger equilibrium*. Precisely, a premerger Bertrand equilibrium $p^0 = \{p_i^0 | \ i \in N\}$ satisfies

---

[12] Note that internal substitution within a firm i has identical rate γ_i (i.e., between any m and $t \in S_i$), and external substitution between two firms $i \neq j$ has identical rate γ_{ij} among all of their products (i.e., between any $m \in S_i$ and $t \in S_j$). Even with such simplifications, the model is already complicated enough such that it is insolvable or precisely that the inverse of the matrix B in (14.19) is unknown and remains as an open problem.

[13] The block structure of B in (14.19) follows by rearranging the first-order conditions as

$$2\gamma_{kk}p_k - 2\gamma_i \Sigma_{m \in S_i \setminus \{k\}} p_m - \Sigma_{j \in H \setminus \{i\}} \gamma_{ij} \Sigma_{m \in S_j} p_m = V + \gamma_{kk}c_k - \gamma_i \Sigma_{m \in S_i \setminus \{k\}} c_m$$

for all $k \in S_i$ and each $S_i \in \Delta$. B contains $[n + h(h+1)/2 - h_1]$ constants, where h_1 is the number of singleton coalitions (i.e., single-product firms) in Δ. See equation (1) in Zhao (2012) for details.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

$$p_i^0 \in ArgMax\{\pi_i(p_i, p_{-i}^0)|p_i \geq 0\}, \text{ all } i, \tag{14.21}$$

and a premerger Cournot equilibrium $q^0 = \{q_i^0|\ i \in N\}$ satisfies

$$q_i^0 \in ArgMax\{\pi_i(q_i, q_{-i}^0)|q_i \geq 0\}, \text{ all } i. \tag{14.22}$$

In this regard, the general Bertrand equilibrium or the solution $p = B^{-1}d$ of (14.19) (assuming the inverse exists) can be called the postmerger equilibrium for Δ. This leads to three kinds of mergers given by the following models 1–3, respectively:[14]

$$\begin{array}{ll} (1) & \pi_S(p) = \Sigma_{k \in S}(p_k - c_k)q_k(p), \text{ all } S \in \Delta; \\ (2) & \pi_S(p) = \Sigma_{k \in S}(p_k - c_S)q_k(p), \text{ all } S \in \Delta; \\ (3) & \gamma_i = Max\{\gamma_{km}|k \neq m \in S_i\}, \gamma_{ij} = \gamma_{ji} \equiv Min\{\gamma_{km}|k \in S_i, m \in S_j\}; \\ (4) & q_k(p) = V - p_k + \gamma\,\Sigma_{m \neq k} p_m, \text{ all } k \in N; \\ (5) & q_k(p) = V - p_k - \gamma(p_k - \overline{p}), \text{ all } k \in N, \end{array} \tag{14.23}$$

where $c_S = Min\{c_k|\ k \in S\}$ in model 2 is the smallest marginal cost of each merger or each multi-product firm $S \in \Delta$, and $\overline{p} = (\Sigma p_k)/n$ in model 5 is the average price.

In model 1, the profits of each merger $S \in \Delta$ are simply the sum of its members' profit (i.e., $\pi_S(p) = \Sigma_{k \in S}\pi_k(p)$), so it represents a set of *simultaneous mergers without synergy*, from the premerger equilibrium p^0 in (14.21). On the other hand, model 2 represents *simultaneous mergers with weak cost-synergy*, because each merger $S \in \Delta$ can use its smallest marginal cost $c_S = Min\{c_k|\ k \in S\}$ in producing all its outputs, or precisely $\pi_S(p) = \Sigma_{k \in S}(p_k - c_S)q_k(p)$. Finally, model 3 represents *simultaneous mergers with marketing-synergy*, because marketing outcomes such as

$$\gamma_i = Max\{\gamma_{km}|k \neq m \in S_i\} \tag{14.24}$$

increases the demands for internal products, and

$$\gamma_{ij} = \gamma_{ji} \equiv Min\{\gamma_{km}|k \in S_i, m \in S_j\} \tag{14.25}$$

reduces the demand for competitors' products.

Model 4 in (14.23) is called Dixit demand (1979) and is a special case of (14.17) when $h = n$ (or there is no γ_i), $\gamma_{kk} \equiv 1$ and $\gamma_{ij} \equiv \gamma$ $(i \neq j, k \in N)$; model 5 is called Shubik demand (1980) and is another special case of (14.17), when $h = n$, $\gamma_{kk} \equiv [n + (n-1)\gamma]/n$, $\gamma_{ij} \equiv \gamma/n$, $i \neq j, k \in N$. Thus, the Dixit and Shubik demands are two models of *premerger Bertrand equilibria*.

[14] To facilitate future studies, the models are arranged in the same order as in Zhao (2012), which can be expanded to include strategic complements surveyed in Vives (Chapter 2 in this *Handbook*).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

The quantity-setting or Cournot equivalents of models 1–5 in (14.23) are:

$$\begin{array}{rl}
(6) & \pi_S(q) = \Sigma_{k\in S}(p_k(q) - c_k)q_k,\ \text{all } S \in \Delta; \\
(7) & \pi_S(q) = \Sigma_{k\in S}(p_k(q) - c_S)q_k,\ \text{all } S \in \Delta; \\
(8) & \widehat{\gamma}_i = Min\{\widehat{\gamma}_{km}|k \neq m \in S_i\},\ \widehat{\gamma}_{ij} = \widehat{\gamma}_{ji} \equiv Max\{\widehat{\gamma}_{km}|k \in S_i, m \in S_j\}; \\
(9) & p_k(q) = \widehat{V} - q_k - \widehat{\gamma}\Sigma_{m\neq k}q_m,\ \text{all } k \in N; \\
(10) & p_k(q) = \widehat{V} - q_k + \widehat{\gamma}(q_k - \overline{q}),\ \text{all } k \in N,
\end{array} \tag{14.26}$$

where $\overline{q} = \sum q_j/n$ in model 10 is the average output, and the synergy in model 8 reduces (increases) the negative effects of an output increase on own (rivals') profits. The details of (14.26) are similar to those of (14.23) and are thus skipped.

It is useful to note the following three remarks. First, most Cournot models in (14.26) and Bertrand models in (14.23) can be inverted from each other. For example, inverting model 5 in (14.23) yields the inverse demand in model 10 in (14.26), with $\widehat{V} = V$ and $\widehat{\gamma} = \gamma/(1+\gamma)$. Second, the Shubik demand or model 5 has an intuitive interpretation: a firm k that charges more (less) than the average price will be penalized (rewarded) by an amount equal to $\gamma|p_k - \overline{p}|$. This intuition is the reason why the Shubik demands or models 5 and 10 are used in the three mergers in (14.23) and (14.26) (and in Lemma 3 below). If one uses the Dixit demands or model 4 and model 9, one will get six additional merger models. Thus (14.23) and (14.26) actually provide a total of 16 oligopoly models (eight models each in both Cournot and Bertrand competitions).[15]

Third, there are two main reasons why the supermajority of previous studies have focused on quantity competition or Cournot oligopoly: (1) the expressions of known Bertrand equilibria are in general more involved and less tractable than the Cournot equilibria, and (2) the expressions of many Bertrand equilibria are unknown. Although the inverse B^{-1} in (14.19) is unknown in the general cases and remains as an open mathematical problem, the partial solution or the inverse A^{-1} in equation (16) of Zhao (2012) is sufficient to yield the equilibrium $p = A^{-1}d$ in most linear oligopolies that are relevant for empirical or theoretical studies. This matrix $A =$

$$A_{n\times n} = \begin{pmatrix} A_{11} & \cdots & A_{1h} \\ \vdots & \ddots & \vdots \\ A_{h1} & \cdots & A_{hh} \end{pmatrix} \tag{14.27}$$

has the same structure of B in (14.19). It is a very small class of B in that it reduces $n(n-1)/2$ constants in B in (14.19) to only three constants in A in (14.27): A_{ii}, $i = 1, \ldots, h$, is an $n_i \times n_i$ square matrix whose diagonal entries are a constant a and other entries a constant $-b$, and A_{ij}, all $i \neq j$, is an $n_i \times n_j$ matrix of a constant $-c$, or precisely $\gamma_{kk} \equiv a/2$, $\gamma_i \equiv b/2$ and $\gamma_{ij} \equiv c$ for all k, i and j in (14.17).

[15] Dixit demand has the advantage of being the solution to a simple utility maximization problem. Let $I_{n\times n}$ be the identity matrix, $E_{n\times n}$ the matrix of ones, $G = (1-\gamma)I_{n\times n} + \gamma E_{n\times n}$, and $U(q, y) = y + V\Sigma q_m - q^\top Gq/2$ the utility, where y is a composite measure of all other consumptions. Then, $Max\left\{U(q, y)|p^\top q + y \leq Y\right\}$ yields the inverse version of model 4 in (14.23), or model 9 in (14.26), where Y is fixed income. In addition, the consumer surplus is equal to $CS = q^\top Gq/2$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

The next lemma provides, as an example of $p = A^{-1}d$, the postmerger Bertrand equilibrium without synergy for a single merger $S = \{1, \ldots, t\}$, or precisely the equilibrium for $\Delta = \{S, t+1, \ldots, n\}$ in model 1 in (14.23) with Shubik demand or model 5. Without loss of generality, assume (1) $c_1 \leq c_2 \leq \ldots \leq c_t$, so $c_S = c_1$; and (2) the following assumption ($A0$) holds. $A0$ guarantees a positive output for all firms at both premerger and postmerger equilibria:[16]

A0 (Assumption 0) *For each* $S = \{1, \ldots, t\}$,

$$\frac{nV + (n + (n-t)\gamma)\overline{c_S} + \gamma(n-t)\overline{c_{-S}}}{(2n + (2n - 2t)\gamma)} > \overline{c_S} \tag{14.28}$$

holds, where $\overline{c_S} = \Sigma_{i \in S} c_i / t$ *and* $\overline{c_{-S}} = \Sigma_{j \notin S} c_j / (n-t)$.

Lemma 3 *Let* p^* *be the postmerger equilibrium for* $S = \{1, \ldots, t\}$ *in model 1* with *Shubik demand,* $\overline{c_S}$ *and* $\overline{c_{-S}}$ *be given in* (14.28). *Then, for each* $k \in S$, $j \in N \backslash S$,

$$\begin{aligned} p_k^* &= \frac{n(2n(1+\gamma) - \gamma)V}{\omega_0} + \frac{\gamma^2 t\,(n-t)\,\overline{c_S}}{2\omega_0} + \frac{(n-t)\gamma(n(1+\gamma) - \gamma)\overline{c_{-S}}}{\omega_0} + \frac{c_k}{2}, \\ p_j^* &= \frac{n(2n(1+\gamma) - t\gamma)V}{\omega_0} + \frac{t\gamma\,(n(1+\gamma) - t\gamma)\,\overline{c_S}}{\omega_0} \\ &\quad + \frac{\gamma\,(n-t)\,(n(1+\gamma) - \gamma)\,(2n(1+\gamma) - t\gamma)\,\overline{c_{-S}}}{(2n(1+\gamma) - \gamma)\omega_0} + \frac{(n(1+\gamma) - \gamma)c_j}{2n(1+\gamma) - \gamma}, \end{aligned} \tag{14.29}$$

where $\omega_0 = \gamma^2\,(n-t)\,(t + 2n - 2) + 2n\gamma\,(3n - t - 1) + 4n^2$.

The above expressions become the premerger equilibrium when $t = 1$ in (14.29); the postmerger equilibrium with weak synergy in model 2 when $c_k = c_1 = c_s$, all $k \in S$; and the postmerger equilibrium with Dixit demand when V in (14.29) is replaced by $V/\,(1 + \gamma - n\gamma)$ and γ by $n\gamma/(1 + \gamma - n\gamma)$.[17]

3.2 The Core in Oligopoly TU Games

A homogeneous Cournot oligopoly is given by an inverse demand $p(\Sigma q_j)$ and n cost functions $C_i(q_i)$, $0 \leq q_i \leq z_i$, with $z_i > 0$ as firm i's capacity, or by a normal form game $\Gamma = \{N, X_i, u_i\}$ in (14.2) in which

$$u_i = \pi_i(q) = p(\Sigma q_j)q_i - C_i(q_i),\ X_i = [0, z_i], \text{ all } i. \tag{14.30}$$

Under the usual conditions of a Cournot oligopoly such as decreasing demand and continuity, both weak and strong separability in Lemma 2 hold because the outsiders' choices

[16] This is an extension of the conditions in a single-product Cournot oligopoly in Zhao (2001a). See Pham Do and Folmer (2003) for discussion and Zhao (2009a, p. 377) for an application.

[17] This follows from $q_k = V - p_k + \gamma\,\Sigma_{m \neq k} p_m = (1 + \gamma - n\gamma)\left[\frac{V}{1+\gamma-n\gamma} - p_k - \frac{n\gamma\,(p_k - \overline{p})}{(1+\gamma-n\gamma)}\right]$, and the observation that the first term $(1 + \gamma - n\gamma)$ does not enter the first-order conditions. See footnote 9 in Zhao (2012) for details.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

in both $v_\alpha(S)$ in (14.6) and $v_\beta(S)$ in (14.7) are equal to their full capacity at $x_{-S} = q_{-S} = z_{-S} = \{z_j | j \notin S\}$. This implies $v_\alpha(S) = v_\beta(S)$ and $Core(\Gamma_\alpha) = Core(\Gamma_\beta)$. Therefore, there is no need to make the α- and β-distinction in oligopoly models, and one can simply use the term core, which will be non-empty under the additional assumption that each $\pi_i(q)$ is concave in $q = (q_1, \ldots, q_n)$. Such results are first obtained by the author and are given here:

Proposition 1 (Zhao 1999a, p. 160) *Let C_α and C_β be the α- and β-cores of an oligopoly* Γ *in* (14.30). *Assume $p(\Sigma q_j)$ is decreasing and each $\pi_i(q)$ is continuous. Then, (i) $C_\alpha = C_\beta = C(\Gamma)$; and (ii) $C(\Gamma) \neq \varnothing$ if each $\pi_i(q)$ is concave.*

Although concavity in part *(ii)* is a strong condition,[18] it can be weakened in large classes of oligopolies such as the linear version of (14.30) in (14.34) in subsection 3.2.2.

3.2.1 Non-empty core as a precondition for horizontal mergers

A fundamental role of the core theory in industrial organization is that a non-empty core is a precondition for horizontal mergers. This argument is summarized in the next proposition. For simplicity, define a monopoly merger contract in an oligopoly (14.30) as a triplet $(N, \overline{q}, \theta)$ of the set of firms N, monopoly supply $\overline{q}$ and a split of monopoly profits θ (i.e., $\overline{q} \in ArgMax\{\Sigma_{j=1}^{n} \pi_j(q) | q_j \in [0, z_j], \text{all } j\}$, $\theta \geq 0$ and $\Sigma\theta_j = \Sigma\pi_j(\overline{q}) = \pi^m = v(N)$).

Proposition 2 (Two preconditions for the monopoly merger, Zhao 2009a, p. 378) *Let q^0 be the premerger equilibrium in* (14.30), *$C(\Gamma)$ its core, and $(N, \overline{q}, \theta)$ an observed monopoly merger. Then, (i) $\theta_j \geq \pi_j^0 = \pi_j(q^0)$, all j; and (ii) $\theta \in C(\Gamma)$.*

Part *(i)* is the well-known profitability precondition (or incentive to merge), and part *(ii)* is the non-empty core precondition. The merger would have made at least one firm worse off (i.e., a firm *j* gets less than its premerger profits π_j^0) if part *(i)* fails, and at least one coalition worse off (i.e., a coalition *S* gets less than its guaranteed or worst profits $v(S)$) if part *(ii)* fails. Therefore, the failure of either precondition will violate a firm's or a coalition's rationality, so both must hold in successful mergers.

Keep in mind that these are necessary, rather than sufficient, conditions for a monopoly merger. Failing either or both will result in a merger failure, and meeting both will not guarantee a merger success. In addition, they are independent of each other: Example 1 in the next subsection reports a profitable monopoly with an empty core, and Example 2 an unprofitable monopoly with a non-empty core.

These preconditions make it possible to study how merging costs or the transaction costs of a merger affect merger formation and how to empirically estimate the sizes of such merging costs. Let $mc(S) \geq 0$ denote the merging cost of each merger $S \subseteq N$. For simplicity, assume $mc(S) = 0$ for all non-monopoly merger $S \neq N$, to focus on the monopoly merging cost $mmc = mc(N) \geq 0$. The next proposition shows that

$$mmc^* = \pi^m - Max\left\{\Sigma\pi_j^0, mnbp\right\} \text{ and} \tag{14.31}$$

$$mmc^0 = \pi^m - Min\left\{\Sigma\pi_j^0, mnbp\right\} \tag{14.32}$$

[18] Continuity can also be weakened, in the same manner of Uyanik (2015) on the TU α-core in a normal form game (14.2).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

are, respectively, an upper bound of the merging cost for a successful monopoly merger and a lower bound for a failed or unobserved monopoly merger, where π^m and π_j^0 are the same monopoly and premerger profits as in Proposition 2, and *mnbp* is given in (14.5) for the oligopoly game (14.30).

Proposition 3 (Zhao 2009a, 2009b) *Given a monopoly merger in* (14.30)*, let* mmc^* *and* mmc^0 *be given in* (14.31–14.32)*. Then, (i)* $mmc \le mmc^*$ *if the merger is successful; and (ii)* $mmc > mmc^0$ *if the merger is prevented by failed preconditions.*

3.2.2 The core with weak synergy

The following assumption ($A0.1$) modifies the concept of weak synergy in the oligopoly (14.30). Though the synergy such as c_S in (14.23) and (14.26) might be quite large in reality, it is called *weak synergy* in Farrell and Shapiro (1990) for comparison with strong synergies involving economies of scale:

A0.1 *(i) Each* $\pi_i(q)$ *is continuous in* q *and quasi-concave in* q_i*, and* $p(\Sigma q_j)$ *is decreasing; (ii) all equilibria are positive and interior solutions; and (iii) the capacity and cost function for each merger* S *are*

$$z_S = \Sigma_{j\in S} z_j,\ C_S(y) = Min\{\Sigma_{j\in S} C_j(q_j) | y = \Sigma_{j\in S} q_j \le z_S, q_S \ge 0\}. \tag{14.33}$$

Most results reviewed in this chapter deal with a linear (14.30) or a subset of model 9 with $\widehat{\gamma} = 1$ and capacities in (14.26): $p(\Sigma q_j) = a - \Sigma q_j$, $C_i(q_i) = c_i q_i$, $q_i \in [0, z_i]$, which can be given by a $(2n + 1)$-vector

$$(a, c, z) \in R_+^{2n+1}, \text{ where } c = (c_1, \ldots, c_n) \text{ and } z = (z_1, \ldots, z_n) \tag{14.34}$$

are the vectors of marginal costs and capacities, and $a > 0$ is the intercept of the demand. Without loss of generality, assume $c_1 \le \ldots \le c_n < a$. Then, above $A0.1$ becomes for all S, $C_S(y) = c_S y$, $y \le z_S = \Sigma_{j\in S} z_j$,

$$0 < (a - c_S - z_{-S})/2 \le z_S \text{ and } (a - c_1)/2 \le z_N = \overline{z} = \Sigma_{j\in N} z_j, \tag{14.35}$$

where $z_{-S} = \Sigma_{j\notin S} z_j$, and $c_S = Min\{c_k | \ k \in S\}$ is the same as in (14.23) and (14.26).

A symmetric linear Cournot oligopoly is the case when $c_i = c$ *and* $z_i = z$*, all* i, or $(a, c, z) \in R_+^3$.[19] In such symmetric cases, the conditions in (14.35) become

$$0 < (a - c)/(n + 1) \le z \ \le (a - c)/(n - 1), \tag{14.36}$$

[19] The same letter c is used here as a scalar and in (14.34) as a vector. This should cause no confusion because the meaning will be clear in the contexts. Similar simplification holds for letter z.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

which leads to, as shown in Zhao (2009a), $mnbp = n(a-c-z)^2/[4(n-1)] < v(N) = (a-c)^2/4$, so the core is not only non-empty but also has a non-empty (relative) interior. Such core results are summarized in the next proposition:

Proposition 4 (Zhao 2009a, p. 381) *Let* $mmc \geq 0$ *be the monopoly merging cost in* (14.34), $C(\Gamma)$ *its core, and assume parts (ii–iii) of* A0.1 *or* (14.35). *Then, (i) the core has a non-empty (relative) interior if* $mmc = 0$*; (ii)* $C(\Gamma) \neq \varnothing$ *if*

$$mmc \leq (a-c_1)^2/4 - \{n(a-c_1-z_{\min})^2/[4(n-1)]\}; \text{ and} \tag{14.37}$$

(iii) in symmetric case with $c_i = c$ *and* $z_i = z$*, all i,* $C(\Gamma) \neq \varnothing$ *if and only if*

$$mmc \leq (a-c)^2/4 - \{n(a-c-z)^2/[4(n-1)]\}, \tag{14.38}$$

where $c_1 = Min\{c_k |\ k \in N\}$ *and* $z_{\min} = Min\{z_k |\ k \in N\}$ *are the minimal marginal cost and minimal capacity.*

The core's interior has important implications in empirical studies. In the event of small shocks to the market, a core with a non-empty interior remains non-empty, but a non-empty core with an empty interior could become empty. Thus, a long-lived merger or trust suggests that the core has a non-empty interior, and short-lived ones suggest either an empty-core or a non-empty core with an empty interior, which are the causes for both merger failure and the breakup of completed mergers such as the breakup of AOL-Time Warner.

The following two examples (Zhao 2009a) show the independence of the two merger preconditions; they also illustrate part *(iii)* in Proposition 4.

Example 1 $n = 3, (a, c, z) = (6, 0.8, 1.5)$; *or* $p = 6 - \Sigma x_j, C_i(x_i) = 0.8x_i, 0 \leq x_i \leq 1.5, i = 1, 2, 3$*; and* $mmc = 1.65$. *The premerger and monopoly profits are* $\pi_i^0 = 1.69$, $\pi^m = 6.76$*, and* $mnbp = 5.13$. *By* $v(123) = \pi^m - mmc = 5.11 > \Sigma\pi_i^0 = 5.07$*, the merger is profitable.*[20] *By* (14.38) *and by* $mmc = 1.65 > (a-c)^2/4 - n(a-c-z)^2/[4(n-1)] = \pi^m - mnbp = 1.63$*, the core is empty.*

Example 2 $n = 3, p = 6 - \Sigma x_j, C_i(x_i) = 0.5x_i, 0 \leq x_i \leq 2$*, all i, and* $mmc = 2$. *The premerger and monopoly profits are* $\pi_i^0 = 1.89$, $\pi^m = 7.56$*, and* $mnbp = 4.59$. *The core is non-empty because* $mmc = 2 < \pi^m - mnbp = 2.97$*, and the merger is not profitable because* $v(123) = 5.56 < \Sigma\pi_i^0 = 5.67$.

The next proposition shows how excessive capacity affects the estimated bound of monopoly merging costs. Let $\tau \geq 0$ be the rate of excessive capacity as defined in

$$z = (1+\tau)(a-c)/(n+1), \tag{14.39}$$

so $\tau = 0$ means full capacity at premerger equilibrium: $q_i^0 = (a-c)/(n+1) = z$, all i.

[20] Coalition $\{1, 2, 3\}$ is simplified as 123. Similar simplifications hold in other places where no confusion arises.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Proposition 5 (Zhao 2009a, p. 383) *Let* $\tau_1 = n - 2\sqrt{n\text{-}1}$*, and* mmc^* *and* τ *be given in* (14.31) *and* (14.39). *Then,* $mmc^* = n(a-c)^2\,(n-1)^2\,/[4n\,(n+1)^2]$ *if* $\tau \geq \tau_1$*, and*

$$mmc^* = \frac{n(a-c)^2}{(n+1)^2}\,\frac{(n-1)\,(n+1)^2 - n\,(n-\tau)^2}{4n(n-1)}\ \text{if}\ \tau < \tau_1. \tag{14.40}$$

The above results indicate that a larger capacity will strengthen the non-empty core precondition, so the monopoly merger is more likely to be formed. This is consistent with and thus provides a new explanation for the stylized fact that mergers are likely to occur in markets plagued by excess capacities.

3.2.3 The convexity in oligopoly games

Convex games or supermodular set functions are interesting in both economics and mathematics and have generated a large literature. Only the less technical results in oligopoly TU games are reviewed here. For non-technical readers, it is sufficient to know three conclusions in a convex oligopoly game: (1) a convex game exhibits increasing returns to scale in coalition size or the property that each i's marginal contribution to a coalition increases as the coalition expands, so there is an incentive to get larger and eventually form the grand coalition; (2) the core is non-empty; and (3) both the nucleolus (Schmeidler 1969) and Shapley value (Shapley 1953) are perfect answers to the question of how to split the monopoly profits.[21]

Definition 4 *The game* Γ *in* (14.1) *or* (14.9) *is convex if for any* $S, T \subseteq N$,

$$v(S) + v(T) \leq v(S \cap T) + v(S \cup T). \tag{14.41}$$

Assume the conditions in (14.35) for a linear oligopoly (14.34), one has $v(S) + v(T) \leq v(S \cup T)$ for any $S \cap T = \varnothing$ (see Theorem 1 in Zhao 1999b), or that the oligopoly game Γ for (14.34) is superadditive. Because convexity in (14.41) implies the preceding superadditivity, convex games are stronger than superadditive games.

The first main result in convex oligopoly games is a necessary and sufficient condition for an oligopoly (14.34) reported in Zhao (1999b),[22] which has been extended along several directions in Norde, Pham and Tijs (2002), Driessen and Meinhardt (2005, 2010), Lardon (2010), and Hou, Driessen and Lardon (2011). Let

$$\Omega = \{(S, T, i) | S \subset T \subset N, i \in N \backslash T \text{ and } c_S - c_{S \cup i} > c_T - c_{T \cup i}\} \tag{14.42}$$

[21] See Driessen and Meinhardt (2010), Meinhardt (2002, 2013), Vives (Chapter 2 in this *Handbook*) and Zhao (1999b) for reviews. In such cases, the nucleolus coincides with both the pre-kernel and kernel and thus satisfies additional nice properties. See Meinhardt (2013, p. 32) for more discussion.

[22] The linear model in Zhao (1999b) contains fixed costs and is given by a $(3n+1)$-vector $(a, c, d, z) \in R_+^{3n+1}$, or $p(\Sigma q_j) = a - \Sigma q_j$, $C_i(q_i) = d_i + c_i q_i$, $q_i \in [0, z_i]$, with $d = (d_1, \ldots, d_n)$ as the vector of fixed costs and (a, c, z) the same as in (14.34). Because fixed costs have no effects on convexity, this review sticks with (a, c, z) or assumes $d = 0$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

denote the set of coalitions whose marginal costs exhibit strict supermodularity.[23] It is not difficult to show that the game is convex if $\Omega = \varnothing$ (see Theorem 2 in Zhao 1999b) . If $\Omega \neq \varnothing$, for each $(S, T, i) \in \Omega$, define

$$\begin{aligned} f(S,T,i) &= c_S^2 - c_{S\cup i}^2 - (c_T^2 - c_{T\cup i}^2) + 2z_i(c_S - c_T + \Sigma_{j\in T\setminus S} z_j) \qquad (14.43)\\ &\quad +2[(c_S - c_{S\cup i})\Sigma_{j\notin S, j\neq i} z_j - (c_T - c_{T\cup i})\Sigma_{j\notin T, j\neq i} z_j],\\ F(S,T,i) &= f(S,T,i)/[2(c_S - c_{S\cup i} - (c_T - c_{T\cup i}))], \text{ and}\\ \omega &= Min\{F(S,T,i) | (S,T,i) \in \Omega\}. \qquad (14.44) \end{aligned}$$

Under the conditions in (14.35), one has $\omega > 0$ (see Lemma 5 in Zhao 1999b). Although the economic meaning of ω is still not well understood, it nevertheless fully characterizes the convexity.

Proposition 6 (Zhao 1999b, p. 195) *Given* (a, c, z) *in* (14.34) *and* ω *in* (14.44)*, assume* (14.35)*. Then,* Γ *in* (14.9) *is convex if and only if* $a \leq \omega$*.*

Example 3 (Zhao 1999b) $n = 3, (a, c, z) = (7, 4, 2.25, 2.25, 1.3, 1.3, 1.3)$. *One has* $\omega = 6.6907$, $v(1) = 0.04$, $v(2) = v(3) = 1.1556$, $v(12) = v(13) = v(23) = 2.9756$, $v(123) = 5.6406$. *By Proposition* 6 *and* $\omega < a = 7$*, the game is not convex. Indeed,* (14.41) *fails for* $S = 12, T = 13$: $v(S) + v(T) = 5.9512 > v(S \cup T) + v(S \cap T) = 5.6806$. *Let a be decreased to* $a = 6.65$*, all other parameters be unchanged, one has the same* $\omega = 6.6907$ *and new coalitional values*: $v(1) = 0.0006$, $v(2) = v(3) = 0.81$, $v(12) = v(13) = v(23) = 2.4026$ *and* $v(123) = 4.84$. *By* $a = 6.65 < \omega$*, the game is now convex. Indeed, one can verify that* (14.41) *holds for all S and T.*

Norde, Pham and Tijs (2002) extend the above results to mergers without weak synergy (or *without transferable technologies*). They show that such oligopoly games are always convex. Precisely, let $\pi_S(q_S, z_{-S}) = \Sigma_{k\in S}[p(\Sigma_{j\in S} q_j + \Sigma_{j\notin S} z_j) - c_k]q_k$ be the same as in model 6 in (14.26), and a coalition's payoff be given by

$$v(S) = v_\alpha(S) = v_\beta(S) = Max\{\pi_S(q_S, z_{-S}) | q_k \in [0, z_k], \text{ all } k \in S\}. \qquad (14.45)$$

Proposition 7 (Norde, Pham and Tijs 2002, p. 203) *Given* (a, c, z) *in* (14.34)*, let* $v(S)$ *in* Γ *in* (14.9) *be given in* (14.45)*. Then* Γ *is convex.*

Driessen and Meinhardt (2005) re-establish Proposition 7 using a new technique that has an economic interpretation. This effective technique allows them to obtain more general convex games with weak synergy without assuming interior solutions. This is summarized in the next proposition:

Proposition 8 (Driessen and Meinhardt 2010, p. 330) *Given* (a, c, z) *in (*14.34)*, assume part (iii) of* A0.1*, and let* $\underline{\Omega}$ *and* $\underline{\omega}$ *be defined as the special case of* Ω *and* ω *in* (14.42–14.44) *when* $T = S \cup j$ *for all* $j \notin T \cup i$*. Then,* Γ *is convex if* $a \leq \underline{\omega}$*.*

[23] A coalition's marginal costs exhibit supermodularity if $c_{S\cup i} - c_S \leq c_{T\cup i} - c_T$ for $S \subset T$ and $i \notin T$. Thus, $(S, T, i) \in \Omega$ in (14.42) implies strict supermodularity or $c_{S\cup i} - c_S < c_{T\cup i} - c_T$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Lardon (2010) extends Propositions 1 and 7 to Bertrand competition and establishes this result: in a symmetric linear Shubik model or in a symmetric model 5 in (14.23), the α-core and β-core are identical and convex. Hou, Driessen and Lardon (2011) give further extensions that similar results hold in an asymmetric Shubik model under reasonable assumptions.

3.3 Refinements of the Core

Refining the core in an oligopoly provides deeper understandings about monopoly stability and helps search for the sufficient conditions of monopoly formation. Such refinements are achieved in two steps: (1) As Lekeas (2013, p. 2) puts it, "Convert the oligopoly (or normal form game) to a partition function game in Thrall and Lucas (1963) by computing the quasi-hybrid solution (each coalition chooses an efficient solution, given others' choices) for each partition in Zhao (1991)"; and (2) study one of the core solutions defined in Definition 3 or other core solutions for the converted partition function game (14.3).[24]

Conversions in the first step implicitly assume that the following simpler version of $A4$ holds for each partition:

A4.0 (Assumption 4.0) *Given a partition* $\Delta = \{S_1, S_2, \ldots, S_h\}$, $A1$ *holds for each* $S \in \Delta$, *and* $A2$ *holds for each* $T \notin \Delta$ *such that there are* $i \neq j$, $T \cap S_i \neq \varnothing$ *and* $T \cap S_j \neq \varnothing$.

Because each coalition $S \in \Delta$ in (14.19) or (14.20) chooses an efficient solution, the same as that under $A4.0$, the solutions in (14.19–14.20) have been called quasi-hybrid solutions (Zhao 1991), as compared with hybrid solutions[25] under $A4$ in which each $S \in \Delta$ chooses a core solution. Such quasi-hybrid solutions are, as pointed out in Zhao (1991), the same as the non-cooperative solution of Shapley (1956, 1959) for the following multi-objective game (*MOG*):

$$\Gamma_m = \{H, Y_j, v_j\}, \tag{14.46}$$

where $H = \{1, \ldots, h\} =: \{S_1, S_2, \ldots, S_h\}$ is the set of new players (i.e., new names for coalitions in $\Delta = \{S_1, S_2, \ldots, S_h\}$), $Y_j = Y_{S_j} = \Pi_{i \in S_j} X_i$ (in (14.2) or (14.30)) is j's choice set, and $v_j = v_{S_j} = \{u_i | i \in S_j\}$ in (14.2) (or $= \{\pi_i | i \in S_j\}$ in (14.30)) is j's vector payoff function. Thus, the solutions in (14.19–14.20) should be cited either as a non-cooperative solution of the *MOG* (14.46) or a quasi-hybrid solution for (14.2) and (14.30).[26]

Given a linear oligopoly in model 4 or 5 in (14.23) or model 9 or 10 in (14.26), the inverse A^{-1} of A in (14.27) readily yields eight classes of postmerger equilibria (i.e., using model 4 or 5 with model 1 or 2, and model 9 or 10 with model 5 or 6) for each partition $\Delta = \{S_1, S_2, \ldots, S_h\}$, and thus leads to eight partition function games, only one of which (i.e., symmetric linear Cournot with no synergy or models 6 and 9) is well studied.

---

[24] As already discussed in subsection 2.3, an exception is the leader–follower core or *lf*-core (Currarini and Marini 2003), see Proposition 20 at end of this subsection for an *lf*-core result.

[25] See Allen (2000, p. 147), Diamantoudia and Xue (2007, p. 108) and McCain (2008) for discussion about the significance and generality of hybrid solutions.

[26] Chander and Tulkens (1997) study a class of normal form games. They gave such quasi-hybrid equilibria a new name without citing Shapley (1956) or Zhao (1991). See Folmer and Mouche (1994) and Zhao (2018) for more discussions about the connection between *MOG* and the solutions in (14.19–14.20).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

The following results are obvious and can be found in the discussions in Yong (2004) and Zhao (1996, or 2013).

Lemma 4 *Given* Γ_p *in* (14.3) *for a symmetric* (a, c, z) *in* (14.34)*, assume* $a > c$ *and* $z_i = z = \infty$*. Let* $C = C_\alpha = C_\beta$ *be its core in* (14.9)*, and* C_γ*,* C_δ*,* $C_{\alpha *}$ *and* C_e *be its* γ*-,* δ*-,* α^**- and e-cores in Definition* 3*. Then,* $C_\delta \subseteq C_e \subseteq C_\gamma \subseteq C_{\alpha *} \subseteq C$*.*

The first major refinement of the core in oligopoly games is Yong's characterization of his *e*-core given below:

Proposition 9 (Yong 2004, p. 10) *Given* Γ_e *in* (14.14) *for a symmetric* (a, c, z) *in* (14.34)*, assume* $a - c > 0$ *and* $z_i = z = \infty$*. Then, (i)* $\theta \in C_e = Core(\Gamma_e) \iff \Sigma\theta_j = v(N)$ *and* $\theta_j \geq v(j) \equiv v(1)$*, all j; and (ii)* $C_e \neq \varnothing \iff n \leq 4$*.*

Thus, in a standard symmetric linear Cournot oligopoly, the *e*-core and imputation set are identical, monopoly can possibly be formed under efficient-belief with four or fewer firms, and will not be formed under efficient-belief with five or more firms.

The next proposition is an extension to diseconomies of scale given by $C_k(q) \equiv C(q) = cq + dq^2$. By $MC/AC = (c + 2dq)/(c + dq) > 1$, average cost is increasing so there are diseconomies of scale, and it is cheaper for a merger of m firms to produce a smaller quantity $\Sigma_{k \in S} q_k/m$ at each of its m plants than to produce the sum $\Sigma_{k \in S} q_k$ at one large plant (i.e., $C_S(\Sigma_{k \in S} q_k) = mC(\Sigma_{k \in S} q_k/m) < C(\Sigma_{k \in S} q_k)$).

Proposition 10 (Yong 2004, pp. 21–24) *Given* $p = a - \Sigma q_j$ *and* $C_k(q) = cq + dq^2$*, all k, assume part (iii) of* A0.1 *or* $\pi_S(q) = p(\Sigma_{j \in N} q_j)\Sigma_{k \in S} q_k - C_S(\Sigma_{k \in S} q_k), S \in \Delta$ *for all* Δ*. Let* $n^0 = [5 + 5d - 2d^2 + (1 + d)\sqrt{(1 + 2d)(9 + 2d)}]/[2(1 + 2d)]$*. Then,* $C_e \neq \varnothing \iff n \leq n^0$*.*

Below is another interesting and non-trivial extension in Yong (2004), which studies a Bertrand-Shubik model with zero costs (i.e., model 5 in (14.23) with zero costs):

Proposition 11 (Yong 2004, p. 14) *Given* $\pi_S(p) = \Sigma_{k \in S} p_k[V - p_k - \gamma(p_k - \overline{p})]$, $S \in \Delta$ *for all* Δ*, let* $\gamma^0 = 2n[(n-2)^2 - 3 + (n-2)\sqrt{(n-2)^2 + 3}]/[9(n-1)]$*. Then,* $C_e \neq \varnothing \iff \gamma \geq Max\{0, \gamma^0\}$*.*

Yong (2004) shows that the conditions of Proposition 11 always hold if $n \leq 3$. Readers can find other extensions in Yong (2004) such as with capacity constraints. Proposition 11 appears to imply the *r*-core result in Huang and Sjostrom (2003):[27]

Corollary 1 (Huang and Sjostrom 2003, p. 208) *In the same oligopoly of Proposition* 11*, r-core* $\neq \varnothing \iff \gamma \geq \widehat{\gamma}(n)$*, where for* $i = 1, \ldots, 7$*,* $\widehat{\gamma}(i) = Max\{0, \gamma^0(i)\}$*,* $\widehat{\gamma}(8) = 19$*,* $\widehat{\gamma}(9) = 43.75$*, and* $\widehat{\gamma}(n) = \infty$*, all* $n \geq 10$*.*

[27] The concept of *r*-core in Huang and Sjostrom (2003) is confusing because it is not clear whether the cooperative or the non-cooperative approach to cooperation is used. The relation of *r*-core with other cores is unknown because none of the known α- or β-core results are cited in the paper. Based on the facts that $\gamma^0(n)$ generates most of the values of $\widehat{\gamma}(n)$, the author suspects that some values of $\widehat{\gamma}(n)$ in Huang and Sjostrom (2003) (i.e., some $n \geq 8$) are incorrect, and conjectures that the *e*-core and *r*-core are identical. The author thanks Giorgos Stamatopoulos for bringing this issue to his attention.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Lardon (2012) provides the second major contribution in refining the core:

Proposition 12 (Lardon 2012, p. 403) *Let C_γ be the γ-core of* (14.30). *Assume $p(\Sigma q_j)$ is decreasing and each $\pi_i(q)$ is continuous and concave. Then, $C_\gamma \neq \varnothing$.*

Thus, under the same conditions of Zhao (1999a) for a non-empty β-core in (14.30), γ-core is also non-empty. This implies Rajan's γ-core result with three or four firms (1989):

Corollary 2 (Rajan 1989, p. 871) *Under the conditions of Proposition* 12*, $C_\gamma \neq \varnothing$ if $n = 3$ or* 4.

The results in Proposition 12 have been extended to coalitional interval games in which the payoffs of each coalition are given by a closed interval (Lardon 2016). Below is another interesting result in Lardon (2012):

Proposition 13 (Lardon 2012, p. 406) *Given Γ_γ in* (14.11) *for* (14.34) *with symmetric $c \in R^n_+$ and asymmetric $z \in R^n_{++}$, assume part (iii) of A*0.1. *Let q^0 be the premerger equilibrium, $\pi^m = v(N)$ the monopoly profit, and $\theta_j = v(N)q^0_j/\Sigma q^0_k$, all j. Then, $\theta \in C_\gamma = Core(\Gamma_\gamma)$.*

Thus, in linear cases, proportional split of the monopoly profit by premerger market shares is in the γ-core. This implies a symmetric γ-core result in Currarini and Marini (2015):

Corollary 3 (Currarini and Marini 2015, p. 11) *In symmetric (a, c, z) with A*0.1*, equal split of the monopoly profit is in the γ-core.*

Currarini and Marini (2006) study a class of symmetric normal form games and provide two γ-core results with applications in oligopoly models. The next two propositions are the non-technical versions of their results. Readers are referred to their paper for the technical assumptions and details.

Proposition 14 (Currarini and Marini 2006, Theorem 3.1, p. 119) *Let C_γ be the γ-core in a class of* (14.2) *in which $X_i \equiv X \subset R$, all $i \in N$. Then, $C_\gamma \neq \varnothing$ if (i) the players are symmetric and X is convex, and (ii) all $u_i(x)$ exhibit monotone externalities and increasing differences.*

This result seems to hold in classes of Bertrand oligopolies with strategic complementarity, but such claims need to be verified in future studies.

Proposition 15 (Currarini and Marini 2006, Theorem 3.2, p. 122) *Consider the same model of Proposition 14. Then, $C_\gamma \neq \varnothing$ if (i) the players are symmetric and X is convex, and (ii) all $u_i(x)$ are strictly quasi-concave, satisfy contraction property, and exhibit monotone externalities.*

This proposition holds in standard linear Cournot models and thus implies the γ-core result in Corollary 3.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Watanabe and Matsubayashi (2013) study a differentiated linear Cournot model with three or four firms and give a positive γ-core result. Chander (2014) studies the non-cooperative foundation of γ-core in (14.2) and provides a γ-core result in oligopoly (14.30). His main result, Theorem 7 in Chander (2014), appears to be identical to Proposition 12. It is not clear at present whether concavity in profit functions can be removed while still maintaining a non-empty γ-core in the oligopoly (14.30).

Zhao (2013) gives several δ-core results in a three-firm asymmetric linear Cournot oligopoly, one of which is given in the next proposition.[28] Note that complication arises once asymmetry is allowed and many of the intuitions in symmetric models no longer hold. Given $(a, c, z) = (a, c_1, c_2, c_3, z_1, z_2, z_3) \in R^7_+$, with $c_1 \leq c_2 \leq c_3$ (so firm 1 is the most efficient, and 3 the least efficient), define

$$\varepsilon_2 = (c_2 - c_1)/(a - c_1),\ \varepsilon_3 = (c_3 - c_1)/(a - c_1), \tag{14.47}$$

$$\omega_1(\varepsilon_2) = [2 - \sqrt{1 + 8\varepsilon_2 - 20\varepsilon_2^2}]/4, \tag{14.48}$$

where ε_2 and ε_3 represent the (relative) cost advantages of firm 1 over 2 and 3. The larger the value of ε_i, the less efficient (or smaller) the firm i. Obviously, $\varepsilon_2 = \varepsilon_3 = 0$ is the symmetric case, and it is easy to check that $0 \leq \varepsilon_2 \leq \varepsilon_3 \leq 0.5$ holds. These two intermediate parameters simplify the original and impossible task of characterizing the δ-core in seven dimensions to a manageable though still difficult task of characterizing the δ-core in only two dimensions.

Proposition 16 (Zhao 2013, p. 12) *Given $(a, c, z) \in R^7_+$, assume parts (ii–iii) of* A0.1*, and let $\varepsilon_2, \varepsilon_3$ and $\omega_1(\varepsilon_2)$ be given in* (14.47–14.48)*. Then, (i) $C_{\alpha*} \neq \varnothing$, and (ii) $C_\delta \neq \varnothing \iff \varepsilon_3 \geq \omega_1(\varepsilon_2)$, and $\varepsilon_3 \geq \omega_1(\varepsilon_2)$ holds if $\varepsilon_2 \in [1/6, 1/2]$.*

Thus, monopoly is always α^*-stable or can possibly be formed under the α^*-belief; it can possibly (will not) be formed under the δ-belief if firms 2 and 3 are sufficiently small, e.g., $\varepsilon_3 \geq \varepsilon_2 \geq 1/6$ (sufficiently large, e.g., $\varepsilon_3 < \omega_1(\varepsilon_2)$). In particular, it will not be formed under the δ-belief in symmetric case (i.e., $\varepsilon_3 = \varepsilon_2 = 0$). The next corollary shows that monopoly is both δ-stable and socially optimal if there are large cost savings. Here, optimality is in the sense of second best, which has the maximal welfare (= total profits + consumer surplus) among the five partitions:

Corollary 4 (Zhao 2013, p. 12) *If $\varepsilon_2 \geq 5/22$, monopoly is both δ-stable and optimal:*

Currarini and Marini (2015) give a negative δ-core result in symmetric linear Cournot oligopolies:

Proposition 17 (Currarini and Marini 2015, p. 12) *In symmetric (a, c, z) in* (14.34) *with $z = \infty$, $C_\delta = \varnothing$.*

This negative result is consistent with the symmetric case of Proposition 16.

[28] Gabszewicz, Marini and Tarola (2016) give a δ-core result in vertically differentiated markets with n firms.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Lekeas (2013) studies a differentiated symmetric linear Cournot oligopoly or a general model 9 in (14.26) and provides existence results on the *j*-core:

Proposition 18 (Lekeas 2013, pp. 9–10) *Let C_j be the j-core of* (14.3) *for a symmetric model* 9 *in* (14.26) *given by $p_k(q) = \widehat{V} - q_k - \widehat{\gamma}\Sigma_{m\neq k}q_m$ and $C_k(q_k) = cq_k$, all k. (i) Assume $\widehat{\gamma} = 1$ and $j(s) \geq 2(\sqrt{n/s} - 1)$, all s. Then, $C_j \neq \varnothing$. (ii) Assume $0 < \widehat{\gamma} < 1$ and $n \geq 2$. Then, there exists $j^* = j^*(n, \widehat{\gamma})$, $0 < j^* \leq 1$ such that for $j(s) > j^*$, all s, $C_j \neq \varnothing$. (iii) Assume $-1/(n-1) < \widehat{\gamma} < 0$. Then, $C_j \neq \varnothing$.*

Thus, by part *(iii)*, the *j*-core is always non-empty if goods are complements ($\widehat{\gamma} < 0$) and if the complementation parameter is small ($|\widehat{\gamma}| < 1/(n-1)$). If goods are substitutes ($\widehat{\gamma} > 0$), by parts *(i–ii)*, a non-empty *j*-core requires that outsiders are divided into large number of coalitions. This condition makes it hard to have a non-empty *j*-core because the belief function $j(s)$ is bounded from above by $n - s$. Note that the condition in part *(ii)* is originally stated as $j^* \leq n - s$, all s, in Theorem 1 in Lekeas (2013, p. 9), which has been simplified to $j^* \leq 1$ in the above proposition.

Lekeas and Stamatopoulos (2014) study a homogeneous Cournot model with $C_k(q_k) = cq_k$ and $Q = 1 - p^b$, $b > 0$. In this case, the game Γ_f in (14.16) is well defined. They consider a reasonable belief $f(s) = \{f_j(s) | j = 1, \ldots, n-s\}$ defined by the Sterling number of the second kind (Lekeas and Stamatopoulos 2014, p. 258), and provide an *f*-core result in linear cases ($b = 1$). However, due to the involved complexity, no *f*-core result is available in non-linear cases ($b \neq 1$).

Proposition 19 (Lekeas and Stamatopoulos 2014, p. 262) *Let C_f be the f-core of* (14.3) *for $Q = 1 - p$ and $C_k(q_k) = cq_k$, all k, with the above f-belief. Then, $C_f \neq \varnothing$ if n is sufficiently large.*

Finally, Currarini and Marini (2003) study the leader–follower belief and give an *lf*-core result (called λ-core in Currarini and Marini, 2003, 2015) in symmetric linear Cournot oligopolies. For each $S \neq N$, let $q_{-S}(q_S) = \{q_j | j \in N\backslash S\}$, $q_j \in ArgMax\{\pi_j(q_j, q_{-j}) | q_j \geq 0\}$, all $j \in N\backslash S$, be the followers' reaction function, and $q_S^* \in ArgMax\{\Sigma_{i\in S}\pi_i(q_S, q_{-S}(q_S)) | q_S \geq 0\}$ be the leaders' optimal choices. Then, the leaders' payoff and the *lf*-coalitional game are $v_{lf}(S) = \Sigma_{i\in S}\pi_i(q_S^*, q_{-S}(q_S^*))$ and $\Gamma_{lf} = \{N, v_{lf}(\cdot)\}$.

Proposition 20 (Currarini and Marini 2003) *Let C_{lf} be the core of Γ_{lf} for a symmetric* (14.34). *Then, $C_{lf} \neq \varnothing$ and equal split is its unique core vector.*

It follows from Lemma 4 and Propositions 13 and 20 that $C_{lf} \subseteq C_\gamma \subseteq C_{\alpha *} \subseteq C = C_\alpha = C_\beta$ hold in symmetric linear Cournot oligopolies.[29] It remains to be seen if new non-trivial inclusion results among the above seven core refinements (i.e., $C_{\alpha *}$, C_γ, C_e, C_δ, C_j, C_f and C_{lf}) can be found in future research.

[29] Currarini and Marini (2004) provide related existence results on the *lf*-core and the γ-core, Driessen, Hou and Lardon (2011) also provide a related *lf*-core result. See Currarini and Marini (2015, pp. 13–14) for more discussion about the *lf*-core.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

3.4 Extensions

This subsection lists seven areas for future research or extensions of the earlier core results:

1. Extend the special cases in Propositions 1–20 in the previous two subsections to more general cases of model 9 and then extend these core results in standard Cournot models to the remaining nine models or precisely to models 1–5 in (14.23) and models 6–8 and 10 in (14.26). This includes the core and its seven refinements in the multi-products or multi-markets in Grossmann (2007), Kao and Menezes (2009), Lapan and Hennessy (2006), Wang and Zhao (2010) and Zhang and Zhang (1996).
2. Extend the reviewed results to mixed oligopolies. Kamaga and Nakamura (2007) provide a core result in a three-firm mixed Cournot oligopoly with linear demand and quadratic costs, but it is not clear what their core is because none of the known α-, β-, γ- and δ-core results are cited in their paper.
3. Extend the reviewed results to oligopolies with indivisibility (*OI*, this simplifies oligopoly market with indivisibility or *OMI* in Zhao 2000), which are small markets for indivisible or discrete goods (such as superstars in sports and ocean shipping with a small number of large orders) where a one-unit change in demand or supply will have a non-negligible effect on equilibrium. Motivated by Telser's flight game (Telser 1994; see Example 4 below), Zhao (2000) models a m-buyer n-seller *OI* as

$$OI = \{A, B; C_i(x_i), [0, z_i], p_\alpha\}, \tag{14.49}$$

where $A = \{1, \ldots, n\}$ is the set of firms or sellers, $B = \{1, \ldots, m\}$ is the set of buyers; $C_i(x_i)$, $x_i \in [0, z_i]$, with $z_i > 0$ as capacity, is the cost function of each firm $i \in A$, and $p_\alpha \geq 0$ is the reservation price of each buyer $\alpha \in B$ for one unit of the homogeneous good. This differs from (14.30) only in that both x_i and z_i are integers and the inverse demand is replaced by a vector of reservation prices $p \in R^m_+$. A coalitional game $\Gamma_c = \{N, v(\cdot)\}$, $N = A \cup B$, can be defined by computing the profit $v(S)$ for each $S = \{T_A, T_B\} \subseteq \{A, B\}$, with $T_A \subseteq A$ and $T_B \subseteq B$.[30]

Let a linear cost be given by $C_i(x_i) = b_i$ if $x_i = 0$; $= d_i + c_i x_i$ if $x_i = 1, \cdots, z_i$; and $= \infty$ if $x_i > z_i$, where d_i and c_i are the fixed and marginal costs, $b_i \gtreqless 0$ is the opportunity cost if $b_i < 0$, the sunk cost if $b_i = d_i$, and it makes d_i an avoidable cost if $b_i = 0$. Such *OI* with linear costs can be defined by a $(4n + m)$-vector $\{b, c, d, z, p\} \in R^n \times R^{2n}_+ \times R^n_{++} \times R^m_+$, with p, z, d, c and b as the vectors of reservation prices, capacities, fixed, marginal and opportunity costs, respectively. In *OI* with only opportunity cost (i.e., $c = d = 0$), (14.49) is reduced to a $(2n + m)$-vector $\{b, z, p\} \in R^n_+ \times R^n_{++} \times R^m_+$. Telser's flight game is a two-seller three-buyer OI with only opportunity cost given here.

Example 4 (Flight game, Telser 1994) $\{b, z, p\} = \{(85, 150), (2, 3), (70, 60, 55)\}$. *There are three passengers whose reservation prices for a trip are \$70, \$60 and \$55, respectively, and there are two private jets (or cabs), one with an opportunity cost* $b_1 = \$85$ *and capacity* $z_1 = 2$, *and the other with* $b_2 = 150$ *and* $z_2 = 3$.

[30] See Zhao (2000, pp. 184–186) for details. Note that A and B have been switched from those in Zhao (2000) to emphasize that B is the set of buyers.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Proposition 21 (Zhao 2000, p. 191) *Given* $\{b, z, p\} = \{(b_1, b_2), (z_1, z_2), (p_1, p_2, p_3)\}$, *with* $b_1 < b_2$ *and* $p_1 \geq p_2 \geq p_3$. *Under usual conditions, the core is empty* $\Longleftrightarrow 3b_1/2 < b_2 < 3p_3$.

By $3b_1/2 = 127.5 < b_2 = 150 < 3p_3 = 165$, the core is empty in Example 4. Bejan and Gómez (2009) provide a non-core solution or core extension for such empty core games.

4. Extend the reviewed results to uncertainty and asymmetric information along the lines line of Allen (2006), Currarini and Feri (Chapter 18 in this *Handbook*), Wilson (1978) and Yannelis (2005).
5. Extend the reviewed results to dynamic games as surveyed in Bischi, Lamantia and Radi (Chapter 12 in this *Handbook*) and Long (2010).
6. Study the external stability (called comparative statics in economics) of the core such as its upper semi-continuity (*u.s.c*) and lower semi-continuity (*l.s.c*). Applying the *l.s.c* condition for an optimal set (Zhao 1997) to *mnbp* in (14.5) should lead to non-trivial results, which could shed light on studying merger contracts under uncertainty. This should not be confused with internal stability caused by coalitional deviations.
7. Connect the above core results to the huge literature on the non-cooperative approach to coalition formation (Bloch 1997; Currarini and Marini 2006, 2015; Ray and Vohra 1997, 2015). Some of the stable monopolies formed in such studies appear to be a refinement of the core ($= \alpha$-core $= \beta$-core), but this is not totally clear and needs to be verified in future studies, because the known α- or β-core results and connections to the partition function game (14.3) via the *MOG* (14.46) or quasi-hybrid solution have not been cited in such studies. If it is verified as a core refinement, these results form a new refinement of the core.

4 STABLE PARTITIONS AS CANDIDATES OF NON-MONOPOLY SOLUTIONS

There are no published and only two unpublished studies on the stability of a general non-monopoly partition. The main ideas in the author's 20-year-old working paper (Zhao 1996) are surprisingly still new. Only the basic idea in this old paper and one result in (Zhao 2013) are reviewed here.

Given a non-monopoly partition $\Delta = \{S_1, S_2, \ldots, S_h\} \neq \{N\}$ in (14.30), any notion of its stability must have two basic elements: an unprofitable monopoly merger caused by merging costs, and a payoff vector $\theta = \theta(\Delta) = \{\theta_S |\ S \in \Delta\} \in R^n_+$ satisfying $\Sigma_{j \in S} \theta_j = \phi_S(\Delta)$ for each $S \in \Delta$, where $\phi_S(\Delta)$ is given in (14.3) for (14.30). Thus, this section assumes that for each $\Delta \neq \{N\}$, $\Sigma_{S \in \Delta} \phi_S(\Delta) > v(N) = (\pi^m - mmc)$ with some positive monopoly merging cost $mmc > 0$. The concept of *hybrid solution with a distribution rule* (*HSDR* in Zhao 1996 or Zhao 1999a) was introduced to define the other basic element.

Let $D =$ {core, equal surplus split, nucleolus, proportional split, Shapley value} be the set of five solutions, which are restricted to (14.30) and exclude other solutions whose general existences are either unknown or too involved. A distribution rule (*DR*) for the given Δ specifies a solution $DR(S) \in D$ for each $S \in \Delta$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Definition 5 (Zhao 1996, 1999a) *Given a partition* $\Delta \neq \{N\}$ *and its DR in an oligopoly* (14.30), *its hybrid solution with a distribution rule or HSDR is a pair* (q^*, θ) *such that* q^* *is the solution of* (14.20) *and for each* $S \in \Delta$, θ_S *is its solution defined by DR*(S).

Keep in mind that θ_S *or DR*(S) solves the parametric normal form game

$$\Gamma_S = \Gamma_S(q^*_{-S}) = \{S, X_i, \pi_i(q_S, q^*_{-S})\} \tag{14.50}$$

for each $S \in \Delta$, where $X_i = [0, z_i]$, $\pi_i(q_S, q^*_{-S}) = p(\Sigma_{j \in S} q_j + \Sigma_{j \notin S} q^*_j)q_i - C_i(q_i), i \in S$, and $q^* = \{q^*_S | \ S \in \Delta\}$ is the quasi-hybrid solution or postmerger equilibrium in (14.20). Such *HSDR* can be called the merger contracts for Δ, which extends the monopoly merger contract $(N, \overline{q}, \theta)$ to a partition contract (Δ, q^*, θ), specifying that each merger $S \in \Delta$ distribute its profits by a solution $DR(S) \in D$.

Now, given a partition contract (Δ, q^*, θ) for $\Delta \neq \{N\}$, consider the possible deviation by a coalition $S \notin \Delta$, $S \neq N$. S has incentives to move to a new partition $\Delta' = \{S, T_1, \ldots, T_m\} \in \Pi(S)$ in (14.10) if its payoff at the new partition is higher than the sum of its members' current payoffs or if $\phi_S(\Delta') > \Sigma_{j \in S}\theta_j$, with $\phi_S(\Delta')$ given in (14.3) for (14.30). A stable contract (Δ, q^*, θ) should rule out all such deviations.

Note that most stable partitions identified in the non-cooperative approach require symmetry, and some of them (such as the equilibrium-binding agreement in Ray and Vohra 1997) only rule out a small set of possible coalitional deviations; such stable partitions are thus very weak and are not really stable. This is the reason why they have been excluded here as candidates of non-monopoly solutions.

For simplicity, only the γ-, δ-, α^*- and e-deviations are evaluated here. The new partitions $\Delta_{\alpha*} = \Delta_{\alpha*}(S) = \{S, T_1^{\alpha*}, \ldots, T^{\alpha*}_{m(\alpha*)}\}$ and $\Delta_e = \Delta_e(S) = \{S, T^e_1, \ldots, T^e_{m(e)}\}$ are the same as in (14.13–14.14), due to their independency of the current Δ. The new partitions $\Delta_\gamma = \Delta_\gamma(S, \Delta)$ and $\Delta_\delta = \Delta_\delta(S, \Delta)$ under γ- and δ-beliefs are given by

$$\Delta_\gamma = \Delta_\gamma(S, \Delta) = \{S, T^\gamma_1, \ldots, T^\gamma_{m(\gamma)}\} \in \Pi(S), \text{ and} \tag{14.51}$$

$$\Delta_\delta = \Delta_\delta(S, \Delta) = \{S, T^\delta_1, \ldots, T^\delta_{m(\delta)}\} \in \Pi(S), \tag{14.52}$$

where for each $i = 1, \ldots, m(\gamma)$, $T^\gamma_i = T$ for each $T \in \Delta$ with $S \cap T = \varnothing$, $= \{j\}$ for each $j \in T \backslash S$ and each $T \in \Delta$ with $S \cap T \neq \varnothing$; and for each $i = 1, \ldots, m(\delta)$, $T^\delta_i = T \backslash S = \{j | j \in T, j \notin S\}$ for each $T \in \Delta$. As an example, for $\Delta = \{1, 2345, 67\}$ and $S = \{1, 2\}$, one has $\Delta_\gamma = \{12, 3, 4, 5, 67\}$ and $\Delta_\delta = \{12, 345, 67\}$.

Definition 6 (Zhao 1996) *A partition contract* (Δ, q^*, θ) *for* $\Delta \neq \{N\}$ *or* Δ *with* $\theta(\Delta)$ *is* γ*-stable (*δ*-,* α^**- and e-stable) if for all* $S \notin \Delta$, $\Sigma_{j \in S}\theta_j \geq \phi_S(\Delta_\gamma)$ $(\geq \phi_S(\Delta_\delta), \geq \phi_S(\Delta_{\alpha*})$ *and* $\geq \phi_S(\Delta_e))$, *where* $\Delta_{\alpha*}$ *and* Δ_e *are given in* (14.13–14.14), Δ_γ *and* Δ_δ *are given in* (14.51–14.52).

Let the *mnbp* under the above four notions of stability be given by

$$mnbp_k = \{Min \Sigma^n_{i=1} x_i | x \geq 0, \Sigma_{i \in S} x_i \geq \phi_S(\Delta_k), all\ S \notin \Delta, S \neq N\} \tag{14.53}$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

for $k = \gamma$, δ, α^**and* e, and let the optimal set or the set of minimal solutions in (14.53) be given, respectively, by

$$Y_\gamma(\Delta),\ Y_\delta(\Delta),\ Y_{\alpha*}(\Delta)\ \text{and}\ Y_e(\Delta). \tag{14.54}$$

Then, the stability of $\Delta \neq \{N\}$ with $\theta(\Delta)$ is fully characterized by (14.54) or (14.53).

Proposition 22 (Zhao 1996) *Given a partition contract* (Δ, q^*, θ) *for* $\Delta \neq \{N\}$ *in* (14.30), *assume* $v(N) = (\pi^m - mmc) < \Sigma_{i=1}^n \theta_i$. *Then, for* $k = \gamma$, δ, α^**and e,* Δ *with* θ *is k-stable* $\Leftrightarrow$ $\theta \in Y_k(\Delta)_+$, *where* $Y_k(\Delta)$ *is given in* (14.54) *and* $Y_k(\Delta)_+ = \{x + y | x \in Y_k(\Delta), y \in R_+^n\}$.

If players are allowed to freely redistribute θ in (Δ, q^*, θ) among the n players, the above conclusions can be simplified as $\Delta \neq \{N\}$ is k-stable if and only if for $k = \gamma$, δ, α^**and e*, $\Sigma\theta_i = \Sigma_{S\in\Delta}\phi_S(\Delta) \geq mnbp_k$, which is given in (14.53).

The next proposition concludes this section with a δ-stable non-monopoly partition in a three-firm asymmetric linear Cournot oligopoly by applying Proposition 22.

Given $(a, c, z) \in R_+^7$ and $\Delta_1 = \{1, 23\}$, $\Delta_2 = \{2, 13\}$, $\Delta_3 = \{3, 12\}$, one can verify that their γ-, δ- and α^*-stabilities are identical, so there is no need to make such distinction here. The outsider's or the single firm's profit at each Δ_k $(k = 1, 2, 3)$ is

$$\begin{aligned}
\phi_1(\Delta_1) &= (a - c_1)^2(1 + \varepsilon_2)^2/9,\\
\phi_2(\Delta_2) &= (a - c_1)^2(1 - 2\varepsilon_2)^2/9,\\
\phi_3(\Delta_3) &= (a - c_1)^2(1 - 2\varepsilon_3)^2/9,
\end{aligned} \tag{14.55}$$

where ε_2 and ε_3 are given in (14.47). The merger's gain in each Δ_k is:

$$\begin{aligned}
d_{23} &= \phi_{23}(\Delta_1) - (\pi_2^0 + \pi_3^0),\\
d_{13} &= \phi_{13}(\Delta_2) - (\pi_1^0 + \pi_3^0),\\
d_{12} &= \phi_{12}(\Delta_3) - (\pi_1^0 + \pi_2^0),
\end{aligned} \tag{14.56}$$

where π_i^0 is i's premerger profit, and $\phi_S(\Delta)$ is the postmerger profit in (14.3) for (14.34). For $S = 12$, 13, and 23, let the efficient member's share of the above gains be $t \in [0, 1]$. Then, the three-dimensional payoff vector $\theta(\Delta_k) = \theta(t) \in R_+^3$ for $k = 1, 2, 3$, respectively, is

$$\begin{aligned}
&\text{for } \Delta_1, \theta_1 = \phi_1(\Delta_1), \theta_2 = \pi_2^0 + td_{23}, \theta_3 = \pi_3^0 + (1 - t)d_{23},\\
&\text{for } \Delta_2, \theta_1 = \pi_1^0 + td_{13}, \theta_2 = \phi_2(\Delta_2), \theta_3 = \pi_3^0 + (1 - t)d_{13},\ \text{and}\\
&\text{for } \Delta_3, \theta_1 = \pi_1^0 + td_{12}, \theta_2 = \pi_2^0 + (1 - t)d_{12}, \theta_3 = \phi_3(\Delta_3).
\end{aligned} \tag{14.57}$$

Proposition 23 (Zhao 2013, p. 16) *Given* $(a, c, z) \in R_+^7$, *suppose* $\Sigma_{i=1}^3\theta_i > (\pi^m - mmc)$, $d_S > 0$ *for* $S = 12$, 13 *and* 23, *and assume parts (ii–iii) of* A0.1. *Then, the following three claims hold:*

(i) Δ_1 *with* $\theta(t)$ *is stable (i.e.,* δ*- or* γ*- or* α^**- stable)* $\Longleftrightarrow$ $\varepsilon_3 \leq \mu_1(\varepsilon_2, t)$*;* $\varepsilon_3 \leq \mu_1(\varepsilon_2, t)$ *if* $0 \leq \varepsilon_2 \leq 1/11$, *and* $\varepsilon_3 > \mu_1(\varepsilon_2, t)$ *if* $113/316 < \varepsilon_2 \leq 1/2$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

(ii) Δ_2 *with* $\theta(t)$ *is stable* $\Longleftrightarrow \varepsilon_3 \leq \mu_2(\varepsilon_2, t)$; $\varepsilon_3 \leq \mu_2(\varepsilon_2, t)$ *if* $0 \leq \varepsilon_2 \leq 1/11$, *and* $\varepsilon_3 > \mu_2(\varepsilon_2, t)$ *if* $e_2(t) < \varepsilon_2 \leq 1/2$, *where* $e_2(t) = (2t - 9)/[14(2t - 3)]$.

(iii) Δ_3 *with* $\theta(t)$ *is stable* $\Longleftrightarrow \varepsilon_2 \leq \mu_3(\varepsilon_3, t)$, *which holds if* $0 \leq \varepsilon_3 \leq 3/14$.

The details of $\mu_1(\varepsilon_2, t)$ and $\mu_2(\varepsilon_2, t)$ (note $\mu_3(\varepsilon_3, t) = \mu_2(\varepsilon_3, t)$) can be found in (A27–A29) in Zhao (2013). Although such results appear to be technical, they have intuitive interpretations. Consider, for example, part *(i)* or the stability of $\Delta_1 = \{1, 23\}$. Observe first that $\Delta_0 = \{1, 2, 3\}$ and $\Delta_m = \{123\}$ are ruled out by the assumptions and $\Delta_2 = \{2, 13\}$ has the same postmerger profits of Δ_1; thus, one only needs to evaluate the deviation by $S = 12$ in Δ_3. Because a larger share t by firm 2 or a smaller ε_3 makes the merger of $S = 12$ less profitable, Δ_1 *with* $\theta(t)$ will be stable with a smaller ε_3 or a larger t (i.e., $\varepsilon_3 \leq \mu_2(\varepsilon_2, t)$, which is increasing in t).

5 EMPIRICAL STUDIES OF THE CORE

Early empirical studies of the core include Bittlingmayer (1982), Sjostrom (1989), Pirrong (1992), and McWilliams and Keith (1994). Reading these papers, it is clear that the authors had an intuitive understanding of the core for which there was no precise model and that their intuition was based on the empty-core examples in their and other early studies such as Faulhaber (1975), Shapley and Shubik (1969) and Telser (1978, 1994). They understood that the core theory assumes $A3$ and that the empty-core was caused by economies of scale in Addyston Pipe (Bittlingmayer 1982), or by low demand plus indivisible supply or avoidable cost in Ocean Shipping Conferences (Pirrong 1992, Sjostrom 1989) and Trust Industries (McWilliams and Keith 1994); they equated *empty-core* to *market failure* or *ruinous competition*.

The core in these studies involves a small number of sellers and buyers (similar to the oligopoly with indivisibility in (14.49)), so it is different from the core in oligopolies as defined earlier. The documented evidences on the sellers' arrangements were interpreted as a solution for the empty-core to avoid cut-throat competition. However, these same evidences on sellers' arrangements might be interpreted alternatively as supports for a non-empty core in games involving only the sellers. The author believes that a non-empty core of the sellers can be established by revisiting the evidences of Joint Traffic in Addyston Pipe (Bittlingmayer 1982), prices and quotas in Ocean Shipping (Pirrong 1992, Sjostrom 1989) and share allocations in Trusts (McWilliams and Keith 1994).

Recent applications of the core in Propositions 2–5 include liner shipping alliances (Shi and Voss 2011, Yang, Liu and Shi 2011), insurance (Stoyanova and Gruendl 2014) and sugar monopoly (Zhao 2009b). Stoyanova and Gruendl (2014) study the EU legislation called Solvency II, which replaced 13 old EU insurance directives on January 1, 2016. Their conclusion is that Directive Solvency II will reduce merging costs and drive more mergers and acquisitions in the EU insurance industry.

Zhao (2009b) applies Proposition 5 to the 1887–1914 sugar monopoly (Eichner 1969, Genesove and Mullin 1998, and Wang 2008), which replaced the 1882–87 Sugar Trust (McWilliams and Keith 1994). The monopoly consolidated 18 refineries in 1887, with an excess capacity rate of about 20 percent; it dissolved into 12 refineries in 1914, with a near full capacity. Using $n = 18$, $\tau = 0.20$, and the estimated linear model in Genesove and

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Mullin (1998), the estimated monopoly merging costs are at most 35 percent of pre-merger total profits at its formation in 1887. Using $n = 12$ and $\tau = 0$, the estimated monopoly's organizational costs (i.e., the costs of keeping monopoly and avoiding dissolution) are at least 252 percent of post-dissolution total profit at its dissolution in 1914. These results provide a new understanding of the formation and dissolution of the sugar monopoly: it was formed in 1887 because its merging cost was sufficiently low, and it was dissolved in 1914 because enforcing the Sherman Act increased its organizational costs to a level that was too high to be operational. Zhao (2009b) also reports similar results using linear demand and quadratic costs.

6 CONCLUSION AND FUTURE STUDY

The process from early division of labor or specialization to modern industrial organization is long and dates back at least 170,000 years.[31] Such a long process in human history is powered and pulled forward by its two indisputable driving horses or driving wheels called *competition* and *cooperation*.

The previous literature in industrial organization has largely focused on *competition* or the application of non-cooperative game theory, with the exception of a small group of scholars whose works on *industrial cooperation* have been reviewed with some details in this survey. Readers are encouraged to extend the surveyed results to more general and more realistic models. In addition to the seven extensions listed in subsection 3.4, applied scholars are encouraged to extend the few empirical core studies to all industries or sectors with merger activities or joint ventures, and theoretically minded scholars are encouraged to extend the existing core results to non-monopoly partitions.

REFERENCES

Allen, B. (2000), The future of microeconomic theory, *Journal of Economic Perspectives* 14, 143–150.

Allen, B. (2006), Market games with asymmetric information: the core, *Economic Theory* 29, 465–487.

Aumann, R. (1959), Acceptable points in general cooperative n-person games, in A.W. Tucker and R.D. Luce (eds), *Contributions to the Theory of Games IV*, *Annals of Mathematics Studies* 40. Princeton, NJ: Princeton University Press.

Bejan, B., and J. Gómez (2009), Core extensions for non-balanced TU-games, *International Journal of Game Theory* 38, 3–16.

Bertrand, J. (1883), Review of "Théorie mathématique de la richesse sociale" and "Recherches sur les principes mathématiques de la théorie des richesses", *Journal des Savants* (Paris), 499–508.

Bittlingmayer, G. (1982), Decreasing average cost and competition: a new look at the Addyston Pipe Case, *Journal of Law & Economics* 25, 201–229.

Bloch, F. (1997), Non cooperative models of coalition formation in games with spillovers, in C. Carraro and D. Siniscalco (eds), *New Directions in the Economic Theory of the Environment*. Cambridge, UK: Cambridge University Press.

Bloch, F., and A. van den Nouweland (2014), Expectation formation rules and the core of partition function games, *Games and Economic Behavior* 88, 339–353.

Bondareva, O. (1962), The theory of the core in an N-person game [in Russian], Vestnik Leningrad Univ. 13, 141–142.

[31] One of the earliest evidences of specialization is a workshop or small factory for making stone tools, recently confirmed to be at least 170,000 years old, at the Dingcun Site (Ding village in Xiangfen) in Shanxi, China. See Wang (2015). The entries on Dingcun on both Wiki and the UNESCO World Heritage Tentative List are about ancient civilian residential houses in the same village.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Chander, P. (2014), The core of a strategic game. *Working Paper*, Jindal School of Government and Public Policy, Sonipat, India: Jindal Global University.
Chander, P., and H. Tulkens (1997), A core of an economy with multilateral environmental externalities, *International Journal of Game Theory* 26, 379–401.
Cournot, A. (1838), *Recherches sur les principes mathématiques de la théorie des richesses*. Paris: L. Hachette.
Currarini, S., and M. Marini (2003), A sequential approach to the characteristic function and the core in games with externalities. In M. Sertel and A. Kara (eds), *Advances in Economic Design*. Berlin: Springer Verlag.
Currarini, S., and M. Marini (2004), A conjectural cooperative equilibrium in strategic form games, in C. Fragnelli V. (ed.), *Game Practice and the Environment*. Cheltenham, UK and Northampton, MA: Edward Elgar Publishing.
Currarini, S., and M. Marini (2006), Coalition formation in games without synergies, *International Game Theory Review* 8, 111–126.
Currarini, S., and M. Marini (2015), Coalitional approaches to collusive agreements in oligopoly games, *The Manchester School* 83, 253–287.
Diamantoudia, E., and L. Xue (2007), Coalitions, agreements and efficiency, *Journal of Economic Theory* 136, 105–125.
Dixit, A. (1979), A model of duopoly suggesting a theory of entry barriers, *Bell Journal of Economics* 10, 20–32.
Driessen, T., and H. Meinhardt (2005), Convexity of oligopoly games without transferable technologies, *Mathematical Social Sciences* 50, 102–126.
Driessen, T., and H. Meinhardt (2010), On the supermodularity of homogeneous oligopoly games, *International Game Theory Review* 12, 309–337.
Driessen, T., D. Hou and A. Lardon (2011), Stackelberg oligopoly TU-games: characterization of the core and 1-concavity of the dual game. *Working Paper*, Saint-Etienne, France: University of Saint-Etienne.
Edgeworth, F. (1881), *Mathematical Psychics*. London: Kegan Paul.
Eichner, A.S. (1969), *The Emergence of Oligopoly: Sugar Refining as a Case Study*. Baltimore, MD: Johns Hopkins Press.
Farrell, J., and C. Shapiro (1990), Horizontal mergers: an equilibrium analysis, *American Economic Review* 80, 107–126.
Faulhaber, G. (1975), Cross-subsidization: pricing in public enterprises, *American Economic Review* 65, 966–977.
Folmer, H., and P. Mouche (1994), Interconnected games and international environmental problems II, *Annals of Operations Research* 54, 97–117.
Funaki, Y., and T. Yamato (1999), The core of an economy with a common pool resource: a partition function form approach, *International Journal of Game Theory* 28, 157–171.
Gabszewicz, J., M. Marini and O. Tarola (2016), Core existence in vertically differentiated markets, *Economics Letters*, 149, 28–32.
Genesove, D., and W. Mullin (1998), Testing static oligopoly models: conduct and cost in the sugar industry, 1890–1914, *RAND Journal of Economics* 29, 355–377.
Gillies, D. (1959), Solutions to general non-zero-sum games, in A. Tucker and R. Luce (eds), *Contributions to the Theory of Games IV*. Princeton, NJ: Princeton University Press.
Grossmann, V. (2007), Firm size and diversification: multiproduct firms in asymmetric oligopoly, *International Journal of Industrial Organization* 25, 51–67.
Hart, S., and M. Kurz (1983), Endogenous formation of coalitions, *Econometrica* 51, 1047–1064.
Hou, D., T. Driessen and A. Lardon (2011), Convexity and the Shapley value in Bertrand oligopoly TU-games with Shubik's demand functions. *Working Paper*, Saint-Etienne, France: University of Saint-Etienne.
Huang, C., and T. Sjostrom (2003), Consistent solutions for cooperative games with externalities, *Games and Economic Behavior* 43, 196–213.
Kajii, A. (1992), A generalization of Scarf theorem: an α-core existence theorem without transitivity or completeness, *Journal of Economic Theory* 56, 194–205.
Kamaga, K., and Y. Nakamura (2007), The core and productivity-improving mergers in mixed oligopoly, *International Journal of Business and Economics* 6, 181–198.
Kao, T., and F. Menezes (2009), Endogenous mergers under multi-market competition, *Journal of Mathematical Economics* 45, 817–829.
Lapan, H., and D. Hennessy (2006), A note on cost arrangement and market performance in a multi-product Cournot oligopoly, *International Journal of Industrial Organization* 24, 583–591.
Lardon, A. (2010), Convexity of Bertrand oligopoly TU-games with differentiated products. *Working Paper*, Saint-Etienne, France: University of Saint-Etienne.
Lardon, A. (2012), The γ-core in Cournot oligopoly TU-games with capacity constraints, *Theory and Decision* 72, 387–411.
Lardon, A. (2016), Endogenous interval games in oligopolies and the cores, *Annals of Operations Research* (published online: April 29, 2016).
Lekeas, P. (2013), Coalitional beliefs in Cournot oligopoly TU-games, *International Game Theory Review* 15, 1–21.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Lekeas, P., and G. Stamatopoulos (2014), Cooperative oligopoly games with boundedly rational firms, *Annals of Operations Research* 223, 255–272.
Long, N. (2010), *Dynamic Games in Economics: A Survey*. Singapore: World Scientific.
Marini, M. (2009), Games of coalition and network formation: a survey, in A. Naizmada et al. (eds), *Networks, Topology and Dynamics*, *Lectures Notes in Economics & Mathematical Systems* 613, 67–93. Berlin: Springer-Verlag.
McCain, R. (2008), Cooperative games and cooperative organizations, *Journal of Socio-Economics* 37, 2155–2167.
McWilliams, A., and K. Keith (1994), The genesis of the trusts: rationalization in empty core markets, *International Journal of Industrial Organization* 12, 245–267.
Meinhardt, H. (2002), Cooperative decision making in common pool situations. *Lecture Notes in Economics and Mathematical Systems* 517. Berlin: Springer.
Meinhardt, H. (2013), *The Pre-Kernel as a Tractable Solution for Cooperative Games: An Exercise in Algorithmic Game Theory.* Berlin: Springer.
Nash, J. (1950), Equilibrium points in n-person games, *Proceedings of the National Academy of Sciences of the United States* 36, 48–49.
Norde, H., D. Pham and S. Tijs (2002), Oligopoly games with and without transferable technologies, *Mathematical Social Sciences* 43, 187–207.
Pham Do, K., and H. Folmer (2003), International fishery agreements: the feasibility and impacts of partial cooperation. *CentER Discussion Paper* No. 2003–52, Tilburg University.
Pirrong, S. (1992), An application of the core theory to the analysis of ocean shipping markets, *Journal of Law and Economics* 35, 89–131.
Rajan, R. (1989), Endogenous coalition formation in cooperative oligopolies, *International Economic Review* 30, 863–876.
Ray, D., and R. Vohra (1997), Equilibrium binding agreements, *Journal of Economic Theory* 73, 30–78.
Ray, D., and R. Vohra (2015), Coalition formation, in P. Young and S. Zamir (eds), *Handbook of Game Theory with Economic Applications* 4, 239–326. Amsterdam, North-Holland.
Scarf, H. (1971), On the existence of a cooperative solution for a general class of n-person games, *Journal of Economic Theory* 3, 169–181.
Schmeidler, D. (1969), The nucleolus of a characteristic function game, *SIAM Journal of Applied Mathematics* 17, 1163–1170.
Shapley, L. (1953), A value for N-person games, in H.W. Kuhn and A.W. Tucker (eds), *Contributions to the Theory of Games II*, *Annals of Mathematics Studies* 28. Princeton, NJ: Princeton University Press.
Shapley, L. (1955), Markets as cooperative games, *Rand Corporation Papers* P–629, 1–5.
Shapley, L. (1956), Equilibrium points in games with vector payoffs, *Rand Corporation Research Memorandum* RM-1818, 1–7.
Shapley, L. (1959), Equilibrium points in games with vector payoff, *Naval Research Logistics Quarterly* 6, 57–61.
Shapley, L. (1967), On balanced sets and cores, *Naval Research Logistics Quarterly* 14, 453–460.
Shapley, L., and M. Shubik (1969), On the core of an economic system with externalities, *American Economic Review* 59, 678–684.
Shi, X., and S. Voss (2011), Game theoretical aspects in modeling and analyzing the shipping industry, in J.W. Bose et al. (eds), *Computational Logistics*, *Lecture Notes in Computer Science* 6971, 302–320.
Shubik, M. (1980), *Market Structure and Behavior*. Cambridge, MA: Harvard University Press.
Sjostrom, W. (1989), Collusion in ocean shipping: a test of monopoly and empty core models, *Journal of Political Economy* 97, 1160–1179.
Stoyanova, R., and H. Gruendl (2014), Solvency II: a driver for mergers and acquisitions? *Geneva Papers on Risk and Insurance* 39, 417–439.
Telser, G. (1978), *Economic Theory and the Core*. Chicago, IL: University of Chicago Press.
Telser, L. (1994), The usefulness of core theory in economics, *Journal of Economic Perspectives* 8, 151–164.
Thrall, R., and W. Lucas (1963), N-person games in partition function form, *Naval Research Logistics Quarterly* 10, 281–298.
Uyanık, M. (2015), On the nonemptiness of the α-core of discontinuous games: transferable and nontransferable utilities, *Journal of Economic Theory* 158, 213–231.
von Neumann, J., and O. Morgenstern (1944), *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.
Wang, X. (2008), A study of the sugar refinery industry in the United States: 1887–1914, *MA Project Paper*, Department of Economics, Saskatoon, SK: University of Saskatchewan.
Wang, X., and J. Zhao (2010), Why are firms sometimes unwilling to reduce costs? *Journal of Economics* 101, 103–124.
Wang, Y. (2015), Dingcun relics confirmed by the academic conference, *Chinese Cultural Relics* (November).
Watanabe, T., and N. Matsubayashi (2013), Note on stable mergers in markets with asymmetric substitutability, *Economics Bulletin* 33, 2024–2033.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Wilson, R. (1978), Information, efficiency, and core of economy, *Econometrica* 46, 807–816.
Yang, D., M. Liu and X. Shi (2011), Verifying liner shipping alliance's stability by applying core theory, *Research in Transportation Economics* 32, 15–24.
Yannelis, N. (2005), The core of an economy with differential information, in D. Glycopantis and N. Yannelis (eds), *Differential Information Economies*. *Studies in Economic Theory* 19, 73–86. Berlin: Springer.
Yong, J. (2004), Horizontal monopolization via alliances. *Working Paper*, Melbourne Institute of Applied Economic and Social Research, Parkville, VIC: University of Melbourne.
Zhang, A., and Y. Zhang (1996), Stability of a Cournot-Nash equilibrium: the multiproduct case, *Journal of Mathematical Economics* 26, 441–462.
Zhang, J. (2016), Personal communication, July 2016.
Zhao, J. (1991), The equilibria of multiple objective games, *International Journal of Game Theory* 20, 171–182.
Zhao, J.(1992), The hybrid solutions of an N-person game, *Games and Economic Behavior* 4, 145–160.
Zhao, J. (1996), The stability and the formation of coalition structures in norm form TU games. *Working Paper* 96–13, Department of Economics, Columbus, OH: Ohio State University.
Zhao, J. (1997), The lower semicontinuity of optimal solution set, *Journal of Mathematical Analysis and Application* 207, 240–254.
Zhao, J. (1999a), A β-core existence result and its application to oligopoly markets, *Games and Economic Behavior* 27, 153–168.
Zhao, J. (1999b), A necessary and sufficient condition for the convexity in oligopolistic games, *Mathematical Social Sciences* 37, 189–204.
Zhao, J. (1999c), The existence of TU alpha-core in norm form games, *International Journal of Game Theory* 28, 25–34.
Zhao, J. (2000), The core in an oligopoly market with indivisibility, *Economic Theory* 16, 181–198.
Zhao, J. (2001a), A characterization for the negative effects of cost reduction in Cournot oligopoly, *International Journal of Industrial Organization* 19, 455–469.
Zhao, J. (2001b), The relative interior of base polyhedron and the core, *Economic Theory* 18, 635–648.
Zhao, J. (2009a), Estimating merging costs by merger preconditions, *Theory and Decision* 66, 373–399.
Zhao, J. (2009b, first draft), Estimating the organizational and merging costs of sugar monopoly, 1887–1914. *Working Paper*, Department of Economics, Saskatoon, SK: University of Saskatchewan.
Zhao, J. (2012), An open problem on inverse matrices from industrial organization, and a partial solution, *Linear Algebra and its Applications* 437, 294–306.
Zhao, J. (2013), The most reasonable solution for an asymmetric three-firm oligopoly. *Working Paper*, Department of Economics. Saskatoon, SK: University of Saskatchewan.
Zhao, J. (2018), Three little-known and yet still significant contributions of Lloyd Shapley, *Games and Economic Behavior* (forthcoming, https://doi.org/10.1016/j.geb.2017.05.002).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

PART IV

INFORMATION GAMES

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

15. Trading under asymmetric information: Positive and normative implications

Andrea Attar and Claude d'Aspremont

1 INTRODUCTION

Trading under asymmetric information involves traders (buyers and sellers) some (or all) of whom have private information about characteristics that influence their utilities and their beliefs and that are relevant if trading is to be mutually beneficial. However, the terms of trade being fixed through some market institution (with or without public intervention), some traders might have an interest in hiding or distorting their private information. Such behaviors might affect the equilibrium allocation, its existence and its efficiency.

The incentive problems above were traditionally linked to collective issues in public expenditure and taxation theory (Samuelson, 1969, Mirrlees, 1971), but, as made clear in the pioneering contributions of Vickrey (1961) and Hurwicz (1973), these issues arise whenever economic decisions are reached through some decentralized process, even when goods are purely private.

The early research on auctions by Vickrey (1961) and others (*e.g.* Griesmer, Levitan and Shubik, 1967, and Wilson, 1967) was particularly instructive in this regard. Auction theory uses (without saying) the Bayesian equilibrium concept (formalized by Harsanyi, 1967, 1968) where strategies are functions of the player's type, and auction design has become a major application of mechanism design (Myerson, 1981). First-best efficiency is replaced by constrained (or second-best) efficiency and, in the case of one seller, an optimal auction is one for which a Bayesian equilibrium exists and maximizes the seller's expected utility. The non-informed seller acts as a principal (and mechanism designer) and the buyers as agents.

Akerlof (1970) was among the first to point out the potentially dramatic implications of incomplete information in competitive markets. His example features a number of non-informed buyers who compete to attract sellers who are informed about the quality of the product (used cars). Sellers with high-quality cars tend to withdraw from the market and at the (competitive) equilibrium price only the "lemons" (or even no car) get traded. Yet, little attention is given to the role of incentive mechanisms to elicit information revelation and (potentially) unfreeze the market. Indeed, in the original Akerlof (1970) example, buyers are restricted to posting linear prices, in the spirit of competitive equilibrium theory. The need for a more explicit representation of agents' strategic behaviors gave rise to two independent lines of research. The first one, acknowledged as "signaling," develops the analysis of extensive form games in which the informed agents move first. This line of research was initiated by Spence (1973) who introduced the possibility for informed sellers to "signal" the quality of their product (labor) by taking a costly action (education).[1] The second one is usually

[1] The possibility of signaling has triggered a lot of game-theoretical research to deal with the multiplicity of equilibria due to the possibility of unanticipated action by the first movers: e.g. Kreps and Wilson (1982), Cho and Kreps (1987), Kohlberg and Mertens (1986).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

referred to as competitive "screening." This amounts to considering games in which several players have the power to design contracts to attract privately informed agents. On those lines, Rothschild and Stiglitz (1976) were the first to model a non-cooperative game between uninformed insurance companies, acting as principals, and offering exclusive contracts to a buyer who knows her own risk type and chooses the contract that is best for her.

The purpose of this chapter is to pursue this second line of research. In this respect, we do not aim to revisit the original approaches to screening, nor to propose an exhaustive survey.[2] We rather start from the remark that situations in which competitors post mechanisms instead of prices are prominent in several modern markets. Examples include: competitive insurance, competing auctions, financial markets (over-the-counter [OTC] markets, interbank market), etc. In some of these markets, the posted mechanisms potentially involve some degree of reciprocity (meet the competition clause) and exclusivity of contractual relationships is not enforceable. A typical example is given by OTC markets where little information on trading volumes is available.[3]

We choose to focus on some selected contributions in the screening approach that may contribute to our understanding of the institutional features described above. As in Rothschild and Stiglitz (1976), we consider settings in which market equilibria may or may not exist and second-best optimality is not guaranteed, depending very much on the specific extensive form and on the contracting assumptions involved.

In this respect, observe that we will only consider optimality and constrained optimality in the typology defined by Holmstrom and Myerson (1983). For instance we will stick to *ex post* optimal mechanisms in the classical sense of leading to a Pareto-optimal allocation at every state of nature without taking into account the information derived from observing this allocation (*i.e.* we do not deal with the Forges' 1994 notion of posterior efficiency). Also, we will not review the Walrasian approaches to markets under incomplete information as modeled in Prescott and Townsend (1984).[4]

Specifically, this chapter is divided into two parts. In the first part (Section 2), we adopt the simple mechanism design approach with only one "mechanism designer." When the mechanism designer is an outsider (say a public authority), all traders may have private information and play simultaneously. When the mechanism designer is an insider (a principal, buyer or seller), then he is uninformed and has no private information. Three illustrative applications are introduced: bilateral trade, auctions and insurance. In the second part (Section 3), the model is extended to several principals who are uninformed and have no private information, but compete by designing mechanisms.

2 THE MECHANISM DESIGN APPROACH

Our first approach to trading rules under incomplete information is based on mechanism design. The focus will be on efficiency and incentive efficiency (Holmstrom and Myerson, 1983). Participation constraints will also be taken into account.

---

2 Riley (2001) provides an excellent retrospective of signaling and screening models.
3 Other examples include the US credit card industry (Rysman, 2007), the US life insurance market (Philipson and Cawley, 1999), and the UK annuity market (Finkelstein and Poterba (2002, 2004).
4 As shown by Rustichini and Siconolfi (2008) such an approach works well if types are publicly known but not under adverse selection. Then prices have to depend on types and to be incentive compatible: each type of consumer should want to buy in the market at the corresponding price.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

2.1 A Trading Mechanism

We consider the following scenario. There is a set $\mathcal{I}$ of agents, who are trading: $\mathcal{I} = \{1, \cdots, i, \cdots, I\}$. They are buyers and sellers. The characteristic, or type θ^i, of a trader $i \in \mathcal{I}$, takes values in a set Θ^i. We denote $\Theta = \underset{i \in N}{\times} \Theta^i$ the set of all states of nature (and $\Theta^{-i} = \underset{i \neq j}{\times} \Theta^j$). In this section, we suppose that there is a single principal (the mechanism designer), who may or may not be an outsider (a planner). An allocation is a vector $x = (x^1, \cdots, x^i, \cdots, x^I)$ in some set X, the set of feasible allocations. Each trader i may decide to participate, by taking the decision $a^i = Y$ (yes), or not to participate and take the decison $a^i = N$ (no). The utility for trader $i \in \mathcal{I}$ is given by the real-valued function $u^i(x, a; \theta)$, defined on the set $X \times \{Y, N\}^{\sharp \mathcal{I}} \times \Theta$. We assume that the utility of a trader i, when not participating, $u^i(x, N, a^{-i}; \theta)$ is given by the utility $u_0^i\left(\theta^i\right)$ of some outside option and, without loss of generality, we suppose $u_0^i\left(\theta^i\right) \equiv 0$. Observe that, otherwise, the utility of each trader, as defined, might be affected by the types of all others (common value or interdependent values). A particular case, the private value case, is when $u^i(x, a; \theta) \equiv v^i(x, a; \theta^i)$. Each agent $i \in \mathcal{N}$ knows her true type $\theta^i \in \Theta^i$ and we assume that there is a distribution F on the random variable θ, which is common knowledge (beliefs $F(\theta^{-i} \mid \theta^i)$ are *consistent*). Beliefs are *free* if $F(\theta^{-i} \mid \theta^i) = F(\theta^{-i} \mid \theta'^i)$, $\forall i \in \mathcal{N}$, $\forall \theta^i, \theta'^i \in \Theta^i$ (the so-called independent case).

The principal is supposed to choose a trading mechanism in some set Γ. A mechanism is a pair (M, γ) with $M = \underset{i \in \mathcal{I}}{\times} M^i$, each M^i being the set of messages available to trader i, together with a function $\gamma : M \times \{Y, N\}^{\sharp \mathcal{I}} \to X$. Given a mechanism, each trader $i \in \mathcal{I}$ sends a message m^i to the principal and chooses a decision $a^i \in \{Y, N\}$. For $m = (m^1, \cdots, m^I) \in M$ and $a \in \{Y, N\}^{\sharp \mathcal{I}}$, the allocation $\gamma(m, a) \in X$ is the resulting allocation. A trading mechanism determines a game with incomplete information. A *Bayesian equilibrium* is a vector of strategies $(\tilde{m}, \widetilde{a}) = (\tilde{m}_i, \widetilde{a}_i)_{i \in \mathcal{I}}$ where, for every $i \in \mathcal{I}$, $\tilde{m}^i$ is a measurable function from Θ^i into M^i and $\widetilde{a}^i$ a measurable function from Θ^i into $\{Y, N\}$ such that: $\forall i \in \mathcal{I}$, $\forall \theta^i \in \Theta^i$, $\forall m^i \in M^i$, $\forall a^i \in \{Y, N\}$,

$$\int_{\Theta^{-i}} \left(u^i(\gamma(\tilde{m}(\theta), \widetilde{a}(\theta)), \widetilde{a}(\theta); \theta)\right) dF\left(\theta^{-i} \mid \theta^i\right)$$

$$\geq \int_{\Theta^{-i}} \left(u^i(\gamma(m^i, \tilde{m}^{-i}(\theta^{-i}), a^i, \widetilde{a}^{-i}(\theta^{-i})), a^i, \widetilde{a}^{-i}(\theta^{-i}); \theta)\right) dF\left(\theta^{-i} \mid \theta^i\right).$$

In addition, this equilibrium should ensure participation, that is: $\widetilde{a}^i(\theta) = Y$, $\forall i \in \mathcal{I}$, $\forall \theta \in \Theta$.

Now, by the revelation principle, we can as well consider the associated *direct trading mechanism* $(\Theta, \widetilde{\gamma})$ such that $\widetilde{\gamma}(\theta) \equiv \gamma(\tilde{m}(\theta), \widetilde{a}(\theta))$ where $\widetilde{a}^i(\theta) = Y$, $\forall i \in \mathcal{I}$, $\forall \theta \in \Theta$, and where the vector of strategies $\tilde{\theta} = (\tilde{\theta}^1, \tilde{\theta}^2, \cdots, \tilde{\theta}^i, \cdots, \tilde{\theta}^I)$ of reporting truthfully its type for each i, *i.e.* $\tilde{\theta}^i\left(\theta^i\right) = \theta^i$, is a Bayesian equilibrium in the associated game of incomplete information. For simplicity of notation we let, for all $x \in X$ and $\theta \in \Theta$, $U^i(x; \theta) \equiv u^i(x, Y; \theta)$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

The mechanism $(\Theta, \widetilde{\gamma})$ then satisfies *Bayesian incentive compatibility* (BIC):[5] for all $\theta^i \in \Theta^i, \theta'^i \in \Theta^i$ and $i \in \mathcal{N}$,

$$\int_{\Theta^{-i}} \left[U^i\left(\widetilde{\gamma}\left(\theta\right);\theta\right)\right] dF(\theta^{-i} \mid \theta^i) \geq \int_{\Theta^{-i}} \left[U^i\left(\widetilde{\gamma}\left(\theta'^i, \theta^{-i}\right);\theta^i, \theta^{-i}\right)\right] dF(\theta^{-i} \mid \theta^i),$$

as well as *interim individual rationality (IIR)*:

$$\int_{\Theta^{-i}} \left[U^i\left(\widetilde{\gamma}\left(\theta\right);\theta\right)\right] dF(\theta^{-i} \mid \theta^i) \geq 0,$$

for all $\theta^i \in \Theta^i$ and all $i \in \mathcal{N}$.

When the principal is an outsider, say a public decision maker, her objective function might be chosen so that some efficiency property be satisfied. For example, following Holmstrom and Myerson (1983), we may define the following social welfare function on direct mechanisms

$$W(\widetilde{\gamma}) = \sum_i \int_{\Theta} \mu^i\left(\theta\right)\left[U^i\left(\widetilde{\gamma}\left(\theta\right);\theta\right)\right] dF(\theta^{-i} \mid \theta^i),$$

with every weight $\mu^i(\theta)$ non-negative and some strictly positive. Maximizing W on the set of direct mechanisms leads to *ex post* (Pareto) efficiency. If, for every $i \in \mathcal{I}$, $\mu^i(\theta) = \mu^i\left(\theta^i\right)$, then we get the stronger property of *interim* efficiency, and if for every $i \in \mathcal{I}$, $\mu^i(\theta) = \mu^i$, then we get the even stronger property of *ex ante* efficiency. If the maximization is only done over the subset of BIC direct mechanisms, we get, respectively, the properties of *ex post*, *interim* and *ex ante* incentive efficiency.

Example 1: The Quasi-linear Case Suppose that $X \subset \mathbb{R}^{(K+1)I}$ is the set of the feasible allocations of a finite number K of goods and of the corresponding monetary transfers. An element $(q, t) \in X$ is such that $q = \left(q^1, \ldots, q^i, \ldots, q^I\right) \in \mathbb{R}^{KI}$ and $t = \left(t^1, \ldots, t^i, \ldots, t^I\right) \in \mathbb{R}^I$. Each trader utility function u^i is assumed to be separable and transferable in money and can be written as $u^i(q, a; \theta) + t^i$ (resp. $U^i(q;\theta) + t^i$). In that case a trading mechanism is a triple (M, χ, τ) where $\chi : M \to X$ is called the *allocation rule* and $\tau : M \to \mathbb{R}^{\mathcal{N}}$ is the *payment scheme*. The *direct trading mechanism* associated with the Bayesian equilibrium $(\tilde{m}(\theta), \widetilde{a}(\theta))$ is the triple (Θ, q, t) where q is the allocation rule $q(\theta) \equiv \chi(\tilde{m}(\theta), \widetilde{a}(\theta))$ and t is the payment scheme such that $t^i(\theta) \equiv \tau^i(\tilde{m}(\theta), \widetilde{a}(\theta))$.

The allocation rule q is said to be *(ex post) efficient (EF)* if, for all $q \in Q$ and all $\theta \in \Theta$,

$$\sum_{i \in \mathcal{N}} U^i(q(\theta);\theta) \geq \sum_{i \in \mathcal{N}} U^i(q;\theta),$$

[5] Bayesian incentive compatibility (BIC) is the terminology used in d'Aspremont and Gérard-Varet (1979a) (see also Myerson, 1982). In the following, when we say incentive compatible it will mean BIC.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

The payment scheme t is *budget-balancing (BB)*, if for all $\theta \in \Theta$,

$$\sum_{i\in\mathcal{I}} t^i(\theta) = 0.$$

These two properties taken together imply that the allocation resulting from the mechanism is *Pareto optimal*. In the private value case, a most well-known *ex post* efficient mechanism is the Vickrey-Clarke-Groves (VCG; Vickrey, 1961, Clarke, 1971, Groves, 1973) mechanism. Take any *ex post* efficient allocation rule q^* and define the payment scheme

$$t^i_{VCG}(\theta) = \sum_{j\neq i} U^j(q^*(\theta);\theta^i) + h^i\left(\theta^{-i}\right),$$

where $h^i\left(\theta^{-i}\right)$ is any function independent of θ^i. Because q^* is efficient, agents report truthfully their types (individual and collective objectives coincide) whatever their beliefs. It is a dominant strategy. Choosing high enough h^is ensures interim individual rationality. The problem is that budget balance is generally not achievable. If $\sum_{i\in\mathcal{I}} t^i_{VCG}(\theta) > 0$ (resp. $\sum_{i\in\mathcal{I}} t^i_{VCG}(\theta) < 0$) for some $\theta \in \Theta$, then the mechanism runs a deficit (resp. a surplus).

Example 2: The Private Value Linear Case A subcase (extensively used in applications) is to assume linear utilities and private values and that, for every i, $\theta_i \in \Theta_i = \left[\theta^i_0, \theta^i_1\right]$, a non-degenerate interval in $\mathbb{R}$, and that i's beliefs are free and represented by a continuous density function $f^i\left(\theta^{-i}\right)$ with full support. The utility U^i of trader i (assuming participation) is now of the form $U^i(q)\,\theta^i + t^i$. A well-known result,[6] characterizing BIC mechanisms, is given by the following lemma:

Lemma 1 *A direct trading mechanism* (Θ, q, t) *is BIC if and only if*

$$\overline{U}^i\left(\theta^i\right)\theta^i + \overline{t}_i\left(\theta^i\right) = \overline{U}^i\left(\theta^i_0\right)\theta^i_0 + \overline{t}^i\left(\theta^i_0\right) + \int_{\theta^i_0}^{\theta^i} \overline{U}^i\left(\widehat{\theta}^i\right)d\widehat{\theta}^i,$$

where $\overline{t}^i\left(\theta^i\right) \equiv \int_{\Theta^{-i}} t^i\left(\theta^i, \theta^{-i}\right) f^i\left(\theta^{-i}\right) d\theta^{-i}$ and with $\overline{U}^i\left(\theta^i\right) \equiv \int_{\Theta^{-i}} U^i\left(q\left(\theta^i, \theta^{-i}\right)\right) f^i\left(\theta^{-i}\right) d\theta^{-i}$ a non-decreasing function, since by BIC, $\overline{U}^i\left(\theta^i\right)\theta^i + \overline{t}^i\left(\theta^i\right)$ is a convex (almost everywhere differentiable) function and its derivative is equal to $\overline{U}^i\left(\theta^i\right)$.

2.2 Applications

To illustrate this general model, we turn now to three applications.

[6] This is due to Myerson (1981) and Riley and Samuelson (1981). A characterization under efficiency is given in d'Aspremont and Gérard-Varet (1979b).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

2.2.1 Bilateral trade

If we consider the situation where one seller tries to sell an object to several potential buyers, we can further specify Example 2 by assuming (i) that trader 1 is the seller, all other traders being potential buyers, (ii) that an allocation $q = \left(q^1, q^2, \ldots, q^I\right) \in Q$ determines the probability $q^i \geq 0$ that trader i will get (or keep) the object (with $\sum_{i \in \mathcal{I}} q^i = 1$), and (iii) that the utility function is simply $U^i(q;\theta) = q^i \theta^i$.

For this context, we can adopt the following specification of the functions h^is in the definition of the VCG mechanism (see Krishna, 2010), for some efficient trading rule q^*,

$$h^i\left(\theta^{-i}\right) \equiv -\sum_{j \neq i} q^{*j}\left(\theta_0^i, \theta^{-i}\right)\theta^j - q^{*i}\left(\theta_0^i, \theta^{-i}\right)\theta_0^i.$$

Denote VCG^0 the VCG mechanism with this specification. It satisfies EF and BIC, and by Lemma 1 we get IIR, since

$$\overline{U}^i\left(\theta^i\right)\theta^i + \bar{t}^i_{VCG^0}\left(\theta^i\right) = \int_{\Theta^{-i}} \left[\sum_{j \in \mathcal{I}} q^{*j}\left(\theta^i, \theta^{-i}\right)\theta^j - \sum_{j \neq i} q^{*j}\left(\theta_0^i, \theta^{-i}\right)\theta^j - q^{*i}\left(\theta_0^i, \theta^{-i}\right)\theta_0^i\right] f^i\left(\theta^{-i}\right) d\theta^{-i},$$

is equal to zero for $\theta^i = \theta_0^i$. Also, by Lemma 1 again, for any other mechanism (Θ, q^*, t) satisfying EF, BIC and IIR, $\bar{t}^i\left(\theta^i\right) - \bar{t}^i_{VCG^0}\left(\theta^i\right)$ is a non-negative constant for each i. This observation implies the following:

Proposition 1 (Myerson and Satterthwaite, 1983) *Supposing that $I = 2$, $\theta_0^2 < \theta_1^1$ and $\theta_1^2 \geq \theta_0^1$, there is no direct mechanism satisfying EF, BIC, IIR and, at the same time, balancing the budget (BB).*

Indeed, the VCG^0 mechanism, as just defined, always runs a deficit (with the supposed overlapping intervals $\left[\theta_0^i, \theta_1^i\right]$, $i = 1, 2$), and so does any other mechanism satisfying EF, BIC and IIR (see Krishna and Perry, 1997 and Krishna, 2010).

There are various ways to escape this impossibility result. One is to assume[7] that the VCG^0 mechanism, as just defined, runs a surplus (as it would be the case here if $\theta_0^2 \geq \theta_1^1$). Another is to have more than two agents and to vary the ownership shares of the object (Cramton, Gibbons, and Klemperer, 1987) or to allow for interdependent beliefs (as we will see below). Also, the budget balance condition can be weakened to expected budget balance (McAfee and Reny, 1992).

[7] See Krishna and Perry (1997) (Theorem 2) and (under more specific assumptions) Makowski and Mezzetti (1994) (Theorem 3.1). The argument uses a modified "expected externality" mechanism (or AGV mechanism for Arrow, 1979 and d'Aspremont and Gérard-Varet, 1979a). See also d'Aspremont and Gérard-Varet (1975).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

2.2.2 Auctions

Auctions are widely used in practice for selling a large variety of objects. Consider again the situation where one seller tries to sell an object to several potential buyers with private values. The seller is now supposed to be the mechanism designer (or the principal) and his type is common knowledge. The set of messages M^i that buyer i sends to the seller is the set of possible bids. Several kinds of auctions are possible: first-price sealed-bid auction (where the winner pays the highest bid), second-price sealed-bid auction (where the winner pays the second highest bid), English (or open ascending) auction are common examples. The revelation principle still applies and to each equilibrium of an auction mechanism (direct or not) we can associate a BIC direct auction mechanism.

Assume that properties (i), (ii) and (iii) of the previous subsection still hold, that $\theta_1^0 = \theta_1^1 = \theta_0^i = 0$, $i = 2, \ldots, I$ and that $f(\theta) = \prod_{i=2}^{I} f^i\left(\theta^i\right)$. In this case, the mechanism VCG^0 coincides with the second-price direct auction (and with Clarke's "pivotal" mechanism). The object is sold to a buyer i reporting the highest valuation (i is a "pivotal" agent) and the amount this buyer i pays to the seller, $-\bar{t}^i_{VCG^0}\left(\theta^i\right)$, is equal to $\max_{j \neq i, j > 1} \theta^j$ the second-highest valuation. The other buyers pay nothing. Clearly EF and IIR hold and, since the seller receives $\bar{t}^i_{VCG^0}\left(\theta^i\right)$, we get budget balance without affecting incentives since the type of the seller is common knowledge. All buyers report truthfully (it is the dominant strategy) whatever their beliefs. The mechanism is independent of the traders' characteristics (names, valuations and beliefs) and of the object characteristics: it is anonymous and universal. Also, since values are private and statistically free, the second-price sealed-bid auction is equivalent to the English auction (since the information obtained during the latter auction is irrelevant). The first-price sealed-bid auction (even when formulated as a direct auction mechanism) is different. It satisfies BB and IIR but not, in general, BIC and EF. Vickrey (1961) already mentions the possibility of inefficient allocation in a first-price auction. A simple argument (see Krishna, 2010) is to suppose, with two asymmetric bidders, that the equilibrium bidding strategies are continuous increasing and strictly unequal at some value, say $\tilde{m}^1(\theta) < \tilde{m}^2(\theta)$, for $\theta_0^i < \theta < \theta_1^i$, $i = 1, 2$. Then $\tilde{m}^1(\theta + \varepsilon) < \tilde{m}^2(\theta - \varepsilon)$, for small $\varepsilon > 0$, and bidder 2 still wins although she has a lower value.

However, if we assume symmetry ($\theta_1^i = \theta_1^j$ and $f^i = f^j$, $i, j = 2, \ldots, I$), the first-price auction is *ex post* efficient and, by the revenue equivalence principle (Riley and Samuelson, 1981, Myerson, 1981) the expected revenue of the seller is the same as in any other *ex post* efficient auction, although BIC remains violated (each buyer reports $\theta^i/2$).

More generally, from a mechanism design perspective, one can look for an *ex ante* incentive-efficient auction mechanism. In particular, considering that the seller plays the role of a principal (and as such is the mechanism designer), we can look for the direct auction mechanism maximizing the seller's expected revenue. Following Myerson (1981), and assuming the virtual valuation function $\psi^i\left(\theta^i\right) \equiv \theta^i - \left(1 - F^i\left(\theta^i\right)\right)/f^i\left(\theta^i\right)$ to be increasing, the optimal direct auction is the one attributing the object to the buyer with maximal virtual valuation (if non-negative) and the winner pays the smallest amount that keeps him winning. Since the virtual valuation differs from the value and may be negative, the optimal auction is not efficient in general. In the symmetric case $\psi^i \equiv \psi$ for all i and the optimal auction is simply a second-price auction with reserve price equal to $\psi^{-1}(0)$ (see Proposition 5.4 in Krishna, 2010). The optimal auction is obtained under the

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

assumption of statistical independence, and in that case each buyer always benefits from some informational rent.

Myerson (1981) and Crémer and McLean (1985, 1988) show that some correlation between types is necessary and sufficient for the seller to do much better, namely to extract the full surplus. This result holds even if values are interdependent, namely if we assume the more general utility function $U^i(q;\theta) = q^i U^i(\theta)$ (with the normalization $U_0^i(\theta^i) \equiv 0$). Suppose for simplicity, as in Crémer and McLean (1988), that the set of types Θ^i for each i is finite and that the beliefs $F(\theta^{-i} \mid \theta^i)$ are represented by discrete probability distributions with full support and satisfy a very general condition (implying statistical dependence). To extract the whole surplus means that, for each buyer i, $q^i(\theta) = 0$ if $U^i(\theta) < max_j U^j(\theta)$ or if $max_j U^j(\theta) \leq 0$ (recall $\sum_{i \in \mathcal{I}} q^i(\theta) = 1$) and that each buyer's IIR constraints holds as an equality.

Crémer-McLean's condition simply requires that for each buyer i of type θ^i there is a lottery $s^i(\theta^{-i};\theta^i)$ defined on Θ^{-i} such that

$$\sum_{\theta^{-i}} s^i(\theta^{-i};\theta^i) F(\theta^{-i} \mid \theta^i) > \sum_{\theta^{-i}} s^i(\theta^{-i};\theta^i) F(\theta^{-i} \mid \theta^{\circ i})$$

for every $\theta^{\circ i} \neq \theta^i$. Under this condition, a direct auction mechanism can be constructed to extract the whole surplus, while satisfying BIC and IIR (but not *ex post* individual rationality).[8]

The Crémer-McLean condition can be reinforced and applied in a much larger context (with quasi-linear utilities), in particular to general (direct) trading mechanisms with possibly multiple buyers and sellers, all of multiple types, and with any allocation rule, efficient or not. This is most simply obtained by reinforcing condition B in d'Aspremont and Gérard-Varet (1982) so that IIR can be ensured, in addition to BIC and BB, whenever, for a given allocation rule, the *ex ante* expected surplus is non-negative:

$$\sum_{i \in \mathcal{I}} \sum_{\theta} U^i(q(\theta);\theta) F(\theta^{-i} \mid \theta^i) F(\theta^i)) \geq 0.$$

The condition (introduced as condition B^{IIR} in d'Aspremont-Crémer, 2017) requires that there exists a budget-balanced payment scheme s for all i, all θ^i and $\theta^{\circ i}$, $\theta_i \neq \theta^{\circ i}$, such that

$$\sum_{\theta^{-i}} s^i(\theta^{-i}, \theta^i)\ F(\theta^{-i} \mid \theta^i) > \sum_{\theta^{-i}} s^i(\theta^{-i}, \theta^{\circ i}) F(\theta^{-i} \mid \theta^i),$$

and

$$\sum_{\theta^{-i}} s^i(\theta^{-i}, \theta^i) F(\theta^{-i} \mid \theta^i) = 0.$$

[8] This condition is here stated in its "primal form." It is generic in the finite case. But it implies no freeness: for all i, and any $(\theta^i, \theta^{\circ i})$, if $\theta^i \neq \theta^{\circ i}$ then $F_i(. \mid \theta^i) \neq F_i(. \mid \theta^{\circ i})$, *i.e.* no agent has free beliefs over two types. This is called "belief announcement" in Johnson Pratt, and Zeckhauser (1990) and "beliefs determine preferences" (BDP) in Heifetz and Neeman (2006).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

This condition can be shown to be equivalent to the Crémer-McLean condition plus an identifiability condition introduced by Kosenok and Severinov (2008), entailing that, for any allocation rule generating a non-negative *ex ante* expected surplus, any distribution of this whole surplus (with BB respected) among the traders can be implemented by a BIC and IIR mechanism:

Proposition 2 *Assume $I \geq 3$ and that the beliefs satisfy condition B^{IIR}. For any allocation rule q and any set $\left\{\upsilon^i\left(\theta^i\right), \theta^i \in \Theta_i \text{ and } i \in \mathcal{I}\right\}$ of non-negative utility levels such that*

$$\sum_{i\in\mathcal{I}}\sum_{\theta^i}\upsilon^i\left(\theta^i\right)F^i\left(\theta^i\right)=\sum_{i\in\mathcal{I}}\sum_{\theta}U^i\left(q\left(\theta\right);\theta\right)F\left(\theta^{-i}\mid\theta^i\right)F\left(\theta^i\right)\geq 0,$$

there is a direct trading mechanism (q,t) satisfying BIC, BB, IIR and such that

$$\sum_{\theta^{-i}}\left[U^i\left(q\left(\theta\right);\theta\right)+t^i\left(\theta\right)\right]F\left(\theta^{-i}\mid\theta^i\right)=\upsilon^i\left(\theta^i\right),$$

for all $\theta^i \in \Theta i$ and $i \in \mathcal{I}$.

This is equivalent to Corollary 1 in Kosenok and Severinov (2008). Of course, this allows for full surplus extraction by a single trader. Matsushima (2007) has a similar result, but under an assumption that is stronger than B^{IIR}. The proof of the proposition is simple. Since

$$\sum_{i\in\mathcal{I}}\sum_{\theta}U^i\left(q\left(\theta\right);\theta\right)F\left(\theta^{-i}\mid\theta^i\right)F\left(\theta^i\right)=\sum_{i\in\mathcal{I}}\sum_{\theta^i}\upsilon^i\left(\theta^i\right)F\left(\theta^i\right),$$

there is a payment scheme τ satisfying BB and such that

$$\sum_{\theta^{-i}}\left[U^i\left(q\left(\theta\right);\theta\right)+\tau^i\left(\theta\right)\right]F\left(\theta^{-i}\mid\theta^i\right)\geq\upsilon^i\left(\theta^i\right),$$

for all θ^i, all i (see Lemma 1 in Matsushima, 2007). Now, using the budget-balanced payment scheme s given by condition B^{IIR}, we have a family of budget-balanced payment schemes $t \equiv \tau + Ks$, $K \geq 0$. With K large enough, BIC is satisfied and, since $\sum_{\theta^{-i}} Ks^i(\theta^{-i},\theta^i)F(\theta^{-i} \mid \theta^i) = 0$, we get IIR:

$$\sum_{\theta^{-i}}\left[U^i\left(q\left(\theta\right);\theta\right)+\tau^i\left(\theta\right)\right]F\left(\theta^{-i}\mid\theta^i\right)=\upsilon^i\left(\theta^i\right)\geq 0.$$

The mechanisms that are thus obtained, optimal or not, are interesting from an investigation point of view. As mentioned by Wilson (1985),"it suffices in principle to study direct revelation games in order to find efficient trading rules." But, most importantly, "There often remains a motive, of course, to translate an efficient direct revelation game back

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

into a form of the sort more usually found in practice."[9] Although direct mechanisms use the simplest kind of equilibrium (truth-telling), their rules integrate specific features of the economic environment.[10] They are neither universal, nor anonymous. Although there might be exceptions (*e.g.* a group of advertisers competing on the web for an ad impression), in most standard contexts, trading rules have to be more simple (*e.g.* independent of the number of participants and of their beliefs), and the complexity is shifted to the equilibrium strategies, to be computed by the participants themselves on the basis of their knowledge of the economic environment.

2.2.3 Insurance

In the first two applications we have developed, traders are restricted to being risk neutral. All utility functions are assumed to be quasi-linear (or even linear). If one looks at a single insurer selling an insurance policy to a single buyer with private information about the risk she wants to insure (hence of several types), risk aversion should be an essential ingredient. Our point of departure is the celebrated Rothschild and Stiglitz (1976) insurance economy, as reformulated by Stiglitz (1977) in a monopolistic setting with a single risk-neutral seller (the principal) offering coverage-premium contracts $(q, t) \in \mathbb{R}^2_+$ to ensure BIC and hence having the power to screen different types (the no-trade contract being $(0, 0)$).[11]

We suppose that the buyer (the agent) may be of two types, $\theta \in \{\theta^0, \theta^1\}$, with positive probabilities $F\left(\theta^1\right) = \phi$ and $F\left(\theta^0\right) = (1 - \phi)$. She has initial wealth W_0 and faces the risk of a loss $L > 0$ with a probability given by her type $\theta \in (0, 1)$ and such that $\theta^1 > \theta^0$ and $L < W_0$. Type θ's preferences over aggregate coverage-premium pairs have an expected-utility representation

$$U(q, t; \theta) \equiv \theta u(W_0 - L + q - t) + (1 - \theta)u(W_0 - t), \tag{15.1}$$

where u is a twice continuously differentiable, strictly increasing, and strictly concave von Neumann–Morgenstern utility function. One can check that, since $\theta^1 > \theta^0$, type θ's preferences over coverage-premium pairs $(q, t) \in \mathbb{R}^2_+$ are ordered by single crossing. That is, geometrically, in the (q, t) plane, an indifference curve for type θ^0 crosses an indifference curve for type θ^1 only once from below, implying that her willingness to substitute coverage for premium is everywhere higher than type θ^1's.[12] If the principal provides type θ with coverage q for a premium t, he earns a profit $\pi\,(q, t\,; \theta) = t - v(\theta)q$, with $v(\theta^1) > v(\theta^0)$, and we let $v = \phi v(\theta^1) + (1 - \phi)v(\theta^0)$ be the average price. Thus, this is a case of *common value*: conditional on a trade taking place, the insurer directly cares about the characteristics, or information, of the insured.

If the probability of loss was common knowledge, the insurer (knowing the type θ) would offer a contract $(q\,(\theta)\,, t\,(\theta))$ to each type θ with full coverage $(q\,(\theta) = L)$ and with the

---

9 See Wilson (1985) p. 183.

10 Note that these features can be very general if the type space is rich enough.

11 An alternative approach involves the informed agent moving first. This is the signaling problem originally analyzed by Spence (1973) in a labor market context, and further examined by Cho and Kreps (1987). In this class of problems the informed agent is given little, if any, power to design incentive contracts.

12 More formally, the single-crossing assumption states that: for each $(q, t) \in \mathbb{R}^2_+$, $\tau_{\theta^0}(q, t) > \tau_{\theta^1}(q, t)$, where $\tau_\theta \equiv -\frac{\partial U_\theta/\partial q}{\partial U_\theta/\partial t}$ is type θ's marginal rate of substitution of coverage for premium, which is everywhere well defined and strictly decreasing along her indifference curves.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

participation constraint binding: $U(q(\theta), t(\theta); \theta) = U_\theta(0, 0; \theta)$. This is first best optimal, the risk-neutral seller bears all the risk and extracts the whole surplus. Note that, if contracts were restricted to be actuarially fair, *i.e.* $t(\theta) - v(\theta) L = 0$, the expected profit of the seller would be zero and the surplus would go to the buyer: $U(q(\theta), t(\theta); \theta) > U(0, 0; \theta)$.

With private information, the first best becomes unfeasible. Given the incentive constraints, only the high-risk type can be fully insured in a monopolistic equilibrium.

Only the high-risk type is fully insured. If the low-risk type were buying more coverage at better terms, the high-risk type would switch to that contract and the seller would lose profit. Therefore, the seller maximizes total expected profit $\phi\pi\left(q\left(\theta^1\right), t\left(\theta^1\right); \theta^1\right) + (1-\phi)\,\pi\left(q\left(\theta^0\right), t\left(\theta^0\right); \theta^0\right)$ under the BIC and IIR constraints. At the solution the low-risk type gets partial (or zero) insurance coverage, with the BIC constraint being strict,

$$U(q\left(\theta^0\right), t\left(\theta^0\right); \theta^0) > U(q\left(\theta^1\right), t\left(\theta^1\right); \theta^0)$$

and no surplus: $U(q\left(\theta^0\right), t\left(\theta^0\right); \theta^0) = U(0, 0; \theta^0)$. The high-risk type, indeed, gets full coverage, with the BIC constraint binding,

$$U(q\left(\theta^1\right), t\left(\theta^1\right); \theta^1) = U(q\left(\theta^0\right), t\left(\theta^0\right); \theta^1)$$

and positive surplus: $U(q\left(\theta^1\right), t\left(\theta^1\right); \theta^1) > U(0, 0; \theta^1)$.

This characterization of the monopoly (second-best) allocation has been recently generalized by Schlee and Chade (2012), in which the set Θ of buyer's types is not restricted to be finite (but with θ_0 the smallest and θ_1 the largest element) and the distribution $F(\theta)$ is arbitrary. The seller's problem can be written in more general terms: $\underset{q(\theta),t(\theta)}{max} \int_\Theta \pi\left(q(\theta), t(\theta); \theta\right) dF(\theta)$, subject to $U(q(\theta), t(\theta); \theta) \geq U(q\left(\theta'\right), t\left(\theta'\right); \theta)$ and $U(q(\theta), t(\theta); \theta) \geq U(0, 0; \theta)$, for all θ, θ' in Θ. As in Stiglitz (1977), type θ_0 gets no surplus, type θ_1 gets full coverage, all other types get partial insurance. The seller makes positive expected profit. The coverage and premium are non-negative and co-monotone.

If we now impose that the seller's total expected profit be zero, we get the set of second-best contracts as defined by Harris and Townsend (1981) and characterized for this model by Crocker and Snow (1985). These are the *ex post* incentive-efficient allocations obtained by maximizing the weighted sum $\mu U(q\left(\theta^1\right), t\left(\theta^1\right); \theta^1) + (1-\mu)\, U(q\left(\theta^0\right), t\left(\theta^0\right); \theta^0)$ under the constraint that $\phi\pi\left(q\left(\theta^1\right), t\left(\theta^1\right); \theta^1\right) + (1-\phi)\,\pi\left(q\left(\theta^0\right), t\left(\theta^0\right); \theta^0\right) = 0$, as well as the two BIC constraints, for all values of the non-negative weights μ and $(1-\mu)$. The resulting allocation depends on the relative weight μ/ϕ of the high-risk type. If $\phi \geq \mu$ (resp. $\phi \leq \mu$) the high-risk type (resp. low-risk type) receives full coverage and is indifferent to the contract received by the low-risk type (resp. high-risk type): $U(q\left(\theta^1\right), t\left(\theta^1\right); \theta^1) = U(q\left(\theta^0\right), t\left(\theta^0\right); \theta^1)$ (resp. $U(q\left(\theta^1\right), t\left(\theta^1\right); \theta^0) = U(q\left(\theta^0\right), t\left(\theta^0\right); \theta^0)$). To fix an (individually rational) allocation, the surplus can be divided among the two types (under some distributional conditions). Figure 15.1 depicts the Wilson-Miyazaki-Spence (WMS; Wilson, 1977, Miyazaki, 1977, Spence, 1978) second-best allocation $(q_\theta^{WMS}, t_\theta^{WMS})$ for $\theta \in \{\theta^0, \theta^1\}$, in which $\mu = 0$, that is, utility of the low-risk type is maximized.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

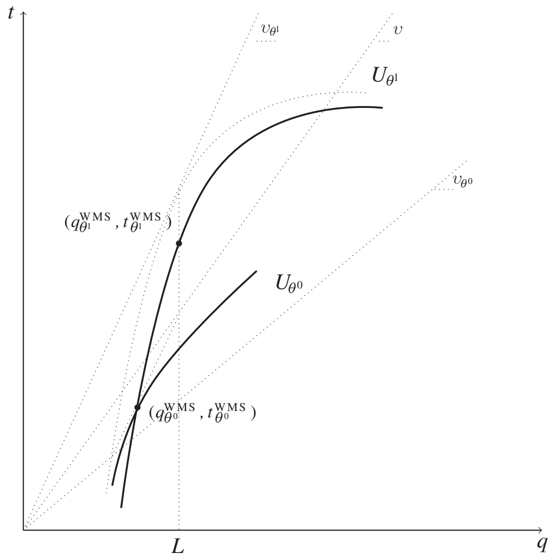


Figure 15.1 The WMS allocation

3 THE STRATEGIC APPROACH

This section revisits some recent extensions of the mechanism design approach. Specifically, we consider markets subject to incomplete information in which several parties have the power to propose incentive schemes. Our aim is then twofold. First, to investigate to what extent the design of an optimal trading mechanism by a single principal is affected by the presence of the mechanisms posted by his competitors. Second, to propose a novel approach to the study of such markets, taking into account *both* the relevant informational frictions and the decentralized nature of the contracting process. To this extent, we frame our analysis in the context of an extensive form game in which several principals compete over mechanisms in the presence of several privately informed agents. The next paragraphs introduce a general version of this game and provide a theoretical reference for the economic applications analyzed in the remaining of the section.

3.1 The Model

We refer to a scenario in which several principals (indexed by $j \in \mathcal{J} = \{1, \ldots, J\}$) contract with several agents (indexed by $i \in \mathcal{I} = \{1, \ldots, I\}$). Each agent i has private information about her type $\theta^i \in \Theta^i$ and $\theta = \{\theta^1, \cdots, \theta^I\} \in \Theta = \underset{i \in \mathcal{I}}{\times} \theta^i$ is a random variable with distribution F.

Each principal j may choose an action $x_j \in X_j$. Agents take no actions, except for their participation decisions, with $a_j^i \in \{Y, N\}$ being the decision of agent i to participate with principal j, in which $\{N\}$ stands for not participating, and we let $a^i = (a_1^i, a_2^i, \ldots, a_J^i)$. We also take $v_j : X \times A \times \Theta \to \mathbb{R}_+$ and $u^i : X \times A \times \Theta \to \mathbb{R}_+$ to be the payoff to principal j and to agent i, respectively, with $X = \underset{j \in \mathcal{J}}{\times} X_j$ and $A = \underset{i \in \mathcal{I}}{\times} A^i$. For a given array of agents' types θ,

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

of actions $a = \left(a^1, a^2, \ldots, a^I\right)$ and of principals' decisions $x = (x_1, x_2, \ldots, x_J)$, the payoffs to agent i and to principal j are $u^i\,(x, a, \theta)$ and $v_j\,(x, a, \theta)$, respectively.

Each principal perfectly observes the set of agents who participate with him. Communication is one-sided: each agent i may send a private message $m_j^i \in M_j^i$ to principal j. We let each set M_j^i be sufficiently rich to include the element $\{\emptyset\}$ corresponding to the information "agent i does not communicate with principal j," and to satisfy the standard size restriction $\sharp M_j^i > \sharp\Theta^i$ for every i and j. Principal j takes his decisions contingent on the array of messages m_j he receives, with $m_j = \left(m_j^1, m_j^2, \ldots, m_j^I\right) \in M_j = \underset{i\in\mathcal{I}}{\times} M_j^i$, and on the participation choices of the agents. Formally, we say that a mechanism proposed by principal j is the measurable mapping $\gamma_j : M_j \times \{Y, N\}^{\sharp\mathcal{I}} \to \Delta\left(X_j\right)$. We take Γ_j to be the set of mechanisms available to principal j and denote $\Gamma = \underset{j\in\mathcal{J}}{\times} \Gamma_j$. The competing mechanism game relative to Γ begins when principals publicly and simultaneously commit to mechanisms.

Given the posted mechanisms $(\gamma_1, \gamma_2, \ldots, \gamma_J) \in \Gamma$ and their privately observed types, agents simultaneously take a participation and a communication decision with respect to every principal. In this incomplete information game, a strategy for principal j is a $\gamma_j \in \Gamma_j$, and $\gamma = (\gamma_1, \cdots, \gamma_J) \in \Gamma$ is a profile of strategies for principals.

A strategy for each agent i associates to every profile of posted mechanisms γ a joint participation and communication decision. In a pure strategy, every agent participates with a subset of principals and sends a non-degenerate message only to the principals she participates with. We let $S^i = \left\{s^i \in M^i \times A^i : m_j^i = \emptyset \text{ if and only if } a_j^i = \{N\}\right\}$ be the strategy set for agent i, with $A^i = \left\{a^i = \left(a_1^i, \ldots, a_J^i\right) \in \{Y, N\}^{\sharp\mathcal{J}}\right\}$ and $M^i = \underset{j\in\mathcal{J}}{\times} M_j^i$ representing the sets of participation and communication decision, respectively. Given a profile γ of posted mechanisms, a strategy for agent i is then the measurable mapping $\sigma^i \equiv \left(\widetilde{m}^i, \widetilde{a}^i\right) : \Gamma \times \Theta^i \to S^i$, with $\widetilde{m}^i\left(\gamma, \theta^i\right) \in M^i$ and $\widetilde{a}^i\left(\gamma, \theta^i\right) \in A^i$. Every $\sigma\,(\gamma, \theta) = \left(\sigma^1\left(\gamma, \theta^1\right), \ldots, \sigma^I\left(\gamma, \theta^I\right)\right)$ induces principal j decision $\gamma_j\,(\sigma\,(\gamma, \theta)) \in X_j$ and $\gamma\,(\sigma\,(\gamma, \theta)) \in X$. The expected payoff to each type θ^i of agent i is:

$$\int\limits_{\theta\in\Theta^{-i}} u_i(\gamma(\sigma\,(\gamma, \theta)), \widetilde{a}\,(\gamma, \theta)\,, \theta^i, \theta^{-i})dF(\theta_{-i} \mid \theta_i),$$

with $F\left(\theta^{-i}|\theta^i\right)$ being the conditional probability of θ^{-i} given θ^i. The expected payoff to principal j when he plays γ_j against his opponents' strategies γ_{-j} is:

$$V_j\left(\gamma_j, \gamma_{-j}, \sigma\right) = \int\limits_{\theta\in\Theta} v_j\,(\gamma\,(\sigma\,(\gamma, \theta)), \widetilde{a}\,(\gamma, \theta)\,, \theta)\;\; dF\,(\theta)\,.$$

The strategies (γ, σ) constitute a perfect Bayesian equilibrium relative to Γ if σ is a continuation equilibrium for every γ and if, given γ_{-j} and σ, for every $j \in \mathcal{J}$: $\gamma_j \in \underset{\gamma_j'\in\Gamma_j}{argmax}\; V_j\left(\gamma_j', \gamma_{-j}, \sigma\right)$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Existence of a (potentially mixed) perfect Bayesian equilibrium of an arbitrary game Γ has been established by Page and Monteiro (2003) for the multiple agent case, and by Carmona and Fajardo (2009) for the single agent case. We focus here on characterization results. That is, we investigate how equilibrium outcomes are affected by the set of mechanisms made available to principals. In this respect, the following remarks will be useful in the balance of the analysis.

Remark 1 *A mechanism available to principal j is direct if agents can only communicate their types to principal j, i.e. if $M_j^i = \Theta^i \cup \{\emptyset\}$ for every i, with $\{\emptyset\}$ representing no communication. We denote a direct mechanism for principal j as $\tilde{\gamma}_j : \times_{i \in \mathcal{I}} \left(\Theta^i \cup \{\emptyset\}\right) \times \{Y, N\}^{\sharp\mathcal{I}} \to \Delta\left(X_j\right)$ and the set of direct mechanisms as $\Gamma_j^D \subseteq \Gamma_j$. We let G^Γ be the competing mechanism game induced by a given Γ, and G^D the game in which principals are restricted to direct mechanisms. As in Myerson (1982), a direct mechanism is incentive compatible from the point of view of principal j if, given the mechanisms offered by the other principals, it induces a continuation equilibrium in which agents truthfully reveal their types to him. A direct mechanism $\tilde{\gamma}_j$ can therefore be incentive compatible for a given array $\tilde{\gamma}_{-j}$, but not for some other $\tilde{\gamma}'_{-j} \neq \tilde{\gamma}_{-j}$. An equilibrium is truth-telling if every principal posts an incentive-compatible mechanism and agents truthfully reveal their private information to the principals they participate with, whenever this constitutes an equilibrium in their continuation game.*

Remark 2 *The model does not put any specific structure on the agents' message spaces $\left(M^i\right)_{i=1}^I$. One is therefore led to ask to what extent the corresponding equilibrium characterization depends on the available modes of communication. Indeed, an agent's report to a given principal may convey information about other principals' mechanisms and this information can be strategically exploited. This suggests that relying on a straightforward application of the revelation principle, by restricting agents to only reveal their (exogenous) private information, may involve a loss of generality. In this perspective, Epstein and Peters (1999) are the first to provide a canonical definition of the set of agents' types to which the revelation principle should apply. This set includes the agents' physical types and a component of market information, which is rich enough to describe what competitors would do under all kinds of different circumstances. Despite its relevance in terms of generality, the result also documents a fundamental difficulty in relying on simple direct mechanisms. These mechanisms indeed turn out to be too complex to be of practical use in applications.*

Remark 3 *The model focuses on "ordinary" contracting games: principals cannot design their mechanism contingent on the proposals of their rivals. An alternative possibility would be to explicitly let them commit to write "contractible" contracts, i.e. contracts that explicitly refer to each other, as done by Peters and Szentes (2012) and Szentes (2015). Such an approach, in turn, would require each of the principals to be able to monitor the entire contracting process, including all relevant off-equilibrium threats.*

3.2 Applications

Our general model encompasses several economic approaches to competition in markets subject to incomplete information, as we illustrate below.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Example 3: Trading Under Adverse Selection The simplest application of our setting features the trade between J buyers and I sellers. Each seller is endowed with one unit of a perfectly divisible good. Let q_j^i be the quantity of the good purchased by buyer j from seller i, and t_j^i the transfer he makes in return. The feasible trades $((q_1^1, t_1^1), \ldots, (q_J^1, t_J^1), \ldots, (q_1^I, t_1^I), \ldots, (q_J^I, t_J^I))$ are such that $\sum_j q_j^i \leq 1$ for all i. As in Samuelson (1984) and Myerson (1985), the profit to seller i from trading $(Q^i, T^i) = \left(\sum_j q_j^i, \sum_j t_j^i\right)$ in the aggregate is $T^i - \theta^i Q^i$, where θ^i is seller i's opportunity cost of giving away her endowment. Each buyer j's profit from trading $(\sum_i q_j^i, \sum_i t_j^i)$ is $\sum_i [v(\theta^i)q_j^i - t_j^i]$; thus, he directly cares about the identity of the sellers he is trading with through the common value component $v(\theta)$. Each seller is privately informed of her opportunity cost. As first pointed out by Akerlof (1970), in such circumstances trade is typically threatened by adverse selection whenever $v(\theta)$ increases with θ, since offering to trade at a given price then only attracts the lowest qualities. In this context, a mechanism γ_j for buyer j associates a profile of individual trades to each array of messages he receives from buyers. Mas-Colell, Whinston, and Green (1995) illustrate how this setting can be naturally exploited to model competition in several market scenarios, and Attar, Mariotti, and Salanié (2011) provide a fully strategic formulation of the multiple-buyer multiple-seller game.

Example 4: Competitive Screening In their canonical analysis of the insurance market, Rothschild and Stiglitz (1976) study strategic competition between intermediaries for the exclusive right to serve a customer facing a binary risk on her endowment $w \in \{w_L, w_H\}$, with probabilities $(\theta, 1-\theta)$ that constitute her private information. Her (expected) payoff is $pu(w_L+d_L)+(1-p)u(w_H+d_H)$, with $(d_L, d_H) \in \mathbb{R}^2$ being the state-contingent transfers issued by the company she trades with. Similarly, market-microstructure models in the tradition of Glosten (1994) consider several market makers who compete to sell shares of a risky asset to a single insider who can trade with any subset of them (Biais, Martimort, and Rochet, 2000 and Back and Baruch, 2013). The private information θ of the buyer is her willingness to trade the asset. When trading an aggregate quantity Q against an aggregate transfer T, the buyer's payoff is $\theta Q - \frac{\sigma^2}{2} Q^2 - T$, with $\sigma > 0$. The sellers are risk neutral and the cost of selling a share of the asset to type θ is its expected value conditional on the insider's being of type θ. The model of this section can hence be interpreted in terms of competitive screening by letting $I = 1, \theta \equiv (p, 1-p)$ and $\tilde{\gamma}_j : \theta \times \{Y, N\} \to \mathbb{R}^2$.

Example 5: Competing Auctions In a seminal paper, McAfee (1993) analyzes sellers who compete over auctions when buyers' valuation constitute their private information. In these settings, sellers simultaneously and anonymously post their reservation prices and buyers choose at most one auction to participate in. A seller and the buyers who participate in his auction form an isolated corporation. In addition, sellers are restricted to post direct mechanisms, asking each buyer $i \in \mathcal{I}$ to report her valuation $v^i \in [0,1]$. A strategy for seller j is a mechanism $\tilde{\gamma}_j : \left|I_j\right| \times [0,1]^{|I_j|} \to \mathbb{R}$, where $I_j \subseteq I$ is the set of buyers that participate in auction j. A pure strategy for buyer i is a mapping $\lambda^i : \Gamma_1 \times \ldots \times \Gamma_J \times [0,1] \to A^i \times [0,1] \times \mathbb{R}_+$, with Γ_j being a set of second-price auctions for $j \in J$. Given her participation decision, it is always a dominant strategy for each of the buyers to truthfully report their private valuations. Specifically, the model of this section adapts to the competing auction settings of Peters (1997), Peters and Severinov (1997), Burguet and Sakovics (1999), Viràg (2010), Han (2015) and Peck (2015).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

3.3 Equilibrium Trades

We start by taking a normative perspective. In this respect, the following paragraphs analyze a traditional issue: can second-best allocations be supported as equilibrium outcomes of our competitive setting? A positive answer to this question would provide novel insights into a reformulation of the second welfare theorem for incomplete information economies.

3.3.1 The multiple agent case

In recent years, providing a full characterization of the set of equilibrium allocations of games in which several principals have the power to design mechanisms has become a relevant issue in mechanism design. With reference to the class of extensive form games described in Section 3.1, Yamashita (2010) establishes a folk theorem: an allocation is implementable if and only if it is incentive compatible and the payoff of each principal is above a well-chosen threshold value.

We illustrate the logic underlying his result in the context of trading under adverse selection (Example 3), assuming that the type of each seller can be either low, $\theta = \theta^1$, or high, $\theta = \theta^0$, for some $\theta^0 > \theta^1 > 0$. To further simplify the exposition, we assume that the quality of the good increases with the type of the seller, that is, $v(\theta^0) > v(\theta^1)$, and that it would be efficient to trade no matter the type of the seller, that is, $v(\theta) > \theta$ for each θ. Finally, to avoid trivial cases, we assume that $x \equiv prob[\theta = \theta^0] \in (0, 1)$. Suppose that sellers communicate with buyers through the message spaces $M^1 = \ldots = M^i = \ldots = M^I = \{\theta^1, \theta^0, m\}$. We first show that a monopolistic outcome for buyers can be supported in a pure strategy equilibrium. More precisely, let p^m be the price that would be optimally set by a monopsonistic buyer.[13] Suppose now that each buyer commits to buy any quantity from each of the sellers at a constant unit price equal to p^m, unless at least $I-1$ sellers send him the message m. In this last contingency, and for every profile of sellers' participation decisions, he offers to buy a quantity of one from each of the sellers at a constant unit price $p = v(\theta^0)$ so to maximize the sellers' surplus. It is straightforward to check that p^m is supported at equilibrium by having all sellers whose type is $\theta \leq p^m$ trading a quantity of one with the same buyer, and those such that $\theta > p^m$ staying out of the market. The equilibrium is such that each seller, irrespective of her type, sends the message m to each non-deviating buyer she participates with in the subgame following a buyer's unilateral deviation, which constitutes a continuation equilibrium. A similar reasoning guarantees that every incentive-compatible allocation can be supported at equilibrium.

Unfortunately, the analysis in Yamashita (2010) does not allow the derivation of a full equilibrium characterization in general settings. A main drawback of his theorem is that threshold values are not identified in terms of the primitives of the game. Although several works have recently tried to overcome this difficulty, we still lack a general characterization result for the class of incomplete information games analyzed in this section.[14] We do not attempt to fill this gap here, but rather to point out two implications of Yamashita's insights that may be relevant for the economic applications of competing mechanism games.

13 As shown by Samuelson (1984) and Myerson (1985), who extensively analyze this trading setting in the case $I = J = 1$, an optimal mechanism for the buyer involves determining a given price p^m at which he stands ready to trade any quantity between 0 and 1. Indeed, given bilateral linearity of preferences, the buyer cannot further increase his profit by designing a direct revelation mechanism $\tilde{\gamma} : [\theta^1, \theta^0] \to [0, 1] \times \mathbb{R}_+$ which prescribes a quantity and a transfer for each revealed type.

14 See Szentes (2009), Peters and Troncoso-Valverde (2013), Xiong (2013), and the survey of Peters (2014).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

The first implication can be derived from the example above. The reasoning crucially exploits the fact that each buyer uses a message space that is "larger" than each seller's type space $\{\theta^1, \theta^0\}$. Indeed, the message m is used out of equilibrium to deter any profitable deviation by his rivals. Incentive-compatible mechanisms, such as those identified in Remark 1, are actually not rich enough to reproduce the same threats. That is, if one considers the simpler game in which buyers are only allowed to post direct mechanisms, there always exists a profitable deviation for (at least) one buyer against any equilibrium supporting the price p^m. As already documented in earlier examples (see Peck, 1997, Peters, 2001, and Martimort and Stole, 2002), restricting buyers from using incentive-compatible mechanisms involves a loss of generality: there exist *pure strategy* equilibrium outcomes of a game in which they post indirect mechanisms that cannot be reproduced by incentive-compatible ones.[15] The result suggests that the equilibrium predictions of competing mechanism models crucially depend on the set of instruments that are available to competitors. At the same time, but from a more applied standpoint, it calls for the identification of a simple class of mechanisms that allows the characterization of meaningful equilibria. Specifically, we say that a set of mechanisms is "robust" if the corresponding equilibria survive the principals' unilateral deviation towards any indirect mechanism. That is, the corresponding outcomes are supported in an equilibrium of the game in which principals use arbitrary communication mechanisms.

Characterizing the equilibrium outcomes supportable by "robust" mechanisms is relevant for several economic applications. Indeed, economic models of competing mechanisms typically restrict attention to simple incentive-compatible mechanisms. That is, principals commit to message-contingent decisions that induce agents to truthfully reveal their exogenous private information. It is therefore natural to investigate in which contexts such incentive-compatible mechanisms end up being robust. A case in point is provided by competing auction settings. In a pioneering work, Peters (1997) shows that when every seller offers a second-price auction with reserve price equal to his cost, none of his rivals can improve his profits by deviating to an alternative direct mechanism when the number of sellers gets large. The recent work by Han (2015) extends the result by establishing the robustness of second-price auctions against any arbitrary mechanism. Key to his argument is the fact that second-price auctions are dominant strategy incentive compatible. This guarantees that a best reply of a single principal to a given profile of mechanisms posted by his opponents can be characterized by an incentive-compatible mechanism. One should, however, appreciate that the result does not hold in general: Attar et al. (2012) show that two-sided communication is needed to obtain a full characterization of principals' best replies. Their analysis stresses the fact that, whenever principals compete in the face of privately informed agents, then, from the viewpoint of a single principal, the messages that agents send to his rivals can be seen as hidden actions. Given the profile of mechanisms proposed by his opponents, a principal that behaves as if he was interacting with several agents can take some non-contractible actions, *i.e.* the messages they send to the other principals. It is hence possible to show, along the lines of Myerson (1982), that he can gain by using mechanisms that induce agents to correlate on the messages they send to his opponents.

The second relevant implication of Yamashita's (2010) work can be described as follows. His folk theorem-like result exploits the presence of several, at least three, agents. Yet the strategic settings in which several principals compete to serve a single agent are at the

[15] This result is typically acknowledged as a "failure of the revelation principle" in competing mechanism games.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

center-stage of several economic applications. It is therefore natural to investigate to what extent the decentralization of second-best allocations may be successfully performed in such a restricted scenario. We perform this task in the following paragraphs.

3.3.2 The single agent case

We analyze here the situation in which principals compete in the presence of a single agent, that is, $I = 1$. It might be useful to analyze this setting in the context of competitive screening (Example 4), in which a risk-averse agent allocates her consumption over two states of nature by purchasing coverage from several risk-neutral sellers. Specifically, we refer to the insurance framework introduced in Section 2.2.3, and we aim at providing a full strategic analysis of the competition between $J \geq 2$ sellers. In this context, the payoff to the single consumer depends on the total coverage she raises from sellers and on the total premium she provides in exchange. It is therefore useful to denote $Q = \sum_{j \in \mathcal{J}} q_j$, and $T = \sum_{j \in \mathcal{J}} t_j$, with (q_j, t_j) being the coverage-premium pair she trades with seller j. Thus, considering again (15.1), we let $U(Q, T; \theta)$ be the payoff to type $\theta \in \{\theta^0, \theta^1\}$ when purchasing the *aggregate* coverage Q against the *aggregate* premium T, and to refer to the quadruple $\big((Q_{\theta^0}, T_{\theta^0}), (Q_{\theta^1}, T_{\theta^1})\big)$ as an aggregate allocation.

We frame the competitive provision of insurance in the context of a competing mechanism game. Thus, an action x_j, or contract, available to principal j is a coverage-premium pair $(q, t) \in \mathbb{R}^2_+$, and the no-trade contract is $(0, 0)$. The restriction to a single agent setting allows us to simplify the extensive form game described in Section 3.1. In such a scenario, as established by Martimort and Stole (2002) and Peters (2001), every equilibrium outcome of each game Γ can be supported at equilibrium in a "simpler" game in which principals are restricted to post arbitrary menus of contracts, with the agent choosing one item in each menu upon privately observing her type.[16] Our corresponding *menu game* unfolds as follows: first nature selects the consumer's type, then principals simultaneously post menus of coverage-premium pairs. Finally, the agent optimally takes her participation decision and picks one item in the menu of each principal.[17] We make no specific assumption on the structure of the sets of available menus. We only require that they are compact sets so that each type's choice problem admits a solution. To cope with standard applications of competitive screening, we restrict attention to pure strategy perfect Bayesian equilibria (PBE). Equilibrium trades are threatened by the following conflict. Given single crossing, the risk-taker consumer is willing to purchase a higher amount of insurance. Yet, given common values, intermediaries are rather willing to sell higher quantities to the low-risk type θ^0. The tension between these two forces is at the root of adverse selection, and may have destabilizing effects on market equilibria.

Key to our analysis is to specify the agent's participation decisions. Specifically, we say that competition is exclusive if, in a pure strategy, the agent participates with *at most* one principal, in which case she is allowed to pick *at most* one item different from $(0, 0)$, and that competition is non-exclusive otherwise. In single agent contexts, these two market structures deliver different implications for decentralization. We discuss them in the following paragraphs.

---

16 The result is often acknowledged as the delegation principle.

17 A simple way to incorporate the consumer's participation decisions in the analysis is to impose that every menu of each intermediary must include the no-trade contract $(0, 0)$. The decision not to participate with principal j therefore corresponds to choosing the item $(0, 0)$ on his menu.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Exclusive competition If competition is exclusive, our analysis mirrors that originally developed by Rothschild and Stiglitz (1976).[18] Consider the aggregate allocation $RSW = \big((Q_{\theta^0}, T_{\theta^0}), (Q_{\theta^1}, T_{\theta^1})\big)$ such that $T_\theta = v_\theta Q_\theta$ for each $\theta \in \{\theta^0, \theta^1\}$ and

$$\tau_\theta(Q_{\theta^1}, T_{\theta^1}) = v_{\theta^1} \tag{15.2}$$

$$U_{\theta^1}(Q_{\theta^1}, T_{\theta^1}) = U_{\theta^1}(Q_{\theta^0}, T_{\theta^0}). \tag{15.3}$$

The first condition states that the high-risk type θ^1 purchases her first-best allocation, and the second one that type θ^0 has to pay a cost to signal her effective quality, which corresponds to a binding BIC constraint for θ^1. The allocation is "competitive" in the sense that each intermediary gets a zero profit on each consumer's type. The allocation above is the only candidate to be supported in a pure strategy equilibrium of the exclusive competition game.[19] The allocation is depicted in Figure 15.2, which represents the buyer's and sellers' indifference curves.

As first acknowledged by Rothschild and Stiglitz (1976), however, pure strategy equilibria fail to exist in a robust number of circumstances. In these cases, a single intermediary can profitably deviate by attracting both types of consumer in such a way that his gains on the low-risk type θ^0 more than offset the losses on the higher-risk type θ^1.[20]

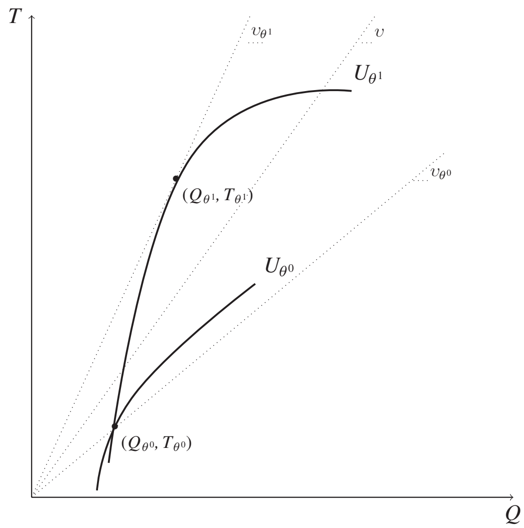


Figure 15.2 The exclusivity outcome

[18] See Attar, Campioni, and Piaser (2016) for a general analysis of competing mechanism games under exclusive competition.

[19] The proof of the result is rather standard (see, for example, Mas-Colell et al., 1995, pp. 460–465), and we do not include it here to ease exposition.

[20] See Fagart (1996) and Luz (2016) for a full characterization of the conditions needed to guarantee existence of a pure strategy equilibrium.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Non-exclusive competition Allowing for non-exclusive competition crucially modifies the strategic behavior of intermediaries. On the one hand, a larger set of deviations becomes available. Indeed, each seller can exploit the offers of his rivals by proposing insurance contracts that the consumer may use to complement her coverage. In principle, this creates new opportunities for undercutting. On the other hand, each intermediary may exploit the consumer as a coordinating device to possibly prevent his rivals' deviations. This is done by introducing additional threats that take the form of *latent* contracts in one's competitors' menu. The interplay of these two forces dramatically shapes the set of equilibrium allocations with respect to the benchmark of the scenario in which exclusivity clauses are enforced from the outset.

The recent work of Attar, Mariotti, and Salanié (2014) provides a full equilibrium analysis of the non-exclusive menu game. In general terms, they show that non-exclusivity worsens the impact of adverse selection, and pure strategy equilibria necessarily feature the market breakdown emphasized by Akerlof (1970). In a simple two-type setting, a positive level of trades for one type of the consumer only obtains if the other type does not trade at all. We revisit their arguments in the following paragraphs. We start by establishing the following:

Lemma 2 *The RSW allocation cannot be supported at equilibrium in the non-exclusive menu game.*

The intuition for the result can be easily understood in a free entry equilibrium.[21] Consider then an inactive intermediary and suppose that he deviates by offering, together with the null contract (0, 0), the additional contract $(q,t) = (\varepsilon, \varepsilon\chi)$ with ε strictly positive and $\chi \in \left(v_{\theta^1}, \tau_{\theta^1}(Q_{\theta^0}, T_{\theta^0})\right)$.[22] One can check that, since $\chi < \tau_{\theta^0}(Q_{\theta^0}, T_{\theta^0})$, ε can be chosen small enough to guarantee that $U_{\theta^1}\left(Q_{\theta^1} + \varepsilon, T_{\theta^1} + \varepsilon\chi\right) > U_{\theta^1}\left(Q_{\theta^1}, T_{\theta^1}\right) = U_{\theta^1}\left(Q_{\theta^0}, T_{\theta^0}\right)$, which ensures that type θ^1 will be trading the contract (q,t). One should observe that, since $\chi > v_{\theta^1}$, we get $t - v_{\theta^1}q = \varepsilon(\chi - v_{\theta^1}) > 0$. That is, the entrant earns a strictly positive profit on the high-risk type θ^1. Given that $v_{\theta^1} > v_{\theta^0}$, the deviation is *a fortiori* profitable if the deviating contract (q,t) is also traded by the low-risk type θ^0. This deviation is illustrated in Figure 15.3.

Thus, the deviation exploits the possibility for sellers to attract the *high*-risk type θ^1 by proposing her to trade additional insurance on top of that chosen by the *low*-risk type θ^0. The profitability of any such deviation guarantees that in any separating equilibrium one should have $\frac{T_{\theta^1} - T_{\theta^0}}{Q_{\theta^1} - Q_{\theta^0}} = v_{\theta^1}$. That is, the BIC constraint of type θ^1 turns out *not* to be binding. Furthermore, Attar et al. (2014) show that a positive level of trade for type θ^1 obtains at equilibrium only if type θ^0 is left out of the market. It follows that the only candidate to be supported in a pure strategy equilibrium of the non-exclusive menu game is the aggregate allocation $AMS = \left((Q_{\theta^0}, T_{\theta^0}), (Q_{\theta^1}, T_{\theta^1})\right)$ with $(Q_{\theta^0}, T_{\theta^0}) = (0,0)$ and $(Q_{\theta^1}, T_{\theta^1})$ such that $\tau_{\theta^1}(Q_{\theta^1}, T_{\theta^1}) = v_{\theta^1}$, and $T_{\theta^1} = v_{\theta^1}Q_{\theta^1}$. Clearly, this allocation involves a strictly positive trade for the high-risk type θ^1 only if $\tau_{\theta^1}(0,0) > v_{\theta^1}$. In the specific context of insurance, the AMS allocation is such that the high-risk type achieves her first-best level of coverage,

---

[21] See Proposition 1 of Attar et al. (2014) for a general argument.

[22] Since the marginal rate of substitution is decreasing along a given indifference curve, and provided that $Q_{\theta^1} > Q_{\theta^0}$ by single crossing, one gets $\tau_{\theta^1}(Q_{\theta^0}, T_{\theta^0}) > \tau_{\theta^1}(Q_{\theta^1}, T_{\theta^1}) = v(\theta^1)$, which guarantees that the interval is non-empty.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

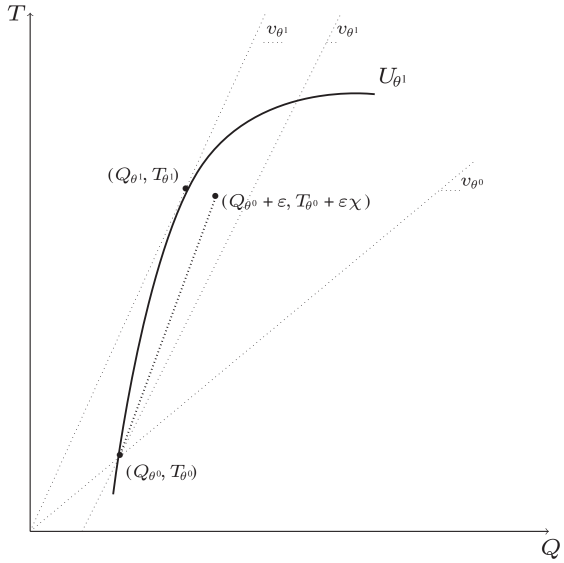


Figure 15.3 The RSW allocation is NOT an equilibrium under non-exclusivity

so that $Q_{\theta^1} = L$ and $T_{\theta^1} = v_{\theta^1}L$. Theorems 1 and 2 in Attar et al. (2014) show that a necessary and sufficient condition for existence of such an equilibrium is that, starting from the no-trade allocation (0,0), type θ^0 should *not* be willing to purchase insurance issued at the fair price $v = \phi v(\theta^1) + (1-\phi)v(\theta^0)$. More formally, they require that $\tau_{\theta^0}(0,0) \leq v$, which corresponds to Akerlof's (1970) condition for a market breakdown in which only the worse-quality goods are traded. When this condition is not satisfied, at least one seller can profitably deviate by exploiting the consumer's ability to engage in multiple trades. To clarify this point, consider a candidate-separating equilibrium in which aggregate trades are such that $Q_{\theta^0} > 0$. In this case, everything happens as if type θ^0 purchases the aggregate quantity Q_{θ^0}, and type θ^1 purchases it together with the additional insurance $Q_{\theta^1} - Q_{\theta^0}$ priced at the unit price v_{θ^1}.[23] Sellers may therefore engage in a Bertrand-like competition on the first layer, implying that Q_{θ^0} must be priced at v_{θ^0}. Overall, we get $T_{\theta^0} = Q_{\theta^0}v_{\theta^0}$, which guarantees zero profit to each of the sellers even though type θ^0 subsidizes type θ^1 at $(Q_{\theta^0}, T_{\theta^0})$. Now, since no seller is indispensable, to provide the consumer with $(Q_{\theta^0}, T_{\theta^0})$, any of them actively trading with type θ^0 has a profitable menu deviation consisting of two non-zero contracts. The first contract, targeted at θ^0, is approximatively the same as the one the consumer trades with θ^0 on the candidate equilibrium path, and makes a profit when traded by type θ^0 only. The second contract, targeted at type θ^1, allows the consumer to purchase the second layer $Q_{\theta^1} - Q_{\theta^0}$ at a unit price slightly less than v_{θ^1}, and makes a small loss when traded by type θ^1. Because the seller now offers the second layer at slightly better terms than his competitors, it is optimal for θ^1 to trade it with him on top of the first layer Q_{θ^0} provided by the other competitors at unit price v. By deviating in this way, the seller almost neutralizes his loss with θ^1, while securing a profit with θ^0. This amounts to dumping bad risks on one's competitors by selling

[23] Clearly, this quantity is strictly positive only if $\tau_{\theta^1}(Q_{\theta^0}, T_{\theta^0}) > v_{\theta^1}$.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

complementary coverage to type θ^1 slightly below the fair premium rate, and basic coverage to type θ^0 significantly above the fair premium rate.

It remains to be shown that the AMS aggregate allocation can be supported at equilibrium. This is done in the following:

Lemma 3 *If $\tau_{\theta^0}(0,0) \leq v$, then the AMS aggregate allocation can be supported in a pure strategy equilibrium of the non-exclusive menu game.*

An intuition for the proof can be provided along the following lines.[24] Consider the following profile of menus: each seller stands ready to provide any amount Q between 0 and an appropriately chosen upper bound $\overline{Q}$ at a unit price v_{θ^1}.[25] Clearly, trading $(Q_{\theta^0}, T_{\theta^0}) = (0,0)$ and $(Q_{\theta^1}, T_{\theta^1}) = (L, v_{\theta^1}L)$ is the unique optimal choice for type θ^0 and θ^1, respectively, as shown in Figure 15.4. Now consider sellers' deviations. Since $\tau_{\theta^0}(0,0) \leq v$, no seller can profitably deviate by attracting both types of the consumer. In addition, since type θ^1 gets her first-best level of insurance, the only deviations to be considered are those that cream-skim type θ^0. Specifically, we say that any insurance contract $(q,t) \in CS \equiv \{(q,t) \in \mathbb{R}^2_+ : \frac{t}{q} > v_{\theta^0} \;\; and \;\; U_{\theta^0}(q,t) > U_{\theta^0}(0,0)\}$ constitutes a cream-skimming deviation. Given any such deviation, one can see that type θ^0 finds optimal to purchase the corresponding contract (q,t). Yet, given equilibrium menus, type θ^1 also finds it optimal to purchase the same contract because she can complement it with some insurance issued by non-deviating intermediaries. Indeed, as depicted in Figure 15.4, starting from (q,t) type θ^1 can buy additional insurance

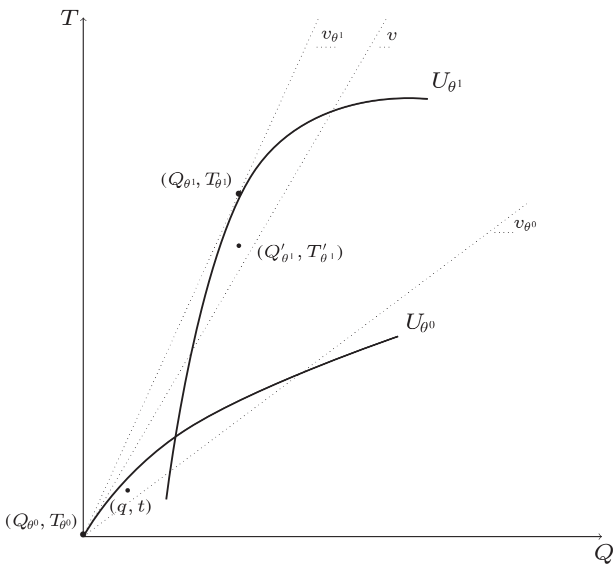


Figure 15.4 Equilibrium under non-exclusivity

[24] See Attar et al. (2014) for a general analysis.

[25] The upper bound is only introduced to make sure that the corresponding menus of contracts are compact, which enables reliance on PBE as a solution concept.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

at price v_{θ^1} so to achieve the allocation $(Q'_{\theta^1}, T'_{\theta^1})$. By doing that, she gets the full insurance quantity $Q'_{\theta^1} = Q_{\theta^1}$ at a smaller unit price $T'_{\theta^1} < T_{\theta^1}$. Overall, the deviation is traded by both types, and it is therefore non-profitable. Key to this reasoning is the possibility for type θ^1 to complement, at the deviation stage, any cream-skimming proposal with further trades provided by incumbent intermediaries. These additional opportunities for insurance, *i.e.* the availability of all quantities between 0 and $Q_{\theta^1} = L$, are usually denoted latent contracts. Despite not being traded at equilibrium, they should be issued to prevent some well-chosen deviations, and to guarantee existence of equilibrium.

Decentralization with a single agent: a discussion We have shown in the previous paragraphs that, in standard single agent contexts, the possibility to enforce exclusive contracting has dramatic implications on equilibrium outcomes. We now evaluate the normative implications of this insight.

Recall first that incentive compatibility is the relevant notion of feasibility when the planner fully observes agents' trades. In such a benchmark situation, the planner is allowed to design incentive-compatible mechanisms while perfectly observing, and therefore being able to monitor, aggregate trades (Myerson, 1979, 1982). As documented in Section 2.2.3, restricting attention to the set of budget-balanced and incentive-compatible trading mechanisms, several works have provided a characterization of the second-best efficiency frontier for insurance economies (see Prescott and Townsend, 1984 and Crocker and Snow, 1985). The corresponding allocations are regarded as a reference point from which to evaluate the performances of insurance markets in which intermediaries are able to enforce exclusivity of contracts. Indeed, as first shown by Crocker and Snow (1985), one can identify a set of conditions on agents' preferences guaranteeing that the *RSW* allocation belongs to the second-best frontier.[26] Importantly, these conditions are necessary and sufficient for the existence of a pure strategy equilibrium in the Rothschild and Stiglitz (1976) economy. This in turn provides an instance of the first theorem of welfare economics under exclusive competition: any allocation supported in a pure strategy equilibrium is constrained (*second-best*) efficient.

Under non-exclusive competition, however, no outside party can monitor the trades between the consumer and any subset of sellers. In general terms, little is known about how the opportunity for privately informed consumers to secretly sign bilateral agreements with sellers further restricts the set of allocations that are feasible to a planner. The recent work of Attar, Mariotti, and Salanié (2016) provides a first step in this direction. Specifically, they require feasible allocations to be not only incentive compatible, but also robust to further trading opportunities provided by private sellers. That is, any price-quantity scheme, or tariff, posted by the planner must be entry proof: no matter the offers subsequently made by an entrant, there is an optimal way for the buyer to combine these offers with the planner's tariff that prevents the entrant from making a profit.

To resume the findings in Attar et al. (2016), it is useful to refer to an allocation, first identified by Jaynes (1978), Hellwig (1988) and Glosten (1994), which we therefore denote *JHG*. In this allocation, both types purchase the same basic coverage, which type θ^1 complements by purchasing additional coverage. A marginal version of Akerlof (1970)

[26] See Bisin and Gottardi (2006) for a recent reformulation of these conditions.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

pricing holds: each layer of coverage is fairly priced given the types who purchase it, and the size of each layer is optimally chosen subject to this constraint. Thus, the first layer Q_{θ^1} is optimal for type θ^0 at unit price v,

$$Q_{\theta^0} \equiv \arg\max \{U_{\theta^0}(Q, vQ) : Q \geq 0\}, \tag{15.4}$$

$$T_{\theta^0} \equiv vQ_{\theta^0}. \tag{15.5}$$

Then the second layer $Q_{\theta^1} - Q_{\theta^0}$ is optimal for type θ^1 at unit price v_{θ^1}, given that she already purchases the first layer Q_{θ^0} at unit price v,

$$Q_{\theta^1} - Q_{\theta^0} \equiv \arg\max \{U_{\theta^1}(Q_{\theta^0} + Q, T_{\theta^0} + v_{\theta^1}Q) : Q \geq 0\}, \tag{15.6}$$

$$T_{\theta^1} - T_{\theta^0} \equiv v_{\theta^1}(Q_{\theta^1} - Q_{\theta^0}). \tag{15.7}$$

The *JHG* allocation is depicted in Figure 15.5.

Clearly, the *JHG* allocation $((Q_{\theta^0}, T_{\theta^0}), (Q_{\theta^1}, T_{\theta^1}))$ makes zero expected profit. However, because the coverage Q_{θ^1} is sold at the average premium rate $v > v_{\theta^0}$, type θ^0 subsidizes type θ^1. This allocation plays a key role in the set of incentive-compatible allocations. Specifically, Theorem 1 in Attar et al. (2016) establishes the following result:

Lemma 4 *The JHG allocation is the unique budget-balanced allocation implementable by an entry-proof tariff.*

That is, the threat of entry severely limits the scope for redistribution: the planner is constrained by his inability to control the buyer's trades with a potential entrant, as the threat of such trades effectively deprives him of any possibility to transfer utility between the

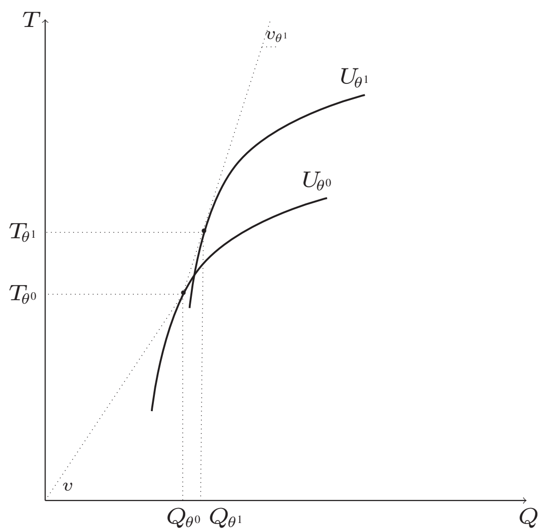


Figure 15.5 The JHG allocation

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

two types. The set of feasible allocations is a singleton. This contrasts with the multiplicity of second-best allocations, which, as discussed earlier, form a non-degenerate frontier. The intuition for the proof of Lemma 4 is as follows. First, to prevent entry, one should have $T_{\theta^0} \leq vQ_{\theta^0}$ and $T_{\theta^1} \leq T_{\theta^0} + v_{\theta^1}(Q_{\theta^1} - Q_{\theta^0})$. Indeed, violating the first inequality would make it profitable for an entrant to profitably attract both types on a contract of unit price above v. Violating the second one would make it profitable for an entrant to profitably attract type θ^1 on a contract of unit price above v_{θ^1}, which she might combine with $(Q_{\theta^0}, T_{\theta^0})$. Second, observe that, given that $T_{\theta^0} - vQ_{\theta^0} + \mu\left[T_{\theta^1} - T_{\theta^0} - v_{\theta^1}(Q_{\theta^1} - Q_{\theta^0})\right]$ by budget balancing, the two inequalities above can be satisfied together only when they hold as equalities. This shows that the *JHG* allocation is the only candidate to be implemented by an entry-proof tariff. In a next step, Attar et al. (2016) prove existence of such a tariff by considering the convex price–quantity schedule

$$T(q) \equiv 1_{\{q \leq Q_{\theta^0}\}} vq + 1_{\{q > Q_{\theta^0}\}} \left[vQ_{\theta^0} + v_{\theta^1}(q - Q_{\theta^0})\right],$$

which is the analogue in our two-type setting of the tariff constructed by Glosten (1994) when demand is continuously distributed.

It is important to clarify the relationship between this allocation and the second-best frontier for insurance economies analyzed by Crocker and Snow (1985) among others. In this respect, observe that, in the *JHG* allocation, only type θ^1 gets fully insured since $\tau_{\theta^1}(Q_{\theta^0}, T_{\theta^0}) > v_{\theta^1}$. In addition, the complementary coverage $Q_{\theta^1} - Q_{\theta^0}$ optimally traded by type θ^1 is strictly positive at the price v_{θ^1}. This in turn implies that her incentive compatibility constraint is slack: $U_{\theta^1}(Q_{\theta^1}, T_{\theta^1}) > U_{\theta^1}(Q_{\theta^0}, T_{\theta^0})$, which guarantees that the *JHG* allocation does *not* belong to the second-best frontier. To clarify this point, observe that, in a *JHG* allocation, one also has $\tau_{\theta^0}(Q_{\theta^0}, T_{\theta^0}) > v_{\theta^0}$, *i.e.* type θ^0 is underinsured. A planner with the ability to fully control trades can then complement the *JHG* allocation by proposing some additional coverage $(q_{\theta^1}, t_{\theta^1})$ at a premium rate $\frac{t_{\theta^1}}{q_{\theta^1}}$ between v_{θ^0} and $\tau_{\theta^0}(Q_{\theta^0}, T_{\theta^0})$, to be traded by type θ^0 only. As long as this additional amount of coverage is small enough, the relevant incentive constraint of type θ^1 would remain slack, inducing this type not to modify her behavior, and letting the planner achieve a positive (expected) budget surplus. This logic does not extend to the case in which the planner cannot perfectly control trades. In that case, any additional coverage $(q_{\theta^0}, t_{\theta^0})$, designed by the planner to attract type θ^0 alone, would be exploited by an entrant to propose further trades with type θ^1 at a premium rate slightly above v_{θ^1}. This would guarantee a profit to the entrant, and induce a deficit for the planner. Such a reasoning induces Attar et al. (2016) to conclude that, under non-exclusive competition, the relevant binding incentive constraint for type θ^1 is $U_{\theta^1}(Q_{\theta^1}, T_{\theta^1}) = \max\{U_{\theta^1}(Q_{\theta^0} + Q, T_{\theta^0} + v_{\theta^1}Q) : Q \geq 0\}$, which states that she is indifferent between trading $(Q_{\theta^1}, T_{\theta^1})$ and trading $(Q_{\theta^0}, T_{\theta^0})$ along with contracts issued by an entrant at the fair price v_{θ^1}.

To conclude, we remark that, whenever $\tau_{\theta^0}(0,0) \leq v$, (15.4) and (15.5) imply that type θ^0 purchases no insurance in a *JHG* allocation, that is, $(Q_{\theta^0}, T_{\theta^0}) = (0,0)$. It hence follows from Lemma 3 that, when the non-exclusive menu game has a pure strategy equilibrium, the corresponding allocation is entry proof. A planner who is in the impossible situation of controlling aggregate trades cannot therefore improve on such allocation without making profitable the entry of at least one seller. To the extent that this notion is interpreted as a form

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

of constrained efficiency,[27] the result suggests a sense in which market equilibria may achieve efficient outcomes under non-exclusive competition.

REFERENCES

Akerlof, G.A. (1970): "The Market for Lemons: Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics*, 84, 488–500.

Arrow, K. (1979): "The Property Rights Doctrine and Demand Revelation under Incomplete Information," in *Economics and Human Welfare: Essays in Honor of Tibor Scitovsky*, New York: Academic Press.

Attar, A., E. Campioni, and G. Piaser (2012): "Two-sided Communication in Competing Mechanism Games," *Journal of Mathematical Economics*, 49, 62–70.

Attar, A., E. Campioni, and G. Piaser (2016): "On Competing Mechanisms under Exclusive Competition," mimeo, Toulouse School of Economics.

Attar, A., T. Mariotti, and F. Salanié (2011): "Non Exclusive Competition in the Market for Lemons," *Econometrica*, 79, 1869–1918.

Attar, A., T. Mariotti, and F. Salanié (2014): "Nonexclusive Competition under Adverse Selection," *Theoretical Economics*, 9, 1–40.

Attar, A., T. Mariotti, and F. Salanié (2016): "Multiple Contracting in Insurance Markets," mimeo, Toulouse School of Economics.

Back, K. and S. Baruch (2013): "Strategic Liquidity Provision in Limit Order Markets," *Econometrica*, 81, 363–392.

Biais, B., D. Martimort, and J.-C. Rochet (2000): "Competing Mechanisms in a Common Value Environment," *Econometrica*, 78, 799–837.

Bisin, A. and P. Gottardi (2006): "Efficient Competitive Equilibria with Adverse Selection," *Journal of Political Economy*, 114, 485–516.

Burguet, R. and J. Sakovics (1999): "Imperfect Competition in Auction Designs," *International Economic Review*, 40, 231–247.

Carmona, G. and J. Fajardo (2009): "Existence of Equilibrium in Common Agency Games with Adverse Selection," *Games and Economic Behavior*, 66, 749–760.

Cho, I.-K. and D.M. Kreps (1987): "Signaling Games and Stable Equilibria," *The Quarterly Journal of Economics*, 102, 179–221.

Clarke, E. (1971): "Multipart Pricing of Public Goods," *Public Choice*, 11, 17–33.

Cramton, P., R. Gibbons, and P. Klemperer (1987): "Dissolving a Partnership Efficiently," *Econometrica*, 55, 615–632.

Crémer, J. and R.P. McLean (1985): "Optimal Selling Strategies Under Uncertainty for a Discriminating Monopolist When Demands Are Interdependent," *Econometrica*, 53, 345–361.

Crémer, J. and R.P. McLean (1988): "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions," *Econometrica*, 56, 1247–1257.

Crocker, K.J. and A. Snow (1985): "The Efficiency of Competitive Equilibria in Insurance Markets with Asymmetric Information," *Journal of Public Economics*, 26, 207–219.

d'Aspremont, C. and J. Crémer (2017): "Full Surplus Sharing With or Without Free Beliefs," CORE Discussion Paper (unpublished manuscript).

d'Aspremont, C. and L.-A. Gérard-Varet (1975): "Incentive Games with Incomplete Information: An Application to a Public Input Model," in *Colloque International de Théorie des Jeux*, Université Libre de Bruxelles.

d'Aspremont, C. and L.-A. Gérard-Varet (1979a): "Incentives and Incomplete Information," *Journal of Public Economics*, 11, 25–45.

d'Aspremont, C. and L.-A. Gérard-Varet (1979b): "On Bayesian Incentive Compatible Mechanisms," in J.-J. Laffont (ed.), *Aggregation and Revelation of Preferences*, Amsterdam: North Holland.

d'Aspremont, C. and L.-A. Gérard-Varet (1982): "Bayesian Incentive Compatible Beliefs," *Journal of Mathematical Economics*, 10, 83–103.

Epstein, L.G. and M. Peters (1999): "A Revelation Principle for Competing Mechanisms," *Journal of Economic Theory*, 88, 119–160.

Fagart, M.-C. (1996): "Concurrence en Contrats, Anti-sélection et Structure d'Information," *Annals of Economics and Statistics*, 1–27.

[27] Under complete information, several works have identified the set of *third-best* efficient allocations with those allocations implementable by an entry-proof tariff by the planner (see Kahn and Mookherjee, 1998 and **?**, **?**).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Finkelstein, A. and J. Poterba (2002): "Selection Effects in the United Kingdom Individual Annuities Market," *Economic Journal*, 112, 28–50.

Finkelstein, A. and J. Poterba (2004): "Adverse Selection in Insurance Markets: Policyholder Evidence from the U.K. Annuity Market," *Journal of Political Economy*, 112, 183–208.

Forges, F. (1994): "Posterior Efficiency," *Games and Economic Behavior*, 6, 238–261.

Glosten, L.R. (1994): "Is the Electronic Open Limit Order Book Inevitable?" *Journal of Finance*, 49, 1127–1161.

Griesmer, J., R.E. Levitan and M. Shubik (1967): "Toward a Study of Bidding Processes, Part IV – Games with Unknown Costs," *Naval Research Logistics Quarterly*, 14, 415–434.

Groves, T. (1973): "Incentives in Teams," *Econometrica*, 41, 617–631.

Han, S. (2015): "Robust Competitive Auctions," *Economics Letters*, 137, 207–210.

Harris, M. and R.M. Townsend (1981): "Resource Allocation Under Asymmetric Information," *Econometrica*, 49, 33–64.

Harsanyi, J.C. (1967): "Games with Incomplete Information Played by Bayesian Players. Part I. The Basic Model," *Management Science*, 14, 159–182.

Harsanyi, J.C. (1968): "Games with Incomplete Information Played by Bayesian Players, Part II. Bayesian Equilibrium Points," *Management Science*, 14, 320–334.

Heifetz, A. and Z. Neeman (2006): "On the Generic (Im)Possibility of Full Surplus Extraction in Mechanism Design," *Econometrica*, 74, 213–233.

Hellwig, M.F. (1988): "A Note on the Specification of Interfirm Communication in Insurance Markets with Adverse Selection," *Journal of Economic Theory*, 46, 154–163.

Holmstrom, B. and R.B. Myerson (1983): "Efficient and Durable Decision Rules with Incomplete Information," *Econometrica*, 51, 1799–1819.

Hurwicz, L. (1973): "The Design of Mechanisms for Resource Allocation," *American Economic Review*, 63, 1–30.

Jaynes, G.D. (1978): "Equilibria in Monopolistically Competitive Insurance Markets," *Journal of Economic Theory*, 19, 394–422.

Johnson, S., J. Pratt, and R. Zeckhauser (1990): "Efficiency Despite Mutually Payoff Relevant Private Information: The Private Value Case," *Econometrica*, 58, 873–900.

Kahn, C.M. and D. Mookherjee (1998): "Competition and Incentives with Nonexclusive Contracts," *RAND Journal of Economics*, 29, 443–465.

Kohlberg, E. and J.-F. Mertens (1986): "On the Strategic Stability of Equilibria," *Econometrica*, 54, 1003–1037.

Kosenok, G. and S. Severinov (2008): "Individually Rational, Budget-balanced Mechanisms and Allocation of Surplus," *Journal of Economic Theory*, 140, 126–161.

Kreps, D.M. and R. Wilson (1982): "Sequential Equilibria," *Econometrica*, 50, 863–894.

Krishna, V. (2010): *Auction Theory*, San Diego, CA: Academic Press.

Krishna, V. and M. Perry (1997): "Efficient Mechanism Design," *Game Theory and Information*, EconWPA.

Luz, V.F. (2016): "Characterization and Uniqueness of Equilibrium in Competitive Insurance," mimeo, University of British Columbia.

Makowski, L. and C. Mezzetti (1994): "Bayesian and Weakly Robust First Best Mechanisms: Characterizations," *Journal of Economic Theory*, 64, 500–519.

Martimort, D. and L.A. Stole (2002): "The Revelation and Delegation Principles in Common Agency Games," *Econometrica*, 70, 1659–1673.

Mas-Colell, A., M. Whinston, and J. Green (1995): *Microeconomic Theory*, Oxford: Oxford University Press.

Matsushima, H. (2007): "Mechanism Design with Side Payments: Individual Rationality and Iterative Dominance," *Journal of Economic Theory*, 133, 1–30.

McAfee, P. (1993): "Mechanism Design by Competing Sellers," *Econometrica*, 61, 1281–1312.

McAfee, R.P. and P.J. Reny (1992): "Correlated Information and Mechanism Design," *Econometrica*, 60, 395–421.

Mirrlees, J. (1971): "An Exploration in the Theory of Optimum Income Taxation," *Review of Economics Studies*, 38, 175–208.

Miyazaki, H. (1977): "The Rat Race and Internal Labor Markets," *Bell Journal of Economics*, 8, 394–418.

Myerson, R.B. (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6, 58–73.

Myerson, R.B. (1982): "Optimal Coordination Mechanisms in generalized Principal–Agent problems," *Journal of Mathematical Economics*, 10, 67–81.

Myerson, R.B. (1985): "Analysis of Two Bargaining Problems with Incomplete Information," in A. Roth (ed.), *Game Theoretic Models of Bargaining*, Cambridge, UK: Cambridge University Press, 59–69.

Myerson, R.B. and M.A. Satterthwaite (1983): "Efficient Mechanisms for Bilateral Trading," *Journal of Economic Theory*, 29, 265–281.

Page, F.H. and P.K. Monteiro (2003): "Three Principles of Competitive Nonlinear Pricing," *Journal of Mathematical Economics*, 39, 63–109.

Peck, J. (1997): "A Note on Competing Mechanisms and the Revelation Principle," mimeo, Ohio State University.

Peck, J. (2015): "Sky-lift Pricing with Imperfect Competition: An Exercise in Competing Mechanisms," mimeo, Ohio State University.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Peters, M. (1997): "A Competitive Distribution of Auctions," *Review of Economic Studies*, 64, 97–123.
Peters, M. (2001): "Common Agency and the Revelation Principle," *Econometrica*, 69, 1349–1372.
Peters, M. (2014): "Survey on Competing Mechanisms 2: The Revelation Principle," *Canadian Journal of Economics*, 47(2), 373–397.
Peters, M. and S. Severinov (1997): "Competition Among Sellers Who Offer Auctions Instead of Prices," *Journal of Economic Theory*, 75, 141–179.
Peters, M. and B. Szentes (2012): "Definable and Contractible Contracts," *Econometrica*, 80, 363–411.
Peters, M. and C. Troncoso-Valverde (2013): "A Folk Theorem for Competing Mechanisms," *Journal of Economic Theory*, 147, 953–973.
Philipson, T. and J. Cawley (1999): "An Empirical Examination of Information Barriers to Trade in Insurance," *American Economic Review*, 89, 827–846.
Prescott, E. and R. Townsend (1984): "Pareto Optima and Competitive Equilibria with Adverse Selection and Moral Hazard," *Econometrica*, 52, 21–45.
Riley, J.G. (2001): "Silver Signals: Twenty-five Years of Screening and Signaling," *Journal of Economic Literature*, 39, 432–478.
Riley, J.G. and W.F. Samuelson (1981): "Optimal Auctions," *American Economic Review*, 71, 381–392.
Rothschild, M. and J. Stiglitz (1976): "Equilibrium in Competitive Insurance Markets," *Quarterly Journal of Economics*, 90, 629–649.
Rustichini, A. and P. Siconolfi (2008): "General equilibrium in economies with adverse selection," *Economic Theory*, 37, 1–29.
Rysman, M. (2007): "An Empirical Analysis of Payment Card Usage," *Journal of Industrial Economics*, 55, 1–36.
Samuelson, P.A. (1969): "Pure theory of public expenditure and taxation," in J. Margolis and H. Guitton (eds) *Public Economics*, London: Macmillan.
Samuelson, W. (1984): "Bargaining Under Asymmetric Information," *Econometrica*, 52, 995–1005.
Schlee, E. and H. Chade (2012): "Optimal Insurance with Adverse Selection," *Theoretical Economics*, 7, 571–607.
Spence, M. (1973): "Job Market Signaling," *The Quarterly Journal of Economics*, 87, 355–374.
Spence, A. (1978): "Product Differentiation and Performance in Insurance Markets," *Journal of Public Economics*, 10, 427–447.
Stiglitz, J.E. (1977): "Monopoly, Non-linear Pricing and Imperfect Information: The Insurance Market," *Review of Economic Studies*, 44, 407–430.
Szentes, B. (2009): "A Note on 'Mechanism Games with Multiple Principals and Three or More Agents' by T. Yamashita," mimeo, UCL London.
Szentes, B. (2015): "Contractible Contracts in Common Agency Problems," *Review of Economic Studies*, 82, 391–422.
Vickrey, W. (1961): "Counterspeculation, Auctions, and Competitive Sealed Tenders," *Journal of Finance*, 16, 8–37.
Viràg, G. (2010): "Competing Auctions: Finite Markets and Convergence," *Theoretical Economics*, 5, 241–274.
Wilson, C. (1977): "A Model of Insurance Markets with Incomplete Information," *Journal of Economic Theory*, 16, 167–207.
Wilson, R.B. (1967): "Competitive Bidding with Asymmetric Information," *Management Science*, 13, 816–820.
Wilson, R.B. (1985): "Efficient Trading," in G. Feywel (ed.) *Issues of Contemporary Microeconomics and Welfare*, Albany, NY: State University of New York Press, 856–893.
Xiong, S. (2013): "A Folk Theorem for Contract Games with Multiple Principals and Agents," mimeo, Rice University.
Yamashita, T. (2010): "Mechanism Games with Multiple Principals and Three or More Agents," *Econometrica*, 78, 791–801.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

16. Moral hazard: Base models and two extensions

*Inés Macho-Stadler and David Pérez-Castrillo**

1 INTRODUCTION

Moral hazard (also called hidden action), the informational asymmetry related to the agent's behavior during a relationship, has been a long-time concern for insurance. It is said that the term moral hazard was coined in the nineteenth century by fire insurers to differentiate among the various hazards that cause a fire: physical hazards, both the ones related to the causes (lightning, short circuits) and the ones affecting the magnitude of the loss (type of construction), and moral hazards associated with insurees' behavior (less precautions or careless behavior).[1] Since Arrow (1963, 1968, 1971) and Pauly (1968), models of moral hazard and its applications have increasingly been recognized as key elements in understanding sharecropping contracts, corporate governance, licensing agreements, and executive compensations, to cite just a few examples.[2] Moral hazard models are now taught in many undergraduate majors and most graduate programs.[3]

In this chapter we review the literature on moral hazard in static environments. In its simplest version, a moral hazard problem is presented in the contractual relationship of a principal (she) and an agent (he) that works for her on a project.[4] The effort of the agent determines the probability distribution of the project's outcome.[5] There is a moral hazard problem when it is not possible to verify the agent's effort. This implies that effort cannot be contracted upon, because in the case of breach of contract, no court of law could know if the contract had really been breached or not. In this case, once hired, the agent will decide the level of effort that he prefers, taking into account how payments change with the outcome, that is, given the payment scheme that he has accepted. The payment scheme is the indirect way in which the principal can sway the agent's behavior.

The shape of the optimal payment scheme comes from the maximization of the principal's benefit subject to two constraints: the agent participation constraint (the agent will only sign the contract if by doing so he obtains at least as much as his opportunities outside this

* We thank Kaniska Dam, Mikhail Drugov, Pau Olivella, Joaquin Poblete, Pedro Rey-Biel, and Nicolas Roux for very useful comments. We also thank the Department FAE2 of the Universidad del Pais Vasco where this project was partially developed. We gratefully acknowledge financial support from the research grants (ECO2012-31962 and ECO2015-63679-P), Generalitat de Catalunya (2014SGR-142), ICREA Academia, and Severo Ochoa Programme (SEV-2015-0563). The authors are fellows of CESifo and MOVE.

1 Aetna Insurance Co. (1867).

2 The first efforts toward understanding and solving the principal–agent problem were due to Zeckhauser (1970), Spence and Zeckhauser (1971), Ross (1973), Stiglitz (1974), Mirrlees (1975, 1999), Harris and Raviv (1979), Holmström (1979), and Shavell (1979a, 1979b).

3 Several textbooks cover moral hazard problems along with adverse selection situations (Mas-Colell, Whinston, and Green, 1985, Macho-Stadler and Pérez-Castrillo, 1997, Salanié, 1997, Laffont and Martimort, 2002, and Bolton and Dewatripont, 2005).

4 The participants can be individuals or institutions. Examples are bank regulator and bank, shareholders and manager, and insurer and insuree.

5 In this chapter we will refer to the agent's effort, but the agent may be taking a decision or an action.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

relationship), and the incentive compatibility constraint (that recognizes that the agent will choose the effort that is best for him given the contract).

The general moral hazard problem is not easy to analyze. However, some simple set-ups have been very successful when adopted to study particular situations. First, it is generally assumed that the principal is risk neutral, and the agent's utility is separable in payment and effort. Still, these hypotheses do not allow us to have a simple enough problem. Second, it is often assumed that the "first-order approach" (FOA) is valid or that the agent chooses among a finite number (usually two) of possible efforts. While several interesting properties of the optimal contract can be derived thanks to these hypotheses, they do not allow the general derivation of explicit solutions. Thus, in many extensions and applications, further simplifications are used in order to find specific solutions. We will describe and bring into play two of these specifications that consider particular functional forms for the agent's utility function combined with certain assumptions on the payment scheme: the case of constant absolute risk aversion (CARA) utility function with linear contracts and the case of risk neutrality with limited liability.

The purpose of this chapter is neither to explain every aspect of the moral hazard problem nor to review each extension or topic. Moreover, for the sake of space, we focus on theoretical models, and we do not cover empirical or experimental results. We have chosen to present the main trade-offs of the principal–agent model and to discuss two extensions that we find particularly interesting: including behavioral considerations and an analysis of the market assignment that determines the partnerships that are formed. The first extension aims to discuss how incorporating behavioral biases in the analysis of incentives may affect the predictions of the classical moral hazard model. We discuss the effect of some of the strands of the literature. We start by considering an agent who not only takes into account his own well-being but also has other-regarding preferences. We then discuss the role of extrinsic and intrinsic motivations and the consequences on the optimal contract. We also cover the literature that concentrates on loss aversion, where the agent evaluates his payoffs not in absolute terms but in comparison with some reference. Finally, we consider the papers that focus on the idea that agents may be optimistic about the production process or overconfident about their ability.

The second extension we present relates to the insertion of the principal–agent problem in a matching market. It is easy to motivate this avenue from the point of view of the agency models. The partial equilibrium approach characterizes the optimal wage scheme when a principal hires an agent (a given pair principal–agent). This approach is well defined for the case where there is a single principal in the economy or when principals are perfectly competitive and hence get zero profits. In the classical approach, the bargaining power is given to principals or agents by assumption, which implies that the reservation utility or the zero profit condition determines the distribution of surplus. In other words, the effects of competition are summarized by a single parameter of the agent's outside option (his reservation utility) or the principal's zero profit condition. However, when we consider explicitly the existence of several heterogeneous principals and several heterogeneous agents, some of the properties obtained in the simple version of the agency problem do not necessarily hold. Thus, empirical work and policy recommendations may be based on the wrong arguments. Moreover, we can address the endogenous determination of the principals and agents that meet.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

2 BASE MORAL HAZARD MODELS

A principal hires an agent to perform a task that we refer to as effort, $e \in E$, in exchange for a wage, w. The final outcome of the relationship, x, depends on the effort e that the agent devotes to the task and some random variable for which both participants have the same prior distribution. The set of possible outcomes is denoted by X, which can be a continuous set, in which case we denote $X = [\underline{x}, \overline{x}]$, or a discrete set. The distribution of the random variable induces a probability $p(x \mid e) > 0$ of outcome $x \in X$ conditional on effort e, where $p(. \mid e)$ is a twice continuously differentiable density function if X is a continuous set and it is a vector of probabilities if X is discrete. We denote by $P(x \mid e)$ the cumulative distribution function, that is, $P(x \mid e) = \int_{y \leq x} p(y \mid e)$ (or $P(x \mid e) = \sum_{y \leq x} p(y \mid e)$ if X is discrete).

Since uncertainty exists, participants may react to risk. We concentrate on the case of a risk-neutral principal and a (possibly) risk-averse agent. Risk preferences are expressed by the shape of their (von Neumann-Morgenstern) utility functions. The principal, who owns the outcome x and must pay the agent w, has preferences represented by the utility function $x - w$. The agent, who receives a monetary payoff w for his participation in the relationship and supplies an effort e, has an additively separable utility function: $U(w, e) = u(w) - v(e)$, where $u(w)$ is assumed to be increasing and concave and $v(e)$ is increasing and convex.[6] The agent can obtain a utility level $\underline{U}$ outside the relationship with the principal. Therefore, he only accepts contracts that give him, in expectation, at least $\underline{U}$.

Under symmetric information, that is, when effort is verifiable, the optimal (first-best) contract includes the first-best effort e^{FB} and the payment scheme $\left(w^{FB}(x)\right)_{x \in X}$, which incorporates the optimal risk sharing among the two participants. If the agent is risk averse then the payment mechanism completely insures the agent: he receives a fixed payment.

Under moral hazard, the effort is not contractible and the agent can choose the effort that is best for him, given the contract. Thus, if the principal proposes a fixed wage, the agent's payment does not depend on his effort and he will choose the effort that is least costly for him, that is, the lowest possible level of effort.

When designing the optimal incentive contract for the moral hazard problem, the principal gets the agent interested in the consequences of his own behavior by making his payoff dependent on the outcome. If the agent is risk averse, given that the outcome is noisy, this entails the cost of distorting the optimal risk sharing among both participants. In this case, the optimal contract solves the trade-off between distorting the efficient allocation of risk and providing incentives.

The optimal contract under moral hazard takes into account the acceptance condition for the agent and his choice of effort. Moreover, it is often the case that arbitrarily low or high payments are not feasible, which would introduce additional constraints into the principal's program. For example, the agent may have limited liability so that it is not possible to impose a penalty on him (or he should receive a minimum legal wage independent of the outcome). Similarly, it may not be possible for the principal to pay the agent more than the value of the outcome, or she may be constrained (by law or by norm) not to pay too much to the

[6] The key characteristic of this class of utility functions is that the agent's risk aversion (preferences over lotteries) is independent of the effort supplied. Grossman and Hart (1983) assume the most general utility function by considering that the agent's utility has the form $U(w, e) = K(e)u(w) - v(e)$. Special cases are $K(e) = 1$, i.e., additively separable preferences, and $v(e) = 0$, i.e., multiplicatively separable preferences.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

agent. An example of upper bounds are the European Union regulatory cap on bankers' bonus payments such that "the maximum ratio between the variable and the fixed part of the total remuneration is limited to 100%." When there are lower and/or upper bounds to the agent's payment, new trade-offs may appear. For example, the implementation of some (high) efforts may not be possible because there is no room for enough variation in payments or it may become very expensive as it requires awarding the agent an expected utility that is higher than his reservation utility. In the latter case, there is a trade-off between giving extra rents and providing incentives.

The timing of the relationship between the principal and the agent is the following. First, the principal decides on the contract she offers to the agent, in particular on the agent's payment scheme $(w(x))_{x\in X}$ as a function of the outcome of the relationship. Then the agent decides whether or not to enter the relationship. Finally, if the contract is accepted, the agent chooses the effort level e that he most desires, given the agreed contract. This is a free decision by the agent because effort is not a verifiable variable. The principal bears this fact in mind when she designs the contract that defines the relationship, and the "game" can be solved by backward induction. Formally, if we first consider a situation without lower or upper bounds on salaries, the optimal contract under moral hazard is the solution to the maximization problem ($P1$):

$$\underset{\{(w(x))_{x\in X},e\}}{Max} \{E(x - w(x) \mid e)\}$$

$$\text{s.t. } E(u(w(x)) \mid e) - v(e) \geq \underline{U} \tag{PC}$$

$$e \in \arg\max_{\widehat{e}\in E} \{E(u(w(x)) \mid \widehat{e}) - v(\widehat{e})\}, \tag{ICC}$$

where $E(y \mid e)$ denotes the expectation of y conditional on the effort e. The first restriction of the program is the *participation constraint* (PC), which states that the agent will not sign a contract that gives him lower expected utility than the alternative market opportunities. The second restriction is the *incentive compatibility constraint* (ICC), which determines the agent's effort under moral hazard. If the ICC is not relevant (either because there is symmetric information or because ICC is not binding at the optimum) then the solution to the program is the first-best contract $\left(\left(w^{FB}(x)\right)_{x\in X}, e^{FB}\right)$.

The solution to program ($P1$) provides the optimal contract under moral hazard and the optimal level of the principal's utility for a given level of the reservation utility $\underline{U}$. As the level of $\underline{U}$ changes, we obtain the Pareto frontier in the space of the utilities of the two participants. Thus, the main quality properties of the optimal contract hold if instead of considering (PC) we maximize the agent's utility subject to a participation constraint for the principal.

The main difficulty in solving the general program ($P1$) is related to the fact that the incentive compatibility constraint is itself a maximization problem. To overcome this obstacle, the literature has adopted two solutions. (a) If the set E is finite (most papers that follow this approach consider E to include two levels of effort) then the ICC can be replaced by a finite set of inequalities (just one inequality in the case of two efforts). (b) If the set E is a continuum, say $E = \left[\underline{e}, \overline{e}\right]$, then we can try to substitute the ICC by its first-order condition, which is a necessary condition of the optimal e if it is interior. This is called the *first-order approach* (FOA). One has to be careful if one follows this approach because the agent's expected utility may fail to be concave in effort. Hence, using the FOA may be incorrect, and finding the

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

optimal effort in this program difficult.[7] A possible way out proposed by Grossman and Hart (1983) is to solve the problem in steps, first identifying the optimal payment mechanism for any effort and then, if possible, finding the optimal effort.[8] The other possibility is to consider situations where the agent's maximization problem is well defined, which requires introducing assumptions for the FOA to be valid.

The moral hazard problem may give rise to several distortions in the optimal contract because it forces the principal to trade-off incentives for the effort of the agent and other objectives. We now discuss characteristics of the solution of ($P1$) for several cases, emphasizing the trade-offs faced by the principal. We will make it clear in some cases that additional constraints are added to ($P1$) owing, for example, to the existence of bounds on the payments.

2.1 Incentives vs Risk Sharing

We first analyze the consequences of moral hazard in situations where the agent is risk averse, that is, $u(w)$ is strictly concave. In this case, the optimal, first-best contract fully insures the agent. However, providing incentives requires that the agent's salary depends on the outcome. Thus, the principal needs to trade-off incentives vs risk-sharing.

We develop the analysis for three different models.

2.1.1 Model 1: Continuous effort

Consider a situation where E is continuous and the FOA is valid. Denote by λ (resp., μ) the Lagrangian multiplier of the PC (resp., the ICC). Then, for a given effort e, Holmström (1979) tells us that the solution to the principal's program ($P1$) with respect to the payoff scheme $(w(x))_{x \in X}$ satisfies, for all $x \in X$,

$$\frac{1}{u'(w^*(x))} = \lambda + \mu \frac{p_e(x \mid e)}{p(x \mid e)}, \tag{16.1}$$

where $p_e(x \mid e)$ is the partial derivative of $p(x \mid e)$ with respect to e. In the optimal contract, both PC and ICC are binding, that is, λ and μ are strictly positive. Their value depends on the effort e.[9] The ratio $\frac{p_e(x|e)}{p(x|e)}$ is the *likelihood ratio* of obtaining outcome x when the effort is e.

The optimal scheme $(w^*(x))_{x \in X}$ and the multipliers λ and μ are characterized by the condition (16.1) for all $x \in X$ together with (PC) and (ICC). Therefore, the optimal wage scheme for a given effort e does not depend on the value that the principal places on the outcome: the value of x does not enter directly into any of these equations. If the wage is a

[7] Mirrlees (1975) shows that the FOA is generally invalid unless the optimum effort derived from the ICC (the solution to the agent's maximum problem) is unique. In the absence of uniqueness, the first-order conditions of the principal's problem when the ICC is substituted by its first-order condition are not even necessary conditions for the optimality of the incentive contract. We describe the conditions at the end of section 2.1.

[8] Grossman and Hart (1983) show that this can always be done for additively or multiplicatively separable utility functions. By using the utilities of the wages instead of the wages, the principal's program with respect to the payment scheme for any effort can be rewritten as a minimization problem where the objective function is a convex cost function subject to (a possibly infinite number of) linear constraints. In particular, when the set of possible efforts E is a finite set, using Kuhn-Tucker one obtains necessary and sufficient conditions for optimality.

[9] If the agent is risk neutral, then the multiplier μ is zero and equation (16.1) only gives the value of the multiplier λ. In this case, any payment scheme whose expected payoff ensures the agent an expected utility level of $\underline{U}$ is optimal.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

function of the outcome it serves only as an incentive for the agent. Hence, it only depends on the outcome as long as the outcome is informative about the effort. In particular, the necessary and sufficient condition for a better outcome to always lead to a higher wage, that is, $w^{*\prime}(x) > 0$, is that the likelihood ratio is increasing in x. This condition is called the *monotone likelihood ratio property* (MLRP), which holds when

$$\frac{p_e(x \mid e)}{p(x \mid e)} \text{ is strictly increasing in } x \tag{MLRP}$$

for all $e > \underline{e}$.[10]

Moreover, MLRP together with CDFC (*convexity of the distribution function condition*), which are often called the Mirrlees-Rogerson sufficient conditions, are sufficient conditions for the validity of the FOA (Mirrlees, 1976, Rogerson, 1985, and Jewitt, 1988). We say that a distribution function satisfies CDFC if the second derivative of the cumulative distribution function $P(x \mid e)$ with respect to e is non-negative, that is,

$$\frac{\partial^2 P(x \mid e)}{\partial e^2} \geq 0.$$

Hence, the validity of the FOA requires demanding conditions on the probability function (MLRP and CDFC).[11,12]

We make two additional remarks about the optimal contract. First, the wage scheme needs to be simpler as the agent has more room to manipulate the outcome. For example, if the agent can freely dispose of the output, the optimal payment mechanism is necessarily monotonic even if the MLRP does not hold. Alternatively, if there are several agents who can trade output among themselves, then only a linear scheme is feasible (any non-linear scheme will be "linearized" by arbitrage).[13]

Second, we have considered payment schemes that only depend on the outcome of the relationship (the outcome related to the effort is the only verifiable variable). However, the principal will base the contract on any signals that reveal information on the agent's effort. Hence, if possible, the contract should be contingent on many other variables. Information

[10] Holmström (1979), Shavell (1979a), and Milgrom (1981) show that under the FOA, if the distribution function satisfies MLRP then the wage scheme is increasing in output. Note that MLRP is stronger than first-order stochastic dominance, which requires that $\frac{\partial}{\partial e}P(x \mid e) < 0$ for all $x \in \left(\underline{x}, \overline{x}\right)$.

[11] MLRP and CDFC are very strong conditions and it is difficult to find distributions that satisfy both of them. The two-step procedure proposed by Grossman and Hart (1983) provides a way of proceeding when the FOA is not valid.

[12] Kirkegaard (2014) recently proposed a reformulation of the moral hazard problem that allows the use of results from the areas of choice under uncertainty. In this way, he can prove the classic results using an unifying methodology and also extend the analysis to larger domains than previous work.

[13] By considering additional properties of the participants' objective function, more information on the optimal contract can be obtained. Imagine that the agent is "prudent," in the sense that $u'' < 0$ and $u''' > 0$. A prudent agent is risk averse and his marginal utility is strictly convex so he is downside risk averse (Menezes, Geiss and Tressler, 1980). This agent applies a heavier discount to downward variations than to upward variations of the payment scheme. Chaigneau (2014) shows that concave contracts tend to provide more incentives to risk-averse agents, while convex contracts tend to be more profitable to motivate prudent ones. The intuition is that concave payment schemes concentrate incentives where the marginal utility of risk-averse agents is highest, while convex contracts protect against downside risk. However, when the principal is also risk averse and prudent, convex contracts are not optimal if the principal is sufficiently prudent relative to the agent (Sinclair-Desgagné and Spaeter, 2013).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

related to the state of nature may be useful if it allows better estimates of the agent's effort, thus reducing the risk inherent in the relationship. This is known as *the sufficient statistic result*, and it is perhaps the most important conclusion in the moral hazard literature (Holmström, 1979). Formally, we say that x is sufficient for $\{x, y\}$ with respect to $e \in E$ if and only if the distribution function p is multiplicatively separable in y and e :

$$p(x, y \mid e) \equiv g(x, e)h(y, x).$$

We say that y is informative about $e \in E$ if x is not sufficient for $\{x, y\}$ with respect to $e \in E$. Finally, if y is informative about $e \in E$ then there is a payment mechanism $w(x, y)$ that strictly Pareto dominates the best $w(x)$.

The empirical content of the sufficient statistic argument is that the optimal contract should exploit all available information in order to optimally filter out risk.[14] In the limit, if by including many variables the agent's effort can be inferred with certainty, then the symmetric information effort can be implemented at no extra cost.

Finally, once we have computed the optimal scheme for each e, which we denote $w^*(x, e)_{x\in X}$, the principal can find the *optimal effort* under the moral hazard problem by solving

$$\max_{e\in E} \left\{E\left(x - w^*(x, e) \mid e\right)\right\}.$$

The main difficulty of this program is that it is not generally concave in effort. If the principal's problem is well defined and has a solution, the optimal effort e^* is determined by the usual condition of equality between marginal revenues and the marginal costs of increasing the effort, which includes the increase in average wages plus the extra cost in terms of the incentives needed to increase the effort.[15] We notice that some efforts may not be implementable under moral hazard and that the lowest effort $\underline{e}$ can always be implemented at no extra cost using the symmetric information wage scheme.

Example 1 CARA risk preferences and linear contract A particularly simple, and very popular, model is one where the principal is risk neutral and the agent has CARA risk preferences:

$$u(w, e) = -\exp\left[-r\left(w - v(e)\right)\right],$$

where r is the coefficient of absolute risk aversion. Additionally, assume that the cost of effort is a quadratic function

$$v(e) = \frac{1}{2}ve^2.$$

[14] For example, when the principal hires several agents, the central question is whether incentives should be provided as a function of all agents' performance. The answer comes from the sufficient statistic result and depends on the linkage of the agents' situation, in particular on whether the agents' outcomes are subject to correlated shocks (informational linkage) or whether the performance of an agent depends on the effort of other agents (technological linkage). See Holmström (1982) and Mookherjee (1984).

[15] It is interesting to note that under symmetric information the PC determines the optimal effort level, while it is the cost implied by the ICC that determines the effort when there is moral hazard. The reason is that under moral hazard and using the FOA the ICC implies that the derivative of the PC with respect to the effort is zero.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

The output x depends on the agent's effort e and a random variable ε that is normally distributed with mean zero and variance σ^2:

$$x = e + \varepsilon.$$

Finally, we restrict attention to linear wage schemes of the form $w = F + sx$, where F is a fixed payment and s is the share of the output that goes to the agent.[16] In this case, it is convenient to solve the program using the agent's certain equivalent income

$$F + se - \frac{1}{2}ve^2 - \frac{r}{2}\sigma^2,$$

in which case the ICC becomes very easy to write: $e = \frac{s}{v}$.

Solving the principal's program, the PC determines the fixed part of the contract F and the variable performance part of the contract is

$$s^* = \frac{1}{1 + rv\sigma^2}, \tag{16.2}$$

which is decreasing in the cost of the effort v, the agent's risk aversion (measured by r), and the variance of the outcome σ^2. Since a higher s translates into a higher effort, the previous expressions reflect the trade-off between efficiency (optimal risk sharing would require $s = 0$) and incentives.[17]

2.1.2 Model 2: Two efforts

Consider a situation similar to the one discussed in Model 1 but with $E = \{e^H, e^L\}$, that is, there are only two possible levels of effort: a high effort whose cost for the agent is $v\left(e^H\right)$ and a low effort with a cost of $v\left(e^L\right) < v\left(e^H\right)$.[18] Implementing e^L is easy because the same

---

[16] Although linear contracts are generally not optimal in the static setting (see Mirrlees, 1975), Holmström and Milgrom (1987) show that the optimal contract is linear in the final outcome if the agent chooses efforts continuously to control the drift vector of a Brownian motion process and he observes his accumulated performance before acting. Linear contracts are also shown to be optimal in models with limited liability and risk neutrality if the principal is uncertain about the technology available to the agent (see Carroll, 2015).

[17] In a multi-agent situation, the sufficient statistic result is easy to illustrate when the principal hires two agents with CARA risk preferences and non-cooperative behavior. Linear contracts would have the form

$$w_i = F_i + s_i x_i + z_i x_{-i} \qquad \text{for } i = 1, 2.$$

When $z_i \neq 0$ there is relative performance evaluation. Suppose that each agent's individual outcome depends on the other agent's random shock:

$$x_i = e_i + \varepsilon_i + \rho\varepsilon_{-i} \qquad \text{for } i = 1, 2,$$

where ε_i, $i = 1, 2$, follows a distribution $N(0, \sigma^2)$, and ρ is the degree of correlation among the agents' outcomes. Then, in the optimal contract,

$$s_i^* = \frac{1 + \rho^2}{1 + \rho^2 + rv\sigma^2\left(1 - \rho^2\right)^2}, \qquad z_i^* = -\frac{2\rho}{1 + \rho^2 + rv\sigma^2\left(1 - \rho^2\right)^2}.$$

Thus, for $\rho \neq 0$, there is relative performance evaluation, since the wage of agent i depends on the individual outcome of agent $-i$.

[18] See, e.g., Grossman and Hart (1983).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

fixed-wage contract that is optimal under symmetric information is also optimal under moral hazard (the Lagrange multiplier of the ICC is zero). On the other hand, implementing e^H requires taking into account the ICC that, in this case, can be written as

$$E\left(u\left(w\left(x\right)\right) \mid e^H\right) - v(e^H) \geq E\left(u\left(w\left(x\right)\right) \mid e^L\right) - v(e^L). \tag{ICC2}$$

The solution to $(P1)$ satisfies, for all $x \in X$,

$$\frac{1}{u'(w^*(x))} = \lambda + \mu \frac{\left(p(x \mid e^H) - p(x \mid e^L)\right)}{p(x \mid e^H)}, \tag{16.3}$$

where $\frac{(p(x|e^H)-p(x|e^L))}{p(x|e^H)}$ is the likelihood ratio in the discrete case.

In this model, once the optimal payment scheme that allows the implementation of each effort has been obtained, finding the optimal effort is straightforward. It comes from the comparison of the principal's profits for each effort.

2.1.3 Model 3: Bounded feasible payments

As discussed above, there are important real-life situations where the principal cannot base the incentives on arbitrarily large bonuses ("carrots") or fines ("sticks"). We consider now a situation that shares all the assumptions of Model 1 but where there are lower and upper bounds for the feasible payments. For each outcome $x \in X$, the salary $w\left(x\right)$ must satisfy

$$\underline{w}(x) \leq w(x) \leq \overline{w}(x) \tag{16.4}$$

where $\underline{w}(x)$ and $\overline{w}(x)$ are continuous, non-decreasing, and piecewise differentiable, with $\underline{w}(x) < \overline{w}(x)$ for all $x \in X$. Moreover, assume that the MLRP holds. Then, the analysis of Jewitt, Kadan, and Swinkels (2008) ensures that the optimal payment scheme $(w^*(x))_{x \in X}$ to implement an effort e satisfies conditions similar to (16.1) "as much as possible":

$$\frac{1}{u'(w^*(x))} = \begin{cases} \dfrac{1}{u'(\overline{w}(x))}, \text{ if } \frac{1}{u'(\overline{w}(x))} < \lambda + \mu \frac{p_e(x|e)}{p(x|e)} \\ \lambda + \mu \frac{p_e(x|e)}{p(x|e)}, \text{ if } \dfrac{1}{u'(\underline{w}(x))} \leq \lambda + \mu \dfrac{p_e(x \mid e)}{p(x \mid e)} \leq \dfrac{1}{u'(\overline{w}(x))} \\ \dfrac{1}{u'(\underline{w}(x))}, \text{ if } \lambda + \mu \frac{p_e(x|e)}{p(x|e)} < \frac{1}{u'(\underline{w}(x))} \end{cases}, \tag{16.5}$$

for some $\lambda \geq 0$ and $\mu \geq 0$. A particularly interesting example corresponds to a situation where there is no upper bound on salaries but there is a minimum wage $\underline{w}$ (that is, the lower bound is independent of the outcome). This may be the case, for example, because of the agent's limited liability. In this case, the first line of (16.5) has no bite and the third line of (16.5) matters for low levels of the outcome, because the MLRP implies that the wage scheme is monotone. Thus, the optimal contract offers the minimum salary wage $\underline{w}$ until some minimum outcome $\widehat{x}$ is reached and, from this level on, the contract follows a pattern similar to that without bounds.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

2.2 Incentives vs Rents

We now assume that both the principal and the agent are risk neutral and that the sets X and E are continuous. Moreover, the payments to the agent are subject to lower and upper bounds.

Without limited liability, and because of the agent's risk neutrality, there is no benefit in insuring the agent and the solution of $(P1)$ would be a franchise contract that would lead to the first-best. The franchise contract has the form $w(x) = w - k$, where k is the constant that makes (PC) binding. However, with limited liability, the principal is often forced to give the agent additional rents so that he has an incentive to provide a high effort.[19] Thus, the optimal contract trades off incentives vs rents.

2.2.1 Model 4: Limited liability

The participants are subject to limited liability so that, in the same spirit as in (16.4), the wage can neither be negative nor higher than the outcome, that is,

$$0 \leq w(x) \leq x \tag{16.6}$$

for all $x \in X$. Following the steps in Innes (1990),[20] we assume that the MLRP holds, $E\{x \mid e = 0\} = 0$, a profitable contract exists involving a positive effort, and the total value of the relationship is strictly concave in e.

Given that higher effort increases the probability of higher outcomes, the contract should give the agent maximal payoffs in high outcomes. A particularly simple contract emerges under the additional monotonicity constraint that the principal's profit cannot be decreasing in the outcome, that is, $x - w(x)$ is non-decreasing in x.[21] In this case, the optimal contract is a "debt contract" for the principal where she obtains $\min\{x, z\}$, for some $z > 0$. Thus, the optimal salary scheme is

$$w^*(x) = \max\{x - z, 0\}$$

for all $x \in X$. The value z corresponds to the one that makes (PC) binding if $\underline{U}$ is high enough.[22] If $\underline{U}$ is very low then (PC) is not binding because (16.6) constrains the payoffs so

[19] When the agent has limited wealth, his level of effort may be constrained. Quérou, Soubeyran, and Soubeyran (2015) study a situation where the principal may need to make an up-front transfer to the agent because the agent may not have enough resources to pay for the cost of the effort, when this cost is monetary.

[20] Holmström (1979) and Lewis (1980) already noted the potential importance of limited liability constraints. However, Innes (1990) is the first paper to study the impact of liability limits on the qualitative properties of the optimal contract. Sappington (1983) and Demski, Sappington, and Spiller (1988) also bring in a limited liability constraint but they assume that the agent chooses the effort e after observing the state of nature. Other papers that study moral hazard problems with a minimum bound on payments under different assumptions are Park (1995), Kim (1997), Oyer (2000), Matthews (2001), and Jewitt et al. (2008).

[21] The monotonicity constraint may be due to the possibility for the principal to "burn" or "hide" profits, or to the possibility for the agent to inflate the outcome at a cost if the payment does not satisfy the constraint.

[22] Matthews (2001) shows that, under the same restrictions as Innes (1990), debt is still the optimal incentive contract if the agent is risk averse, renegotiation cannot be prevented, and the agent has all the bargaining power in the renegotiation game.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

much that the principal prefers to give the agent some extra rent to better provide incentives for effort. In all the cases, the effort implemented under moral hazard is lower than the first-best level e^{FB}.[23]

Without the monotonicity constraint, the optimal contract takes the extreme “live-or-die” contract of the form

$$w^*(x) = \begin{cases} 0, \text{ if } x \leq z \\ x, \text{ if } x > z \end{cases}$$

for some value $z > 0$, whenever this contract leads to an effort level lower than e^{FB}. Otherwise, a contract proportional to the previous one (that is, a contract that gives sx for $x > z$) is a solution to the program (with z and s chosen in an appropriate way) and it implements the first best.

The PC is often not binding in situations where there is limited liability. In other words, when the limited liability constraint is binding the agent may obtain some rents, making the participation constraint slack. This is in contrast to the case without limited liability constraints where, at least when the agent’s utility is additively separable, the agent never receives rents. Thus, the non-verifiability of the effort when there is limited liability and the agent is risk neutral may imply a cost either because the optimal contract leads to an effort lower than e^{FB}, because it gives the agent an informational rent, or both.[24] Example 2 illustrates the trade-offs in a simple model with two outcomes.

Example 2 Limited liability and two outcomes In this example, we consider that only the agent is subject to limited liability, so the only additional constraint to program ($P1$) is $w(x) \geq 0$. Moreover, two outcomes are possible: success (a “good” outcome), in which case $x = x_G > 0$, and failure (a “bad” outcome), $x_B = 0$. The probability of success is $p(x_G \mid e) = e$ and the cost of effort is $v(e) = \frac{1}{2}ve^2$, with $v > 0$.

The agent’s ICC implies that, under a contract $(w(x_B), w(x_G))$, he will select effort $e = \frac{w(x_G)-w(x_B)}{v}$. Once we take this constraint into account, together with the participation and limited liability constraints, the program that the principal solves is

$$\underset{(w(x_B),w(x_G))}{Max} \left\{ -w(x_B) + \frac{(w(x_G) - w(x_B))(x_G - (w(x_G) - w(x_B)))}{v} \right\}$$

$$\text{s.t. } w(x_B) + \frac{(w(x_G) - w(x_B))^2}{2v} \geq \underline{U} \tag{16.7a}$$

$$w(x_B) \geq 0, \tag{16.7b}$$

which has a solution in which the principal makes non-negative profits if $\underline{U} \leq \frac{x_G^2}{2v}$.

[23] Poblete and Spulber (2012) extend the analysis of Innes (1990) by characterizing the optimal agency contract in more general environments using the state-space (or parametric) representation. They assume a technology $x = x(\theta, e)$, where θ is the state, a random variable with some density and distribution function. They do not assume the MLRP and introduce a “critical ratio” from which the form of the optimal contract easily follows. In particular, they provide a weaker condition than the MLRP under which the optimal contract is a debt contract.

[24] This framework is also useful for studying more complex situations. See Fleckinger and Roux (2012) for a comprehensive review of the literature on performance comparison and competition in motivating agents in the framework where all participants are risk neutral, the agents are protected by limited liability, and they choose their effort non-cooperatively.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

In the optimal contract, the base salary is zero, $w^*(x_B) = 0$, because the limited liability constraint (16.7b) always binds. The participation constraint (16.7a) is also binding if the agent's reservation utility $\underline{U}$ is intermediate ($\underline{U} \in \left[\frac{x_G^2}{8v}, \frac{x_G^2}{2v}\right]$). On the other hand, if $\underline{U}$ is low ($\underline{U} < \frac{x_G^2}{8v}$) then the optimal contract gives a rent to the agent: providing incentives requires separating $w(x_G)$ from $w(x_B) = 0$ and the principal prefers to offer a salary $w(x_G)$ higher than the one necessary to satisfy the PC so that the agent chooses a higher effort. The optimal bonus is

$$w^*(x_G) = \begin{cases} \sqrt{2v\underline{U}}, \text{ if } \dfrac{x_G^2}{8v} \leq \underline{U} \leq \dfrac{x_G^2}{2v} \\ \dfrac{x_G}{2}, \text{ if } \underline{U} < \dfrac{x_G^2}{8v}. \end{cases} \tag{16.8}$$

For intermediate outside utility, the effort increases (it gets closer to the first-best effort $e^{FB} = \frac{x_G}{2v}$) while the principal's profit decreases with $\underline{U}$: $e^* = \sqrt{\frac{2\underline{U}}{v}}$ and $\pi^* = x_G\sqrt{\frac{2\underline{U}}{v}} - 2\underline{U}$. When $\underline{U} < \frac{x_G^2}{8v}$ then effort and profits are constant: $e^* = \frac{x_G}{2v}$, the utility of the agent is $\frac{x_G^2}{8v} > \underline{U}$ and the principal profit is $\frac{x_G^2}{4v}$.

2.3 Incentives to a Task vs Incentives to Another Task

The basic theory of moral hazard considers an agent supplying a one-dimensional effort that influences a one-dimensional output. However, relationships are often more complicated and the agent may be responsible for supplying a multi-dimensional effort or performing more than one task. Holmström and Milgrom (1991) study this extension using a model where the agent has CARA utility over wage and effort (as in Example 1) and either the tasks are related in the agent's cost of exerting them or their outcomes may be subject to correlated shocks. In this influential paper they discuss, among other issues, the trade-off between the incentives for different tasks in the extreme case where the outcome is easy to measure in one task while it is very difficult or impossible to measure (or to verify) in another task.[25]

2.3.1 Model 5: Moral hazard with two tasks

Consider a risk-neutral principal who hires an agent with a CARA utility function to provide a vector (e_1, e_2) of efforts. The cost of the efforts for the agent are summarized in the cost function

$$v(e_1, e_2) = \frac{1}{2}v(e_1^2 + e_2^2) + \delta e_1 e_2,$$

with $|\delta| < v$. The output vector (x_1, x_2) depends on the agent's efforts and some random variable:

$$x_i = e_i + \varepsilon_i \text{ for } i = 1, 2.$$

[25] Holmström and Milgrom (1991) also consider limits on outside activities and how to allocate tasks between the agents. See also Feltham and Xie (1994).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

The noise of the output function is assumed to follow a normal distribution $N(0, \Sigma)$, with $\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$. Finally, the principal offers a payment scheme $w(x_1, x_2)$ to the agent, which is assumed to be linear:

$$w(x_1, x_2) = F + s_1 x_1 + s_2 x_2.$$

Focussing on the interior solution, from the ICCs for the two efforts one can derive the agent's decision regarding (e_1, e_2):

$$e_i(s_1, s_2) = \frac{s_i v - \delta s_j}{v^2 - \delta^2},$$

for $i = 1, 2$. Given the expression for the agent's decision, it is easy to check that if $\delta > 0$ then there is a substitution effect: the effort in one task decreases when incentives provided to the other task increase.

The solution of the principal's problem determines the optimal s_1^* (the term s_2^* of the compensation scheme is symmetric):[26]

$$s_1^* = \frac{1 + r\delta\left(\sigma_{12} - \sigma_2^2\right) + rv\left(\sigma_2^2 - \sigma_{12}\right)}{1 + r\left(2\delta\sigma_{12} + v\sigma_1^2 + v\sigma_2^2\right) + r^2\left(v^2 - \delta^2\right)\left(\sigma_1^2\sigma_2^2 - \sigma_{12}^2\right)}.$$

Since there are many effects at work in the expression defining the optimal shares s_1^* and s_2^*, we present two extreme cases of the general situation to better understand the effects at work:

1. If the two tasks are not related to each other in their cost structure, $\delta = 0$, but the random shocks are correlated, $\sigma_{12} \neq 0$, then the incentive mechanism is

$$s_1^* = \frac{1 + r\sigma_2^2 v - rv\sigma_{12}}{\left(1 + rv\sigma_1^2\right)\left(1 + rv\sigma_2^2\right) - r^2v^2\sigma_{12}^2},$$

   which depends on both tasks' variance shocks and their covariance. For $\sigma_1^2 = \sigma_2^2$, s_1^* is decreasing in σ_{12}: the higher the covariance of the two tasks, the lower the weight of each outcome on the payment scheme. The reason is that with a high covariance, the outcomes of the two tasks move together and the incentives for the effort on one task derive from the payments on the output of both tasks.

[26] If the two tasks are independent ($\delta = 0$) and there is no correlation of the random shocks ($\sigma_{12} = 0$), then the incentives are separable,

$$s_i = \frac{1}{1 + rv\sigma_i^2} \text{ for } i = 1, 2,$$

and the payment scheme is the same as the one obtained in the single-task moral hazard problem.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

2. If the two tasks are related to each other in the cost structure ($\delta \neq 0$) but there is no correlation between the random shocks ($\sigma_{12} = 0$), then in the optimal contract,

$$s_1^* = \frac{1 + r(v - \delta)\sigma_2^2}{1 + rv\left(\sigma_1^2 + \sigma_2^2\right) + r^2\left(v^2 - \delta^2\right)\sigma_1^2\sigma_2^2}.$$

As expected, s_1^* is decreasing on σ_1^2. Moreover, it is also decreasing on σ_2^2. Therefore, when efforts in the two tasks are substitutes then the optimal shares s_1^* and s_2^* are complementary. If $\sigma_1^2 = \sigma_2^2 = 0$ then $s_1^* = s_2^* = 1$ and the agent is the residual claimant for both tasks. But if $\sigma_1^2 = 0$ and $\sigma_2^2 > 0$, then

$$s_1^* = \frac{1 + r(v - \delta)\sigma_2^2}{1 + rv\sigma_2^2} \qquad \text{and} \qquad s_2^* = \frac{1}{1 + rv\sigma_2^2},$$

where $s_1^* < 1$ if and only if $\delta > 0$. Therefore, even if the outcome of task 1 is a perfect measure of the effort in this task, the principal decreases the incentives associated with the outcome of task 1 when tasks are substitutable not to harm the effort supplied in task 2. On the other hand, if the tasks are complementary then the optimal s_1^* is higher than 1. In contrast, the parameter s_2^* does not depend on δ and it is the same as in the traditional moral hazard.

Finally, if task 1 can be measured and task 2 cannot (which can be represented by $\sigma_2^2 = \infty$) but the agent has some intrinsic motivation for this task, then the optimal scheme is based on

$$s_1^* = \frac{1 - \delta/v}{1 + r\sigma_1^2\left(v - \delta^2/v\right)}, \text{ and } s_2^* = 0.$$

Here, if the tasks are substitutes ($\delta > 0$) putting effort into one task increases the marginal cost of the other.[27] Therefore, the principal gives the agent a lower incentive to exert effort in task 1 even when it is easily measurable because she does not want to discourage the agent's effort in task 2, which cannot be directly motivated. The higher the cross-effort effect is (that is, the more substitutable the effort levels are), the lower the optimal s_1^*.[28] In contrast, if the tasks are complements ($\delta < 0$) the opposite happens, and the agent will be highly motivated to perform task 1 to encourage effort in the unmeasurable task.

2.4 Incentives to the Agent vs Incentives to the Principal

In many situations, it is not only the agent who must submit an effort or take a decision, but the principal's contribution is crucial for the relationship and, just like the agent's, it is not

[27] In a situation where the principal cares especially about the non-measurable task, Holmström and Milgrom (1991) show that it is best not to provide any incentive to the task with measurable output.

[28] Dam and Ruiz-Pérez (2012) study a model where a risk-neutral agent subject to limited liability exerts effort in two tasks. When the efforts in the two tasks are independent of each other, the optimal contract is a debt contract. However, if the tasks are substitutes, then revenue sharing emerges as an optimal agreement. Ghatak and Pandey (2000) also show the optimality of sharing contracts when the risk-neutral agent has to supply an effort and to choose the riskiness of the production technique.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

verifiable. In these situations, the stronger the incentives to the agent (that is, the more the salary depends on the outcome) the weaker the incentives to the principal (because the less the principal's benefit depends on the outcome).

2.4.1 Model 6: Double-sided moral hazard with risk neutrality

When both the principal and the agent are risk neutral, program $(P1)$ is still interesting if we consider a double-sided moral hazard problem. In this environment, the agent chooses e and, simultaneously, the principal decides on her effort $a \in A$, at a cost of $c(a)$, with $c(.)$ increasing and convex. Following the analysis of Bhattacharyya and Lafontaine (1995), assume that the outcome of the relationship depends on both e and a according to

$$x = h(e, a) + \varepsilon$$

where the function $h(.,.)$ is increasing and concave in both arguments, the cross-partial derivative is positive, $h(0, a) = h(e, 0) = 0$, and ε is a random term with mean zero and variance σ^2.

The new maximization problem $(P1')$ takes into account that the outcome depends on both efforts and that there is also an ICC for the principal:

$$\begin{gathered}\underset{\{(w(x))_{x\in X},e,a\}}{Max} \{E(x - w(x) \mid e, a)\} \\ \text{s.t. } E(u(w(x)) \mid e, a) - v(e) \geq \underline{U} \\ e \in \arg\max_{\widehat{e}\in E} \{E(u(w(x)) \mid \widehat{e}, a) - v(\widehat{e})\} \\ a \in \arg\max_{\widehat{a}\in E} \{E(x - w(x) \mid e, \widehat{a}) - c(\widehat{a})\}.\end{gathered}$$

Bhattacharyya and Lafontaine (1995) show that, without loss of generality, the optimal sharing rule can be represented by a linear contract

$$w(x) = F + sx$$

for some sharing $s \in (0, 1)$. A linear contract is not the unique way to achieve the optimal solution for $(P1')$ but there is always an optimal solution that is linear.[29] In terms of incentives, the crucial element of any optimal contract is its slope at the optimum. By choosing a linear rule with the slope of any optimal sharing rule (and adjusting the fixed fee), exactly the same incentives and total payments can be achieved as with the initial rule.

The optimal sharing s^* makes a trade-off between providing incentives to the agent and providing incentives to the principal. Once s^* is determined, the fixed part of the contract F^* is easily obtained because the agent's participation constraint is binding. While it is not possible to obtain simple, closed-form expressions for the optimal sharing rule and the levels of the optimal efforts on the part of the two parties in general, a very simple example allows us to grasp most of the intuitions.

[29] Romano (1994) obtains a similar result.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Example 3 Double moral hazard with linear outcome Following Ghatak and Karaivanov (2014), consider[30]

$$h(e, a) = \alpha\theta_A\theta_P + \theta_A e + \theta_P a$$

where $\theta_A \geq 1$ and $\theta_P \geq 1$ represent the agent's and principal's ability to perform his or her task, respectively, and α is a parameter capturing the extent of the types' complementarity in production. Moreover, $v(e) = \frac{1}{2}e^2$ and $c(a) = \frac{1}{2}a^2$.[31] Then, the optimal sharing s^* derived from $(P1')$ is

$$s^* = \frac{\theta_A^2}{\theta_A^2 + \theta_P^2},$$

which gives more weight to the relatively more important participant: $s^* > 1/2$ (and the share that goes to the principal satisfies $(1 - s^*) < 1/2$) if and only if θ_A is larger than θ_P, and s^* is increasing in θ_A and decreasing in θ_P. Given the optimal contract, the efforts under double-sided moral hazard are

$$e^* = \theta_A s^* = \frac{\theta_A^3}{\theta_A^2 + \theta_P^2} \text{ and } a^* = \theta_P\left(1 - s^*\right) = \frac{\theta_P^3}{\theta_A^2 + \theta_P^2}. \tag{16.9}$$

Both efforts are lower than the corresponding first-best efforts that in this model are $e^{FB} = \theta_A$ and $a^{FB} = \theta_B$.[32] However, the optimal sharing rule solves the trade-off with respect to the incentives for the principal and the agent by inducing a smaller distortion to the most important participant, that is,

$$e^{FB} - e^* > a^{FB} - a^* \Longleftrightarrow \theta_A < \theta_P.$$

3 BEHAVIORAL APPROACH

The classical moral hazard problem assumes full rationality and standard preferences. Recent behavioral research on moral hazard, encouraged by experimental results, attempts to understand the implications of agents' non–fully rational and non–purely selfish preferences on the shape of the incentive contracts. What follows is not a review of the behavioral literature but provides some examples of how departures from the classical model affect the

[30] We note that Example 3 does not satisfy all the assumptions of Model 6 because $h(0, a) > 0$, $h(e, 0) > 0$, and $\frac{\partial^2 h}{\partial a \partial e}(e, a) > 0$.

[31] The objective in Ghatak and Karaivanov (2014) was to study the contractual choice in agriculture, taking into account that two different tasks are necessary, following a model in the spirit of the classic Eswaran and Kotwal (1985) model.

[32] The first-best cannot be achieved (even though both partners are risk neutral) because there is no "budget-breaker" (or residual claimant) that is, it is not possible to propose a contract where the total remuneration of the principal and agent is higher than the outcome sometimes and lower other times. This is similar to Holmström (1982) who shows that joint production cannot lead to efficiency when all the income is distributed amongst the agents, i.e., if the budget constraint always binds.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

conclusions obtained in section 2. We briefly present the consequences of considering other-regarding preferences, intrinsic motivation, loss aversion, and overconfidence.[33] In each of the following behavioral approaches, a new effect appears. For example, in the inequality aversion extension, the incentive contract is the issue of the trade-off between insurance, incentives, and fairness. Similarly, in the overconfidence extension the optimal payoff scheme makes a trade-off between optimal risk sharing, incentives, and gambling.

3.1 Other-regarding Preferences

There is evidence pointing to the existence of people or institutions who are not just concerned about their own payment scheme but also care about other participants' well-being (see Rabin, 2002, Englmaier, 2005, and Sobel, 2005, for reviews of the literature). To illustrate the meaning of these type of preferences in our framework, consider the interaction between a principal and an agent in which case the payment of the participants are described by the vector $(x - w(x), w(x))_{x \in X}$. If the agent (a similar argument can be done for the principal) is only concerned about the second element of this vector then we are in the classical framework. In contrast, when he cares about the whole vector of payoffs (or of utilities) then we say that the agent has other-regarding preferences.

There are several ways in which the agent can experience other-regarding preferences. A first possibility is that he has a utility function similar to a *weighted social preference* (Segal and Sobel, 2007), such as

$$U(w(x), x - w(x), e) = u\,(w(x) + \delta(x - w(x))) - v(e).$$

In this case, the agent is altruistic if δ is positive, while he is spiteful if δ is negative.[34] The sign of δ also determines whether the contract is more or less costly for the principal, as compared to the classical framework where $\delta = 0$.

Dur and Glazer (2008) study an environment where the agent is other-regarding because he envies the principal. They show that envy tightens the agent's participation constraint and the optimal contract calls for higher wages and lower effort requirements.

Inequality aversion is a second form of other-regarding preferences (see Fehr and Schmidt, 1999, and Bolton and Ockenfels, 2000). Focusing again on the agent's behavior and denoting the difference between the principal and the agent's payoffs as $d(x) \equiv x - 2w(x)$, an easy example of a utility function that represents an inequality-averse agent is

$$u(w(x), x - w(x), e) = w(x) - \delta\,(\max\{d(x), 0\} + \gamma \max\{-d(x), 0\}) - v(e). \tag{16.10}$$

---

[33] See Köszegi (2014) for the latest survey of behavioral contract design.

[34] We concentrate the discussion on the cases where the agent is other-regarding. One can similarly define an other-regarding preference principal as

$$B(x - w(x), w(x)) = x - w(x) + \delta_P\,(w(x) - v(e))\,.$$

Also notice that the coefficient δ_i, for $i = P, A$, could be a function of the distance between $x - w(x)$ and $w(x) - v(e)$, in such a way that, for example, i may care more about the difference in earnings if s/he is the one getting less than if s/he is the one getting more as in Fehr and Schmidt (1999) and Bolton and Ockenfels (2000) discussed below.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

The parameter δ, with $\delta \geq 0$, measures the extent of the agent's concern about the difference in earnings. If $\delta > 0$, the agent is inequality averse and if $\gamma \in [0, 1)$ the agent suffers more from inequality when he is behind than when he is ahead.

Providing incentives to an inequality-averse agent is often more costly than to a classical agent. The intuition is that as the project is more profitable, more inequality is created and it is more expensive to satisfy the ICC. To illustrate this effect, we follow Itoh (2004) and consider a simple model where both participants are risk neutral, the agent is protected by limited liability, there are two efforts, $E = \left\{e^H, e^L\right\}$, and two outcomes, $X = \{x_G, x_B = 0\}$. The probabilities of success are $p(x_G \mid e^H) = p^H$ and $p(x_G \mid e^L) = p^L$, and the cost of effort is $v(e^H) = v > 0 = v(e^L)$. Assume that the principal wants the agent to exert effort e^H. Then, taking into account that the limited liability binds ($w_B = 0$), the ICC takes the form

$$(p^H - p^L)\left[w_G - \delta\left(\max\{d(x_G), 0\} + \gamma \max\{-d(x_G), 0\}\right)\right] \geq v.$$

The incentives (the left-hand side of the ICC) are decreasing in δ; hence, the principal's profits are decreasing in δ and she is in general worse off when hiring an inequality-averse rather than a classical agent.[35]

In a model where the inequity aversion is convex in the difference in the payoffs, Englmaier and Wambach (2010) find a tendency toward linear sharing rules as the agent's concern for inequity become more important, in line with other findings that the more complex the situation is the simpler the optimal incentive scheme tends to be.[36] Interestingly, given that the contract now has to balance among three objectives (risk sharing, incentives, and inequality concerns) Englmaier and Wambach (2010) also find that the sufficient statistics result is violated because optimal contracts may be overdetermined or incomplete. To understand the intuition, consider a situation with two sources of information: the outcome related to the agent's effort and another variable. First, if this other information is not related to the agent's effort but just to the principal's profit, then the second measure will be included in the contract (which will thus include non-informative performance measures) because the set of variables used in the payment scheme no longer serve only as a signal of the agent's effort but also deal with the agent's concern about an equitable treatment. Second, if the second variable (second-order stochastically) dominates the outcome, it may be optimal to concentrate the incentives on the outcome (neglecting informative performance measures) because this is the variable the agent is interested in when he is inequality concerned.

A third form of other-regarding preferences is *reciprocal behavior*, where the agent may take into account the behavior of the principal, in such a way that the agent will reciprocate and take a decision that also benefits the principal (supplying higher effort) if she takes a decision that benefits the agent (paying higher wages) (Rabin, 1993).[37] If an agent follows the previous behavior, monetary incentive and reciprocal motivation are substitute goods and the agent's PC may not be binding at the optimal contract. To illustrate how these two incentive tools are combined in the optimal contract, we present a simplified version of Englmaier and Leider's (2012) model, where the agent has reciprocal references.

---

[35] In this model, the effect is particularly straightforward because there is no inequality if the outcome is x_B.

[36] The convexity of the inequality term $\delta(d(x))$ implies an aversion to lotteries over different levels of inequity.

[37] Akerlof (1982) explained the labor relation as a gift exchange where agents respond to a generous wage scheme offered by the principal by exerting more than minimal effort. See, e.g., Dufwenberg and Kirchsteiger (2004) and Falk and Fischbacher (2006) on reciprocity in sequential games.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

We consider a risk-neutral principal and a risk-averse agent. There are two possible outcomes: success (x_G) and failure (x_B), with $x_G > x_B$. The agent chooses between two efforts: $e \in \{e^L, e^H\}$, with $p(x_G \mid e^H) = p^H > p^L = p(x_G \mid e^L)$. A contract takes the form (w_G, w_B, e^H), where we assume that the principal is interested in obtaining effort e^H. The effort e^H is not enforceable but is "the job description." The agent is reciprocal, in the sense that his expected utility under (w_G, w_B, e^H) if he provides effort e is

$$p(x_G \mid e)u(w_G) + (1 - p(x_G \mid e))u(w_B) - ve + \delta R(e^H)r(e),$$

where $R(e^H) \equiv p(x_G \mid e^H)u(w_G) + (1 - p(x_G \mid e^H))u(w_B) - ve^H - \underline{U}$ is the agent's expected rent under the job description, $r(e) \equiv p(x_G \mid e)x_G + (1 - p(x_G \mid e))x_B$ is the principal's expected revenue if the agent chooses e, and δ is the intensity of the agent's reciprocal behavior. With this utility function, the agent experiences reciprocal motivation only if the contract gives him an expected utility higher than his reservation utility (that is, if the PC does not bind). The agent's ICC is

$$\left(p^H - p^L\right)(u(w_G) - u(w_B)) + \delta R(e^H)\left(p^H - p^L\right)(x_G - x_B) \geq v\left(e^H - e^L\right).$$

It is worthwhile noticing that a high enough fixed payment $\widetilde{w}$ can implement the effort e^H. The condition is that $\widetilde{w}$ satisfies

$$\delta\left[u(\widetilde{w}) - ve^H - \underline{U}\right]\left(p^H - p^L\right)(x_G - x_B) \geq v\left(e^H - e^L\right).$$

The ICC is always binding at the optimal contract, but the agent's PC may be binding or not. The optimal contract is a standard one for low values of δ, providing no rents to the risk-averse agent, whereas it is a reciprocity contract that gives the agent a utility larger than his reservation utility for large values of δ.[38] When rents are provided to the agent, the FOC of the principal's program with respect to $w(x)$ can be written as

$$\frac{1}{u'(w^*(x))} = \mu\left(\frac{p(x \mid e^H) - p(x \mid e^L)}{p(x \mid e^H)} + \delta\left(p^H - p^L\right)(x_G - x_B)\right),$$

which implies that monetary and reciprocity motivations are substitutable.[39]

[38] In fact, for very large values of δ the first-best solution can be arbitrarily closely approximated with a contract that gives the agent an infinitesimal rent (Englmaier and Leider, 2012).

[39] Behavioral models have been very useful in analyzing multi-agent situations. It is interesting to note that since inequality-averse agents care about other agents' remuneration, to reduce inequity among agents, their payments will tend to depend on other agents' performance, even if they are statistically and technologically independent. For example, Englmaier and Wambach (2010), Goel and Thakor (2006), and Bartling (2011) show that inequity aversion or envy among agents may render team incentives optimal. Itoh (2004) finds that inequity aversion when agents are subject to limited liability may allow agency costs to be reduced. Rey-Biel (2008) finds that the principal can always exploit inequity aversion to extract more rents from her agents. Demougin and Fluet (2006) compare group and individual bonus schemes for behindness-averse agents and derive conditions under which either scheme implements a given effort level at least costs.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

3.2 Extrinsic and Intrinsic Motivation

The classical moral hazard model is based on designing incentives to provide extrinsic motivation to the agent. The so-called extrinsic motivation is the one that is derived from the monetary incentive scheme. In contrast, an agent's *intrinsic motivation* comes from the utility obtained from achieving some goal set by himself, the society, the principal, or from working for a particular type of principal, such as one who honors some community (environmental or another form of social) standards. The simplest agent's utility function that represents an agent with both intrinsic and extrinsic motivation is

$$U(w, e, s) = w + Im - v(e),$$

where $I \geq 0$ is the intrinsic motivation ($I = 0$ in the classical model), and m is the source of this motivation. We briefly discuss the consequences of some sources of intrinsic motivation.[40]

Intrinsic motivation may come from the agent's perception of the world, which may depend on the contract the principal offers. The underlying idea in this approach is the following. The agent expects to offer a predetermined effort and to receive a fixed-fee (first-best) payment. However, if he is offered an incentive contract instead, his perception of the relationship changes and he becomes aware of the possibility of shirking.[41] Thus, providing extrinsic incentives for the agent can be counterproductive because it may crowd out his intrinsic motivation, leading to lower effort levels and lower profits for the principal (Kreps, 1997, and James, 2005). For example, Auster (2013) and Von Thadden and Zhao (2012, 2014) study a situation where agents are *unaware* of the full effort problem and they make a default effort when offered an (incomplete) full insurance payoff, while they become aware of the effort problem and behave strategically if they are offered the optimal moral hazard contract (see also James, 2005). Similarly, Bénabou and Tirole (2003) consider a principal–agent model where the principal is better informed than the agent about the agent's characteristic, and show that performance incentives lead to an increase of the agent's effort in the short run but they are negative reinforcements in the long run.[42]

Another source of intrinsic motivation may be due to some characteristic (or to a verifiable decision) of the principal (Murdock, 2002, and Besley and Ghatak, 2005). To better explain this approach, suppose that the two participants are risk neutral, the agent's payoff is constrained to be non-negative (that is, there is limited liability), and there are two possible outcomes: success (x_G) and failure (x_B), with $x_G > x_B = 0$. The set of possible efforts is $E = [0, 1]$ and $p(x_G \mid e) = e$. A principal has a certain type, or a mission orientation, and the agent's preferences can be aligned with a particular mission or with none of them. The public type of the principal is τ, with $\tau \in \{0, M\}$. A type-0 principal has no mission and is the traditional profit-maximizing partner, whereas type-M principals have a mission. The agent is mission-oriented. The source of his intrinsic motivation is that he cares about the

[40] Some cases of other-regarding behavior, such as that of social preferences (for $m = x - w(x)$) or the reciprocal motivation model presented in the previous subsection, can also be understood as models of intrinsic motivation.

[41] This can be seen as a form of bounded rationality.

[42] The intuition is that when the principal pays a bonus to induce low-ability agents' to work (the principal increases the agent's extrinsic motivation), then the agent perceives the bonus as a bad signal about his own ability (she reduces the agent's intrinsic motivation). Kirkegaard (2015) studies a model where the agent works for the principal and simultaneously pursues private benefits. He shows that the optimal contract may offer high rewards but flat incentives to lessen the agent's incentive to pursue private benefits, his intrinsic motivation.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

success of his job when he works for a principal with a mission.[43] His utility function can be represented by[44]

$$U(w, e, x) = w + I(x, \tau) - \frac{1}{2}e^2, \tag{16.11}$$

where $I(x, \tau)$ depends on the outcome x and the type of the principal τ. In case of failure $I(x_B, \tau) = 0$, for all τ. In case of success, $I(x_G, \tau = M) > I(x_G, \tau = 0) = 0$, that is, when the mission-oriented agent works for a type-0 principal, he behaves as a traditional agent. $I(x_G, \tau = M)$ is the intrinsic utility that the agent derives from the success of his work for a type-M principal.

This model is an extension of Example 2 (with $v = 1$) and the expression for the optimal bonus in this environment is also very similar to (16.8):

$$w^*(x_G) = \begin{cases} \sqrt{2\underline{U}} - I(x_G, \tau), \text{ if } \frac{(x_G + I(x_G,\tau))^2}{8} \leq \underline{U} \leq \frac{(x_G + I(x_G,\tau))^2}{2} \\ \max\left\{0, \frac{x_G - I(x_G,\tau)}{2}\right\}, \text{ if } \underline{U} < \frac{(x_G + I(x_G,\tau))^2}{8} \end{cases}. \tag{16.12}$$

The effort implemented by the agent is $e^* = \sqrt{2\underline{U}}$ for the intermediate region of $\underline{U}$ whereas it is $e^* = \max\left\{I(x_G, \tau), \frac{x_G + I(x_G,\tau)}{2}\right\}$ when $\underline{U}$ is low. Thus, a higher intrinsic motivation $I(x_G, \tau)$ results in a higher (or equal) effort by the agent at a lower cost in terms of bonus.[45]

3.3 Loss Aversion

There is also evidence that some individuals do not evaluate payoff in absolute terms but in comparison with some reference point (Kahneman and Tversky, 1984). *Loss aversion* is the reference-dependent preference that has been most studied, a type of preference that may explain why contracts framed as bonuses are much more prevalent than contracts framed as penalties (see, e.g., Aron and Olivella, 1994). The idea is that, evaluated at the reference point, the marginal utility of a loss is larger than the marginal utility of a gain, so that the agent's utility function has a kink at this reference point. We present the basic loss aversion model by De Meza and Webb (2007) with an exogenous reference wage, which fits within the structure of Model 2, where $e \in \{e^L, e^H\}$. Consider that the principal is risk neutral and the agent is risk averse with loss aversion with respect to a reference wage w^R:

---

43 A type-0 agent would be the traditional agent who does not care about the type of the principal and we would be back to the traditional moral hazard problem, whatever the type (mission) of the principal is.

44 The model can be extended by allowing the principal to choose the "mission," taking into account the effect of the choice in the agent's incentives (Besley and Ghatak, 2005).

45 Guo (2016) also analyzes a model where the agent has extrinsic motivation, in addition to the monetary incentives, associated with a principal's decision. In her paper, the extra motivation of the agent (an employee) comes when his principal (a manager) invests in a non-contractible employee-friendly relationship. In Guo (2016), the utility function of the agent has the form

$$U(w, e, m) = u(w) - v(e, I(m))$$

because the agent's extrinsic motivation influences his effort disutility. The investment in the relationship can also benefit the principal because, for example, the employee may support the manager if she faces a replacement threat.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

$$U(w,e)=u(w)-1_{w<w^R}l\left(u\left(w^R\right)-u(w)\right)-v(e),$$

where the index $1_{w<w^R}=1$ if $w<w^R$ and $1_{w<w^R}=0$ if $w\geq w^R$; and $l>0$ is the loss the agent suffers when the wage is lower than the reference wage w^R.

Assume that MLRP is satisfied and that the principal aims to implement the high effort. Then, from the first-order condition of the principal's problem, we obtain

$$\frac{1}{u'(w^*(x))}=(1+\nu(x)l)\left(\lambda+\mu\frac{p(x\mid e^H)-p(x\mid e^L)}{p(x\mid e^H)}\right). \tag{16.13}$$

where $\nu(x)\in[0,1]$ is an instrument to handle the kink in the agent's utility. If $w<w^R$, loss aversion applies and $\nu(x)=1$; if $w>w^R$, loss aversion does not apply and $\nu(x)=0$; and if $w=w^R$, then the ICC holds with equality for $\nu(x)\in[0,1]$.

From (16.13), De Meza and Webb (2007) derive that loss aversion does not affect the condition if w^R is very low ($\nu(x)=0$ for all x) and, as in Model 2, $w^*(x)$ is strictly increasing in outcome. Similarly, if w^R is very high ($\nu(x)=1$ for all x) then $w^*(x)$ is also strictly increasing in outcome. However, in the remaining cases loss aversion affects the optimal payment scheme, and there are zones (for the lower, the intermediary, or the higher outcomes) where the agent receives a flat wage equal to w^R.[46]

In the presence of loss aversion or reference-dependent preferences, the principal designs incentives by taking into account both the induced risk sharing and the agent's loss aversion. The wage scheme will be a function of the outcome at least for certain outcomes, but it tends to have a significant number of outcomes where the payment is flat.[47]

3.4 Overconfidence

Contracts are based on the principal's and agent's beliefs (correct or incorrect) and, in the standard moral hazard model, it is customary to assume that both participants share the same beliefs about the uncertain elements of the relationship. However, we may think of situations where beliefs are different but each knows the view of the other (they "agree to disagree"). This may correspond to environments where the agent is "optimistic" and "overconfident" or he has different beliefs to the principal concerning his abilities (he can have a positive or negative self-image as compared to what the principal thinks). Santos-Pinto (2008) and De la Rosa (2011) consider a moral hazard model when the principal and agent have (public) asymmetric beliefs.

We present the basic elements and results within a structure close to Model 1. Consider that the principal is risk neutral and the agent is risk averse, with utility function $U(w,e)=u(w)-ve$, and that there are two possible outcomes: success (x_G) and failure (x_B), with $x_G>x_B$. We denote $p_P(x\mid e)$ and $p_A(x\mid e)$ the principal's and the agent's beliefs for outcome $x\in\{x_G,x_B\}$ for a given effort $e\in E$. The beliefs are asymmetric if $p_A(x\mid e)\neq p_P(x\mid e)$ for at

[46] They also show that if, in addition, there is a limited liability constraint $w\geq\underline{w}$ (with $\underline{w}<w^R$), then it is optimal not to have payments in the interval $(\underline{w},w^R)$ and $w^*(x)$ is discontinuous.

[47] Köszegi and Rabin (2006, 2007) show that the optimal payment scheme often has two wages (and incentives are based on a bonus). De Meza and Webb (2007) find that when the reference wage is the median wage, the incentives are based on performances over the median. When the reference point is endogenous, Herweg, Müller, and Weinschenk (2010) show that the rational expectation about the wage is the expected wage.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

least some $e \in E$, and we can say that the agent exhibits a positive self-image of own ability (or he is optimistic) if $p_A(x_G \mid e) > p_P(x_G \mid e)$, for all $e \in E$.

The existence of different beliefs affects the contract even under symmetric information. In this case, the wage scheme to implement an effort e satisfies

$$\frac{1}{u'(w^*(x))} = \lambda \frac{p_A(x \mid e)}{p_P(x \mid e)}.$$

Hence, full insurance (which results when $p_A(x \mid e) = p_P(x \mid e)$) no longer holds. Since the principal and agent have different views of the uncertain situation they are involved in, they can agree on a side-bet in such a way that both think they can gain. In the first-best, an optimistic agent will be paid more in the case of success (because the principal thinks this bonus will not be paid that often) and a pessimist will be paid more in the case of failure.[48] In addition, whether the agent is right or wrong in his beliefs, with the contract he will obtain his reservation utility according to his subjective beliefs (PC always binds). From her perspective, the cost to the principal of implementing high efforts is lower than in the standard model and it decreases with the agent's optimism and overconfidence.[49]

After the analysis of the contracts under symmetric information, it is easy to see that under moral hazard it can be the case that it is less expensive for the principal to implement the high rather than the low effort.[50] If the agent is optimistic or overconfident enough, the first-best risk-sharing incentive scheme may induce the agent to exert high effort under moral hazard. In general, as shown by Santos-Pinto (2008), to induce the agent to work the asymmetry of beliefs can be either favorable or unfavorable, depending on whether the agent is overconfident or the opposite. De la Rosa (2011) and Gervais, Heaton, and Odean (2011) highlight that the reason for the asymmetries also matters. Incentive contracts are sensitive to the kind and level of overconfidence, not only to the presence of overconfidence per se. For example, in De la Rosa (2011) beliefs take the functional forms

$$p_P(x_G \mid e) = q_P + \theta_P e, \text{ and } p_A(x_G \mid e) = q_A + \theta_A e,$$

with $q_I > 0$, $\theta_I > 0$ and $q_I + \theta_I < 1$ for $I = P, A$. Then, assuming $e \in \{e^L = 0, e^H = 1\}$, if $q_A > q_P$ the agent is optimistic, if $\theta_A > \theta_P$ the agent is overconfident, and he is overconfident overall if $q_A > q_P$ and $q_A + \theta_A > q_P + \theta_P$. If the high effort is implemented, the principal's expected profit increases in both the agent's level of optimism and overconfidence. But if the low effort is implemented, the principal's expected profit increases in the agent's level of optimism or pessimism, for an optimistic or a pessimistic agent, respectively, and it decreases in the agent's overconfidence if the agent is significantly optimistic.

[48] As Santos-Pinto (2008) points out, if an agent is risk neutral and has mistaken beliefs, the principal's problem does not have a solution because the principal can always increase her profits by raising the stakes of the side-bet. This implies that when the agent is risk neutral but is protected by limited liability then, in the optimal contract, the limited liability constraint is binding.

[49] It is usual to assume that the agent is the one mistaken about the real technical conditions of the production process, but it is also possible that the opposite is true.

[50] Another classical result in moral hazard with symmetric beliefs is that for the lowest effort the optimal contract under moral hazard and under symmetric information coincide. This result may not hold under overconfidence.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

4 PRINCIPAL–AGENT MARKETS

The models that we have discussed above, and almost all the papers that study settings involving a moral hazard problem, take the identity and the characteristics of the participants in the relationship as given. They consider an isolated principal–agent situation (or an isolated relationship among several principals and/or several agents) and analyze the optimal contract (contracts) in this relationship. The principal assumes all the bargaining power, and the agent is ready to accept a contract as long as it guarantees him his exogenously given reservation utility.

The previous description is a good fit for situations where the participants in a contract cannot be easily replaced, as is the case for the relationship between a regulator and a firm. However, most often, a principal can look for alternative agents and an agent can look for alternative principals. When several principals and several agents exist in this "market," in addition to the question about the optimal contracts, we can address the endogenous determination of the identity of the pairs that meet (i.e., the matching between principals and agents). In particular, we can study whether, at equilibrium, there is *positive assortative matching* (PAM) or *negative assortative matching* (NAM). A PAM between principals and agents with respect to, say, ability (or any other characteristic, such as risk, type, etc.) exists if the partner A of a principal P with a higher ability than another principal P' has a higher or equal level of ability than the partner A' of the principal P'. A negative assortative matching is defined in a similar manner.

Furthermore, the alternative relationships that could be formed in the market are crucial to understanding the endogenous level of payoffs that each principal and agent obtain and some of the properties of the contract.

The theory of "two-sided matching models" provides the tool with which to study markets where heterogeneous players from one side (principals) meet with heterogeneous players from the other side (agents).[51] The equilibrium of the market determines the identity of the partners that actually sign contracts (that is, the "matching") together with the profits that they obtain and the characteristics of the contracts.[52] Equilibrium outcomes satisfy two useful properties. First, equilibrium contracts are always Pareto optimal; hence, we can use what we have learned from the analysis of isolated relationships. Second, if utility is transferable (that is, it is possible to transfer one unit of utility from the principal to the agent) then any equilibrium matching is efficient in the sense that it maximizes "total surplus": the sum of all the profits in the market cannot be increased by reassigning principals and agents.

We now discuss some of the new lessons from matching models with contracts.

---

[51] The book by Roth and Sotomayor (1990) made the theory of two-sided matching models popular and accessible. Gale and Shapley (1962) started it by studying "the marriage market," where each participant (in their case, a man or a woman) is only concerned about the characteristics of the members of the other side of the market (women or men, respectively). Shapley and Shubik (1972) broaden the set of applications of this theory by considering, in "the assignment model" that the utility derived from a relationship not only depends on the characteristics of the partner but also on money, which can be exchanged among partners as part of the agreement.

The assignment model can be easily extended to situations where not only money but also contracts are endogeneously decided simultaneously with the matching, as long as utility is transferable one-to-one between principals and agents. In other cases, the analysis is more complex but the contribution by Kaneko (1982) allows us to also use this tool.

[52] Any competitive equilibrium is also a stable outcome and vice versa, where stability means individual rationality together with the property that no principal–agent pair can be better off by leaving their current partners and signing a new contract among them.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

4.1 The Relationship Between Risk and Incentives

A quite robust prediction of the moral hazard literature is the negative relationship between risk and performance pay (e.g., Holmström and Milgrom, 1987): the higher the risk of the project the agent is working on, the lower the incentives included in the contract. When the agent has CARA risk preferences, we have seen in (16.2) that the variable part of the contract s^* is decreasing in the risk of the relationship, represented by the variance σ^2. The CARA assumption also implies that utility is transferable because the principal can give or take away utility directly through the fixed part of the contract F.

To study whether this conclusion also holds when principals and agents interact in a market, consider that there exists a set of principals who are heterogeneous in the risk (variance) of their production process: each principal is associated with the variance of her project σ^2, with $\sigma^2 \in \left[\sigma_L^2, \sigma_H^2\right]$. There is also a set of agents, heterogeneous in their risk-aversion attitude: each agent is identified by his degree of risk aversion r, with $r \in [r_L, r_H]$. Both populations have the same mass.[53]

Serfes (2005) analyzes how the degree of risk aversion of an agent relates to the risk of the project he is involved in at equilibrium. He provides the answer for two interesting cases: (a) if $\sigma_L^2 r_L \geq 1/v$ (that is, the risk and/or the degree of the agent's risk aversion are always large) then there is PAM: low risk-averse agents are matched with low-risk principals (projects) and vice versa; and (b) if $\sigma_H^2 r_H \leq 1/v$ (that is, the risk and/or the degree of the agent's risk aversion are always small) then there is NAM: low risk-averse agents are matched with high-risk principals (projects) and vice versa.

If we now rethink the relationship between risk and performance pay, there are two effects. There is the direct effect of σ^2 on s, the same that is present in the standard principal–agent model, which is always negative. There is also an indirect effect of σ^2 on s through the assignment that may be negative (if PAM, because a high σ^2 is matched with a high r, which leads to a low s) or positive (if NAM). Thus, while the relationship between risk and performance pay is certainly negative if $\sigma_L^2 r_L \geq 1/v$ (because of PAM), it can be positive or have any other shape (like a U shape), otherwise.

Using a similar model, Li and Ueda (2009) analyze the relationship between risk and ability. As in Serfes (2005), each principal is characterized by the variance of her project but, in contrast with that paper, Li and Ueda (2009) assume that agents are heterogeneous in terms of ability. At equilibrium a better agent is matched with a firm whose project has lower variance. In their set-up, this provides an explanation for the fact that safer firms receive funding from more reputable venture capitalists.

4.2 The Nature of the Matching Between Principals and Agents Under Moral Hazard

The presence of moral hazard in a relationship not only changes the characteristics and efficiency of the contract, it may also influence the identity of the principals and agents that decide to establish a partnership. We illustrate a reversal in the nature of the matching using the model introduced as Example 3, due to Ghatak and Karaivanov (2014), where principals

[53] Although the matching models typically involve a finite set of members on both sides, here we present the continuous model, as in Serfes (2005), because the conditions are easier to write. See also Serfes (2008) for a similar analysis with discrete sets.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

and agents are risk neutral and the relationship of a principal with characteristic θ_P and an agent with characteristic θ_A produces an output of $h(e,a) = \alpha\theta_A\theta_B + \theta_A e + \theta_P a + \varepsilon$.

We now consider a finite set of heterogeneous principals, each endowed with a characteristic θ_P and a finite set of heterogeneous agents, each endowed with a characteristic θ_A. Assume for simplicity that the size of the two sets is the same, and θ_A and θ_P are always higher than 1.

If efforts are contractible, then the first-best efforts are $e^{FB} = \theta_A$ and $a^{FB} = \theta_B$, and the expected value of the outcome is $h\left(e^{FB}, a^{FB}\right) = \alpha\theta_A\theta_B + \theta_A^2 + \theta_P^2$. Thus, if we take into account the cost of the effort, the joint surplus in the relationship, as a function of the characteristics (θ_A, θ_P), is

$$S^{FB}(\theta_A, \theta_P) = \alpha\theta_A\theta_B + \frac{1}{2}\left(\theta_A^2 + \theta_P^2\right).$$

For every $\alpha \geq 0$ the function is increasing in the characteristics θ_A and θ_P and the cross-partial derivative $\frac{\partial^2 S^{FB}}{\partial\theta_A\theta_P}(\theta_A, \theta_P)$ is non-negative. Then, applying results by Legros and Newman (2002) (see also Becker, 1973), the equilibrium satisfies PAM: principals with a high characteristic θ_P end up working with agents with a high characteristic θ_A, and vice versa.[54]

If efforts are not contractible, then the optimal sharing rule decided by any partnership formed makes a trade-off between providing incentives to the principal and the agent. The second-best efforts are given by (16.9) and the joint surplus in the relationship is

$$S(\theta_A, \theta_P) = \alpha\theta_A\theta_B + \frac{1}{2}\left(\theta_A^2 + \theta_P^2\right) - \frac{1}{2}\frac{\theta_A^2\theta_P^2}{\left(\theta_A^2 + \theta_P^2\right)}.$$

The cross-partial derivative of $S(\theta_A, \theta_P)$ is now negative for positive but low values of α. Therefore, if α is low then the equilibrium satisfies NAM: principals with a high characteristic θ_P end up working with agents with low characteristics θ_A and vice versa.

Due to the incentive problem, the modularity of the joint surplus under moral hazard depends on both the complementarity of the characteristics in the production function and the endogenous efforts, which depend on the optimal sharing rule. This rule provides incentives to each participant as a function of the magnitude of his/her type relative to the other. Better incentives are provided to θ_P when θ_A is low rather than when it is high. Therefore, the positive effect of an increase in, say, θ_P on e_P is lower the higher θ_A is. This effect induces a certain substitutability between the types that more than compensates the complementarity in the production function when α is low.

Chakraborty and Citanna (2005) and Kaya and Vereshchagina (2015) also study the nature of the matching in a market where each partnership is subject to double-sided moral hazard. In their contributions, the market has only "one side," instead of "two sides," that is, each of the participants can play either of the two roles in the partnerships. Chakraborty and Citanna (2005) propose a model where individuals are heterogeneous in wealth and are subject to limited liability. The wealth level of the individual can matter because of the

[54] Most models that analyze whether the equilibrium satisfies PAM or NAM consider joint surplus functions that are twice differentiable in the characteristics and hence they use the cross-partial derivative of the joint surplus to assert the nature of the matching. However, as Besley and Ghatak (2005) state, non-standard matching arguments are needed in the analysis of horizontal characteristics, that is, when the value function is not twice differentiable in the arguments, for example because it depends on the distance between the characteristics of the principal and the agent.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

limited liability, but everyone has identical incentives to hire a rich individual. Thus, under symmetric information, any matching is efficient. However, one of the tasks in the partnership is more effort-intensive than the other. Under moral hazard, to facilitate incentive provision, richer individuals have to be allocated to more effort-intensive tasks, which results in NAM at equilibrium. In Kaya and Vereshchagina (2015), individuals are heterogeneous because (for a given effort) their contribution is different: some individuals are better than others. They study a repeated interaction where, once a partnership is formed, the partners produce a stochastic output in each period. As before, in the absence of moral hazard, equilibrium sorting is indeterminate. However, there are two cases in which moral hazard leads to the formation of heterogeneous teams: when one of the partners makes an inefficient effort (NAM is due to the same reasons as above), and when the optimal level of effort can be sustained by both partners at the beginning of the relationship, and the partners' types either increase the marginal product of effort or have little impact on the output, so that the output realization is a very informative signal of the effort (NAM is efficient in this case because it allows better punishment strategies).

The moral hazard problem also has an influence on the equilibrium sharing of the surplus between principals and agents. Although total surplus is reduced because of the moral hazard, an agent with a high θ_A may end up obtaining higher rents because of the existence of the moral hazard problem. When there is moral hazard, a "good" agent is more appealing for a principal with low θ_P, who would be ready to pay him more, increasing his "market bargaining power" (and his expected payoff) with a principal with a high θ_P (see Macho-Stadler and Pérez-Castrillo, 2014).

In the previous model, the moral hazard problem induces a reversal of the nature of the partnerships compared to the first-best matching because the need to provide incentives to both participants makes "asymmetric" partnerships more profitable. When only the agent is subject to moral hazard, a reversal may also happen when the principal can choose between two different instruments. Alonso-Paulí and Pérez-Castrillo (2012) study a situation where the agent receives information about the state of the world after having signed the contract, and this information is relevant for the choice of the optimal effort. The principal can offer either an incentive contract or a contract with a verifiable, but rigid, effort. The second type of contract allows for better management control, but makes it hard for the agent to react to market conditions. Although the matching between principals and agents is PAM when only one type of contract is used in all the partnerships,[55] the best principals might be willing to renounce hiring the best agents through incentive contracts, signing rigid contracts with lower-ability agents instead.[56]

4.3 Heterogeneity, Profits, and Efficiency

The utilities that principals and agents obtain at equilibrium are endogenous and depend on the sizes of the populations of principals and agents as well as on their characteristics. To highlight

[55] Legros and Newman (2007) provide sufficient conditions for monotone matchings in environments where, as is the case in the framework of Alonso-Paulí and Pérez-Castrillo (2012), utility is not fully transferable.

[56] In dynamic relationships, the agreements can also be governed by two types of contracts: short-term and long-term contracts. When information on the workers' ability is revealed during the relationships, the market dictates a trade-off between the optimal matching (which requires that principals sign short-term contracts) and incentives (which requires long-term contracts). At equilibrium, the matching is not necessarily PAM because both types of contract can coexist (see Macho-Stadler, Pérez-Castrillo, and Porteiro, 2014).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

some of the implications of the endogenous market power of principals and agents, consider a simple modification of the model studied in Dam and Pérez-Castrillo (2006). There are n_P homogeneous principals and n_A heterogeneous agents. All participants are risk neutral. Agents differ with respect to their initial wealth. An agent a^j has an initial wealth $\underline{w}^j$, which is known to the principals, with $\underline{w}^1 \geq \underline{w}^2 \geq \ldots . \geq \underline{w}^{n_A} \geq 0$. Therefore, a contract with agent a^j needs to satisfy the limited liability constraint $w(x) \geq -\underline{w}^j$. The relationship is similar to the one introduced in Example 3, where two outcomes are possible, i.e., $x \in \{x_B, x_G\}$.

Given that principals are identical, it is necessarily the case that they obtain the same level of profits at equilibrium; we denote it by $\widehat{\pi}$. Therefore, the equilibrium contracts (which are necessarily Pareto optimal) are not governed by the agents' PC, but they will be the contracts that maximize the agents' utility subject to the principal's obtaining $\widehat{\pi}$. If $n_P < n_A$, then there will be n_P relationships and $\widehat{\pi}$ is the maximum benefit that a principal can obtain by contracting with agent a^{n_P}, or with agent a^{n_P+1} (which is the richest agent that does not sign any contract).[57] Even though agents are the long side of the market, those with an initial wealth higher than $\underline{w}^{n_P}$ obtain rents and, in fact, they sign a contract that is more efficient than the principal–agent contract, in the sense that effort is closer to the first-best. The rents and the efficiency of the contract signed by agent a^j do not depend on the absolute value of $\underline{w}^j$ but on the relative value of $\underline{w}^j$ compared to $\underline{w}^{n_P}$. Similarly, if $n_P > n_A$, then there will be n_A relationships, $\widehat{\pi} = 0$, and all the rents will go to the agents.

As principals compete for the wealthier agents, they are compelled to offer better contracts in order to attract them. These agents obtain higher utility, the limited liability constraint is less stringent and hence the effort level approaches the first-best. The effect of competition on the power of incentives and the efficiency of the relationship has already been pointed out by Barros and Macho-Stadler (1988), in a situation where two principals compete for a good agent.[58]

The analysis of Dam and Pérez-Castrillo (2006) also indicates that a larger inequality in the distribution of agent wealth leads to more efficient relationships. In their framework, a public authority that would like to distribute some money that could serve as collateral in tenancy relations may need to induce inequality among the tenants. If it distributed a small amount to every tenant, then the relative differences in initial wealth would not change and the landowners would appropriate the additional amount distributed. On the other hand, an unequal distribution of the money among a few tenants improves the efficiency and the agents appropriate more than the additional money they receive.

4.4 Competition Among Mission-oriented and Profit-oriented Firms

In many markets, principals are heterogeneous not in terms of productivity or costs but in terms of the importance that they give to their mission. Indeed, many public bureaucracies and

[57] Stable outcomes are typically not unique. For example, in the current situation, $\widehat{\pi}$ can be any number in the interval whose lower bound is given by the benefits that a principal obtains by hiring agent a^{n_P} and the upper bound is the benefits she obtains by hiring agent a^{n_P+1}.

[58] The effect of competition on the efficiency of the incentive contracts is also the main objective of Dam (2015). Edmans, Gabaix, and Landier (2009) and Baranchuk, MacDonald, and Yang (2011) study the implications of the assignment of managerial talent to firm size. Also, Hongy, Serfes, and Thiele (2012) study a market with heterogeneous entrepreneurs and venture capital firms. They show that the entry of new venture capital firms has a "ripple effect" throughout the entire market: all start-ups receive more capital in exchange for less equity and the relationships are more efficient.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

private non-profit organizations give more weight to their mission than to profits. Also, some private profit-oriented firms give some weight to an objective other than profits (for example, the use of clean technologies or the development of the community). Similarly, as we discussed in subsection 3.2, the main heterogeneity among agents (workers) may be due not to their ability or risk aversion, but to their intrinsic motivation to work for certain types of firm. In that subsection, we characterized the optimal principal–agent contract for an agent whose intrinsic motivation to work for the firm is $I(x, \tau)$, which depends on the outcome x and the type of the principal τ (see equation (16.11)).

To discuss the role of matching the mission preferences of principals and agents, we present a model similar to Besley and Ghatak (2005). We consider a market with two types of principals and two types of agents. The types of all the participants are perfectly observable. In case of success, a profit-oriented principal receives a monetary payoff of $x_G^0 > 0$. The payoff $x_G^M > 0$ that a mission-oriented principal receives in the case of success may have a non-pecuniary component. Similarly, there are agents who only care about the monetary reward (we will refer to them as type-0 agents) whereas there are mission-oriented agents who receive an intrinsic motivation of $I^M \equiv I(x_G^M, \tau = M) > 0$ if they work for a mission-oriented firm. To simplify the number of cases, we assume that $I^M \leq x_G^M \leq 2x_G^0$, that is, the agent's intrinsic motivation is not larger than the firm's payoff. Also, we assume that the number of mission-oriented agents is the same as the number of mission-oriented principals.

At the equilibrium matching, there is segregation, in the sense that mission-oriented agents work for mission-oriented principals whereas type-0 agents work for profit-oriented principals. The matching is assortative because it raises organizational productivity.

Even though the nature of the matching is the same irrespective of the number of principals and agents in the profit-oriented sector, the agents' bonuses and the principals' profits in both sectors are affected by those numbers. Suppose first that there is full employment in the profit-oriented sector (that is, the number of type-0 agents is lower than the number of type-0 firms). Then, the equilibrium bonuses and the optimal effort levels in this market for the two types of agents are

$$w^{0*}\left(x_G^0\right) = x_G^0 \text{ and } e^{0*} = x_G^0$$
$$w^{M*}\left(x_G^M\right) = x_G^0 - I^M \text{ and } e^{M*} = x_G^0.$$

Thus, competition for the type-0 agents drives the expected payoff of type-0 principals to zero. The utility that the mission-oriented agents obtain is set by what they could obtain by switching to the profit-oriented sector. The mission-oriented principals benefit from the agents' intrinsic motivation through a reduction in the salary they need to attract them.

Second, if there is unemployment in the profit-oriented sector (that is, the number of type-0 agents is higher than the number of profit-oriented firms) then the supply of motivated agents is determined by their unemployment payoff. The bonuses and optimal levels of effort are

$$w^{0*}\left(x_G^0\right) = \frac{1}{2}x_G^0 \text{ and } e^{0*} = \frac{1}{2}x_G^0$$
$$w^{M*}\left(x_G^M\right) = \frac{1}{2}\left(x_G^M - I^M\right) \text{ and } e^{M*} = \frac{1}{2}\left(x_G^M + I^M\right).$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

In this case, the existence of the market does not influence the levels of the bonuses. It only provides information on the nature of the matching between principals and agents.

REFERENCES

Aetna Insurance Co. (Aetna) (1867) *Aetna Guide to Fire Insurance for the Representatives of the Aetna Insurance Co*, Hartford, CT: Aetna.

Akerlof, G.A. (1982) "Labor Contracts as a Partial Gift Exchange," *Quarterly Journal of Economics* 97 (4), 543–569.

Alonso-Paulí, E. and Pérez-Castrillo, D. (2012) "Codes of Best Practice in Competitive Markets for Managers," *Economic Theory* 49, 113–141.

Aron, D.J. and Olivella, P. (1994) "Bonus and Penalty Schemes as Equilibrium Incentive Devices, with Application to Manufacturing Systems," *Journal of Law, Economics and Organization* 10 (1), 1–34.

Arrow, K.J. (1963) "Uncertainty and the Welfare Economics of Medical Care," *The American Economic Review* 53 (5), 941–973.

Arrow, K.J. (1968) "The Economics of Moral Hazard: Further Comment," *The American Economic Review* 58 (3), 537–539.

Arrow, K.J. (1971) "Insurance, Risk and Resource Allocation," in *Essays in the Theory of Risk Bearing*, Amsterdam: North-Holland.

Auster, S. (2013) "Asymmetric Awareness and Moral Hazard," *Games and Economic Behavior* 82, 503–521.

Baranchuk, N., MacDonald, G. and Yang, J. (2011) "The Economics of Super Managers," *Review of Financial Studies* 24 (10), 3321–3368.

Barros, F. and Macho-Stadler, I. (1988) "Competition for Managers and Market Efficiency," *Journal of Economics and Management Strategy* 7, 89–103.

Bartling, B. (2011) "Relative Performance or Team Evaluation? Optimal Contracts for Other-regarding Agents," *Journal of Economic Behavior and Organization* 79 (3), 183–193.

Becker, G.S. (1973) "A Theory of Matching: Part I," *Journal of Political Economy* 81, 813–846.

Bénabou, R. and Tirole, J. (2003) "Intrinsic and Extrinsic Motivation," *Review of Economics Studies* 70, 489–520.

Besley, T. and Ghatak, M. (2005) "Competition and Incentives with Motivated Agents," *The American Economic Review* 95, 617–636.

Bhattacharyya, S. and Lafontaine, F. (1995) "Double-sided Moral Hazard and the Nature of Share Contracts," *RAND Journal of Economics* 26, 761–781.

Bolton, B. and Dewatripont, M. (2005) *Contract Theory*, Cambridge, MA: MIT Press.

Bolton, G.E. and Ockenfels, A. (2000) "ERC: A Theory of Equity, Reciprocity, and Competition," *American Economic Review* 90, 166–193.

Carroll, G. (2015) "Robustness and Linear Contracts," *American Economic Review* 105 (2), 536–563.

Chaigneau, P. (2014) "Risk Aversion, Prudence and Compensation," *The European Journal of Finance* 21 (15), 1357–1373.

Chakraborty, A. and Citanna, A. (2005) "Occupational Choice, Incentives and Wealth Distribution," *Journal of Economic Theory* 122, 206–224.

Dam, K. (2015) "Job Assignment, Market Power and Managerial Incentives," *The Quarterly Review of Economics and Finance* 57, 222–233.

Dam, K. and Pérez-Castrillo, D. (2006) "The Principal–Agent Matching Market," *Frontiers in Economic Theory* 2 (1), Article 1.

Dam, K. and Ruiz-Pérez, D. (2012) "On the Existence of Share Contracts Under Limited Liability," *Economics Letters* 117, 552–555.

De la Rosa, L.E. (2011) "Overconfidence and Moral Hazard," *Games and Economic Behavior* 73, 429–451.

De Meza, D. and Webb, D.C. (2007) "Incentive Design Under Loss Aversion," *Journal of the European Economic Association* 5 (1), 66–92.

Demougin, D. and Fluet, C. (2006) "Group vs. Individual Performance Pay When Workers Are Envious," in D. Demougin and C. Shade (eds), *An Economic Perspective on Entrepreneurial Decision-Making*, Berlin: Duncker & Humblot.

Demski, J., Sappington, D. and Spiller, P. (1988) "Incentive Schemes with Multiple Agents and Bankruptcy Constraints," *Journal of Economic Theory* 44, 156–167.

Dufwenberg, M. and Kirchsteiger, G. (2004) "A Theory of Sequential Reciprocity," *Games and Economic Behavior* 47, 268–298.

Dur, R. and Glazer, A. (2008) "Optimal Incentive Contracts When Workers Envy Their Bosses," *Journal of Law Economics and Organization* 24 (1), 120–137.

Edmans, A., Gabaix, X. and Landier, A. (2009) "A Multiplicative Model of Optimal CEO Incentives in Market Equilibrium," *Review of Financial Studies* 22 (12), 4881–4917.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Englmaier, F. (2005) "Moral Hazard, Contracts and Social Preferences: A Survey," in *Psychology, Rationality and Economic Behavior*, Basingstoke, UK: Palgrave Macmillan, 125–139.
Englmaier, F. and Leider, S. (2012) "Contractual and Organizational Structure with Reciprocal Agents," *American Economic Journal: Microeconomics* 4 (2), 146–183.
Englmaier, F. and Wambach, A. (2010) "Optimal Incentive Contracts under Inequity Aversion," *Games and Economic Behavior* 69 (2), 312–328.
Eswaran, M. and Kotwal, A. (1985) "A Theory of Two-tier Labor Markets in Agrarian Economies," *The American Economic Review* 75 (1), 162–177.
Falk, A. and Fischbacher, U. (2006) "A Theory of Reciprocity," *Games and Economic Behavior* 54, 293–315.
Fehr, E. and Schmidt, K.M. (1999) "A Theory of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics* 114, 817–868.
Feltham, G.A. and Xie, J. (1994) "Performance Measure Congruity and Diversity in Multi-task Principal/Agent Relations," *The Accounting Review* 69 (3), 429–453.
Fleckinger, P. and Roux, N. (2012) "Collective versus Relative Incentives: The Agency Perspective", mimeo, Paris School of Economics.
Gale, D. and Shapley, L.S. (1962) "College Admissions and the Stability of Marriage," *American Mathematical Monthly* 69, 9–14.
Gervais, S., Heaton, J.B., and Odean, T. (2011) "Overconfidence, Compensation Contracts, and Capital Budgeting," *The Journal of Finance* 66 (5), 1735–1777.
Ghatak, M. and Karaivanov, A. (2014) "Contractual Structure in Agriculture with Endogenous Matching," *Journal of Development Economics* 110, 239–249.
Ghatak, M. and Pandey, P. (2000) "Contract Choice in Agriculture with Joint Moral Hazard in Effort and Risk," *Journal of Development Economics* 63, 303–326.
Goel, A.M. and Thakor, A.V. (2006) "Optimal Contracts When Agents Envy Each Other," Working Paper.
Grossman, S.J. and Hart, O.D. (1983) "An Analysis of the Principal–Agent Problem," *Econometrica* 51 (1), 7–46.
Guo, B. (2016) "Manager Replacement, Employee Protest and Corporate Control," *Economics of Governance*, 17 (3), 265–294.
Harris, M. and Raviv, A. (1979) "Optimal Incentive Contracts with Imperfect Information," *Journal of Economic Theory* 20, 231–259.
Herweg, F., Müller, D., and Weinschenk, P. (2010) "Binary Payment Schemes: Moral Hazard and Loss Aversion," *American Economic Review* 100 (5), 2451–2477.
Holmström, B. (1979) "Moral Hazard and Observability," *Bell Journal of Economics* 10, 74–91.
Holmström, B. (1982) "Moral Hazard in Teams," *Bell Journal of Economics* 13, 324–340.
Holmström, B. and Milgrom, P. (1987) "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica* 55 (2), 303–328.
Holmström, B. and Milgrom, P. (1991) "Multitask Principal–Agent Analyses: Incentive Contracts, Assets Ownership, and Job Design," *Journal of Law, Economics, and Organization* 7 (Sup.), 24–52.
Hongy, S., Serfes, K., and Thiele, V. (2012) "The Market for Venture Capital: Entry, Competition, and the Survival of Start-up Companies," SSRN Working Paper.
Innes, R.D. (1990) "Limited Liabilities and Incentive Contracting With Ex Ante Action Choices," *Journal of Economic Theory* 52, 45–67.
Itoh, H. (2004) "Moral Hazard and Other-regarding Preferences," *Japanese Economic Review* 55 (1), 18–45.
James, H.S. Jr. (2005) "Why Did You Do That? An Economic Examination of the Effect of Extrinsic Compensation on Intrinsic Motivation and Performance," *Journal of Economic Psychology* 26, 549–566.
Jewitt, I. (1988) "Justifying the First-order Approach to Principal–Agent Problems," *Econometrica* 56 (5), 1177–1190.
Jewitt, I., Kadan, O., and Swinkels, J.M. (2008) "Moral Hazard with Bounded Payments," *Journal of Economic Theory* 143, 59–82.
Kahneman, D. and Tversky, A. (1984) "Choices, Values, and Frames," *American Psychologist* 39 (4), 341–350.
Kaneko, M. (1982) "The Central Assignment Game and the Assignment Markets," *Journal of Mathematical Economics* 10, 205–232.
Kaya, A. and Vereshchagina, G. (2015) "Moral Hazard and Sorting in a Market for Partnerships," *Economic Theory* 60, 73–121.
Kim, S.K. (1997) "Limited Liability and Bonus Contracts," *Journal of Economics and Management Strategy* 6, 899–913.
Kirkegaard, R. (2014) "A Unifying Approach to Incentive Compatibility in Moral Hazard Problems," University of Guelph Working Paper.
Kirkegaard, R. (2015) "Contracting with Private Rewards," University of Guelph Working Paper.
Köszegi, B. (2014) "Behavioral Contract Theory," *Journal of Economic Literature* 52 (4), 1075–1118.
Köszegi, B. and Rabin, M. (2006) "A Model of Reference-dependent Preferences," *Quarterly Journal of Economics* 121 (4), 1133–1166.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Köszegi, B. and Rabin, M. (2007) "Reference-dependent Risk Preferences," *American Economic Review* 97 (4), 1047–1073.

Kreps, D.M. (1997) "Intrinsic Motivation and Extrinsic Incentives," *American Economic Review* 87 (2), 359–364.

Laffont, J.J. and Martimort, D. (2002) *The Theory of Incentives: the Principal–Agent Model*, Princeton, NJ: Princeton University Press.

Legros, P. and Newman, A. (2002) "Monotone Matching in Perfect and Imperfect Worlds," *Review of Economic Studies* 69, 925–942.

Legros, P. and Newman, A. (2007) "Beauty is a Beast, Frog is a Prince: Assortative Matching with Nontransferabilities," *Econometrica* 75, 1073–1102.

Lewis, T. (1980) "Bonuses and Penalties in Incentive Contracting," *The Bell Journal of Economics* 11, 291–301.

Li, F. and Ueda, M. (2009) "Why Do Reputable Agents Work for Safer Firms?" *Finance Research Letters* 6, 2–12.

Macho-Stadler, I. and Pérez-Castrillo, D. (1997) *An Introduction to the Economics of Information: Incentives and Contracts*, Oxford: Oxford University Press.

Macho-Stadler, I. and Pérez-Castrillo, D. (2014) "Copyright Licensing under Asymmetric Information," *Review of Economic Research on Copyright Issues* 11 (2), 1–26.

Macho-Stadler, I., Pérez-Castrillo, D., and Porteiro, N. (2014) "Coexistence of Long-term and Short-term Contracts," *Games and Economic Behavior* 86, 145–164.

Mas-Colell, A., Whinston, M.D., and Green, J.R. (1985) *Microeconomic Theory*, New York: Oxford University Press.

Matthews, S.A. (2001) "Renegotiating Moral Hazard Contracts Under Limited Liability and Monotonicity," *Journal of Economic Theory* 97, 1–29.

Menezes, C., Geiss, C. and Tressler, J. (1980) "Increasing Downside Risk," *American Economic Review* 70 (5), 921–932.

Milgrom, P.R. (1981) "Good News and Bad News: Representation Theorems and Applications," *The Bell Journal of Economics* 12, 380–391.

Mirrlees, J. (1975) "The Theory of Moral Hazard and Unobservable Behavior – Part I," Oxford University Working Paper.

Mirrlees, J. (1976) "The Optimal Structure of Incentives and Authority within an Organization," *The Bell Journal of Economics* 7, 105–131.

Mirrlees, J. (1999) "The Theory of Moral Hazard and Unobservable Behavior – Part I," *Review of Economic Studies* 66 (1), 3–21.

Mookherjee, D. (1984) "Optimal Incentive Schemes With Many Agents," *Review of Economic Studies* 51 (3), 433–446.

Murdock, K. (2002) "Intrinsic Motivation and Optimal Incentive Contracts," *Rand Journal of Economics* 33 (4), 650–671.

Oyer, P. (2000) "A Theory of Sales Quotas with Limited Liability and Rent Sharing," *Journal of Labor Economics* 18, 405–426.

Park, E.-S. (1995) "Incentive Contracting Under Limited Liability," *Journal of Economics and Management Strategy* 4, 477–490.

Pauly, M.V. (1968) "The Economics of Moral Hazard: Comment," *The American Economic Review* 58 (3), 531–537.

Poblete, J. and Spulber, D. (2012) "The Form of Incentive Contracts: Agency with Moral Hazard, Risk Neutrality, and Limited Liability," *Rand Journal of Economics* 43, 215–234.

Quérou, N., Soubeyran, A., and Soubeyran, R. (2015) "Moral Hazard and Capability," Working Paper, Université de Montpellier.

Rabin, M. (1993) "Incorporating Fairness into Game Theory and Economics," *American Economic Review* 83 (5), 1281–1302.

Rabin, M. (2002) "A Perspective on Psychology and Economics," *European Economic Review* 46, 657–685.

Rey-Biel, P. (2008) "Inequity Aversion and Team Incentives," *The Scandinavian Journal of Economics* 110 (2), 297–320.

Rogerson, W.R. (1985) "The First-order Approach to Principal–Agent Problems," *Econometrica* 53, 1357–1367.

Romano, R.E. (1994) "Double Moral Hazard and Resale Price Maintenance," *RAND Journal of Economics* 25, 455–466.

Ross, S. (1973) "The Economic Theory of Agency: The Principal's Problem," *American Economic Review* 63, 134–139.

Roth, A. and Sotomayor, M. (1990) *Two-sided Matching: A Study in Game-theoretic Modeling and Analysis*, Econometric Society Monographs, New York: Cambridge University Press.

Salanié, B. (1997) *The Economics of Contracts*, Cambridge, MA: MIT Press.

Santos-Pinto, L. (2008) "Positive Self-image and Incentives in Organisations," *The Economic Journal* 118 (531), 1315–1332.

Sappington, D. (1983) "Limited Liability Contracts Between Principal and Agent," *Journal of Economic Theory* 29, 1–21.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Segal, U. and Sobel, J. (2007) "Tit for Tat: Foundations of Preferences for Reciprocity in Strategic Settings," *Journal of Economic Theory* 136 (1), 197–216.
Serfes, K. (2005) "Risk Sharing vs. Incentives: Contract Design under Two-sided Heterogeneity," *Economics Letters* 88, 343–349.
Serfes, K. (2008) "Endogeneous Matching in a Market with Heterogeneous Principals and Agents," *International Journal of Game Theory* 36, 587–619.
Shapley, L.S. and Shubik, M. (1972) "The Assignment Game I: The Core," *International Journal of Game Theory* 1, 111–130.
Shavell, S. (1979a) "On Moral Hazard and Insurance," *Quarterly Journal of Economics* 93, 541–562.
Shavell, S. (1979b) "Risk Sharing and Incentives in the Principal and Agent Relationship," *The Bell Journal of Economics* 10, 55–73.
Sinclair-Desgagné, B. and Spaeter, S. (2013) "The Prudent Principal," GATE Working Paper.
Sobel, J. (2005) "Interdependent Preferences and Reciprocity," *Journal of Economic Literature* XLIII, 392–436.
Spence, M. and Zeckhauser, R. (1971) "Insurance, Information, and Individual Action," *American Economic Review* 61, 380–387.
Stiglitz, J.E. (1974) "Incentives and Risk Sharing in Sharecropping," *Review of Economic Studies* 61, 219–256.
Von Thadden, E.L. and Zhao, X. (2012) "Incentives for Unaware Agents," *Review of Economic Studies* 79 (3), 1151–1174.
Von Thadden, E.L. and Zhao, X. (2014) "Multi-task agency with unawareness," *Theory and Decision* 77 (2), 197–222.
Zeckhauser, R. (1970) "Medical Insurance: A Case Study of the Trade-off Between Risk Spreading and Appropriate Incentives," *Journal of Economic Theory* 2, 10–26.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

17. Learning in markets

*Amparo Urbano**

The desire to acquire economically valuable information provides a powerful explanation for many empirically observed economic phenomena. This chapter surveys the phenomena of market learning and experimentation – or active market learning – in the context of dynamic models that incorporate a Bayesian expectation revision mechanism. There is an extensive literature on this topic. This chapter only surveys some representative papers, without demerit of many others

1 THE BEGINNING: ONE-SIDE, ONE-AGENT EXPERIMENTATION

1.1 The Ignorant Monopolist

There was a time when economists tended to assume that firms have complete demand information or had an exact knowledge of the stochastic knowledge generating demand. This viewpoint changed in the 1970s, in which explanations were given as to how the firm comes by this information.

In the following cited papers, demand information is generated endogenously through the analysis of a monopolist who wishes to learn about the demand process it faces. It may do so by *experimenting* with its decision variable – by adjusting its price or quantity away from the myopically optimal level – and observing the resulting quantity or price. Experimentation seeks to increase the informative level of observations. While such price or quantity experimentation reduces expected profits in the current period, the loss can possibly be recouped in subsequent periods through use of the improved information.

Two approaches have emerged regarding the question of experimentation in the face of a random demand curve featuring unknown parameters. One involves formulating a long - (finite or infinite) horizon model in which attention turns to the limiting expectation of the firm. Examples include Rothschild (1974), Prescott (1972), Kihlstrom, Mirman and Postlewaite (1984), Easley and Kiefer (1988), Trefler (1993), Keller and Rady (1999), among others. A second approach is restricted to two-period models and examines how the opportunity for experimentation affects the period-one output or consumption level. Examples include Grossman, Kihlstrom and Mirman (1977), Mirman, Samuelson and Urbano (1993a), Creane (1994), Alepuz and Urbano (1995), among others.

* The author acknowledges financial support from both the Spanish Ministry of Economics, Industry and Competitiveness under Projects ECO2013-46550-R and ECO2016-75575-R and the Generalitat Valenciana under the Excellence Programs Prometeo 2014II/054 and ISIC 2012/021. I also wish to thank R. Moner for very helpful comments and suggestions.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

1.2 Monopolists' Beliefs (Posteriors) and Optimal Actions – Convergence Over Time

There is a body of literature on whether the monopolist's beliefs (posteriors) and optimal actions converge over time. These works are statistical decision models of an agent trying to optimize her decision while improving her information at the same time. The simplest market in which the monopolist operates is the "reservation price" market. Imagine that the monopolist operates a shop in which one customer arrives in each period and decides whether or not to make a purchase. There is only one indivisible good for sale and each customer buys at most one unit. Due to unobservable customer heterogeneity the monopolist views customer purchase decisions as stochastic. At a given price some customers buy, others do not. Thus, customer purchases follow a binomial process with probability of purchase inversely related to price. It is this inverse relationship that the monopolist seeks to learn through price experimentation.

Rothschild (1974) proposed the binomial reservation price model, in which a single firm is facing a market with unknown demand. The true market demand is given by a specific probability distribution over consumer valuations. However, the firm initially has a prior probability over several possible market demands. A firm that does not know the consequences of charging a particular price has an obvious way of finding it out. It may charge the price and observe the result. If the firm charges price p_i, then the true probability of a sale is Γ_i. Again, if the firm knew parameters Γ_i, it would simply choose p_i to maximize $\Gamma_i r_i$, where $r_i = p_i - c$, is the profit from making a sale at price p_i, but since the store does not know the values of the Γ_i its problem is more complex. However, the firm can, if it chooses to, learn the value of any particular parameter Γ_i. For Γ_i is simply the probability of success in a single binomial trial. If the trial is repeated infinitely often, an observer will, according to the strong law of large numbers, be able to estimate Γ_i exactly (that is, with probability of 1). The firm could come to learn all the probabilities Γ_i, simply by choosing a strategy that involved charging each of the prices an infinite number of times – such as, for instance, by playing them in turn.

However, such experimental determination of its demand curve is costly. Customers turned away by a price that is objectively too high may not return, and sales made at prices that are too low represent losses that cannot be recouped. Formulating the firm's optimal strategy consists of finding some way of weighing the value of new information from charging a particular price against the cost of not charging the price that present information indicates is most profitable. This is one of the classes of problems under the general heading of two-armed bandit problems. In this class of problems, the gambler has to decide which arm of two different slot machines to play in a sequence of trials so as to maximize her reward. This classical problem has received much attention because it provides the trade-off between exploration (trying out each arm to find the best one) and exploitation (playing the arm believed to give the best payoff). The problem for the firm is to find an optimal sequence of prices to learn more about the true demand while maximizing its expected discounted profits. An optimal policy for a firm involves putting a value on the potential gain in information from charging a precise price. This can be done by setting up a dynamic programming equation in which the state variables describe the firm's present beliefs about the demand function of its customers Rothschild's (1974) results are obtained only for a simple and rather special example: customers buy at most a single unit of the commodity and there are only two prices that the firm can charge.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Consider, then, two potential prices to be charged. Each trial on the i-th price yields payoff r_i with probability Γ_i, and nothing with probability $(1 - \Gamma_i)$. The firm does not know the parameters with certainty. It decides which price to play at each stage after consulting its prior beliefs about the parameter Γ_i and examining the record of successes and failures on the prices so far. Sufficient statistics to represent the information in the sample are μ_i and δ_i, defined by: $\delta_i = \frac{1}{1+T_i}$ and $\mu_i = \frac{N_i}{1+T_i}$ where T_i is the number of trials and N_i the number of successes. These definitions imply particularly simple rules for updating (μ_i, δ_i). If price i is charged, δ_i becomes $\frac{\delta_i}{1+\delta_i}$. If there is a success on price i, μ_i becomes $s_i = \frac{(\mu_i+\delta_i)}{1+\delta_i}$, while if there is failure μ_i becomes $f_i = \frac{\mu_i}{1+\delta_i}$. If a success on price 1 is observed, (μ, ρ) is updated to $s_1(\mu_1, \mu_2, \delta_1, \delta_2) = \left(s(\mu_1), \mu_2, \frac{\delta_1}{1+\delta_1}, \delta_2\right)$, with $s_2(\mu, \rho), f_1(\mu, \rho)$ and $f_2(\mu, \rho)$ defined in the obvious manner. The firm's prior beliefs about the parameters are summarized by a prior density function $g(\gamma_1, \gamma_2)$. If a firm has experience (μ, ρ) it will use Bayes' rule to update its prior beliefs from $g(\gamma_1, \gamma_2)$ to $h(\gamma_1, \gamma_2, \mu, \rho)$, the probability density proportional to the probabilities of success on the two prices.

Define $\lambda_i(\mu, \rho) = \int_0^1 \int_0^1 \gamma_1 h(\gamma_1, \gamma_2, \mu, \rho)\, d\gamma_1 d\gamma_2$, as the posterior mean of the firm's beliefs about the value of Γ_i, given the sample information (μ, ρ) and the prior density g. The dynamic programming equation is a continuous real-valued function, $V(\mu, \rho)$, equal to the maximum expected discounted profits that the firm can make if the present state of information is described by (μ, ρ) and satisfies the basic functional equation: $V(\mu, \rho) = Max_i W_i(\mu, \rho)$, with

$$W_i(\mu, \rho) = \lambda_i(\mu, \rho)\, q_i + \delta\left[\lambda_i(\mu, \rho)\, V(s_i(\mu, \rho)) + (1 - \lambda_i(\mu, \rho))\, V(f_i(\mu, \rho))\right].$$

If the discount rate is positive, almost all firms following optimal strategies will, after an initial period of sampling, settle on one price and play it in preference to all others. However, the price chosen will not necessarily be the correct one. With positive probability a firm pursuing an optimal strategy will play the most favorable price a finite number of times while it plays a less attractive price infinitely often. An heuristic argument is as follows. Suppose that the probability of a payoff on one of the prices, say the second one, is known with certainty. The state of the firm's information is described by its estimate of the probability of a payoff on the first price. Its choice of price at any point in time is entirely determined by this estimate. When the first price is played, the firm receives, in addition to a random payoff, information that allows it to revise its estimate of the probability of a success on a first price. When it plays the second price, it receives only a payoff (if it is lucky) or nothing (if it is not). The outcome cannot affect its estimate of the probability of success on the second price since it is assumed this is known with subjective certainty. It also cannot affect its estimate of the probability of a payoff on the first price. Thus, if the state of the firm's information is such that the optimal strategy dictates that second price is played, it is unchanged after the second price been played again. If the price whose payoff probability is known is ever played, it will be played forever more. Rothschild (1974) shows that the optimal policy is in the form of a stopping rule, answering the question: when is information on the unknown arm so disappointing that play on it should be suspended? *Ex ante* optimal pricing rules may well end up using prices that are *ex post* suboptimal if the true distribution were to be known.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Prescott (1972) establishes the direction of experimentation for the case of an additive, normally distributed error term and a normal prior on the unknown demand slope. He sets up a finite-horizon problem with a single control and a single unknown parameter to analyze a model of an uncertain demand with unknown slope coefficient. The process assumed to generate the data is a simple regression model, namely, $y_t = \beta x_t + u_t$, for $t = 1, 2, \ldots, T$, where y_t is the tth observation of the dependent variable, x_t is the tth value of a control variable, β is an unknown scalar parameter, and u_t is the tth unobserved random error term. The u_ts are normally and independently distributed, each with mean zero and common known variance, which without loss of generality are taken to be one. Further, assume that the prior knowledge at the time x_1 can be represented by a normal distribution with mean m_1 and precision (the reciprocal of the variance) h_1. Then, the distribution on the unknown parameter at the time of the tth decision is normal, with precision satisfying the difference equation $h_t = h_{t-1} + x_{t-1}^2$ and mean satisfying $m_t = (m_{t-1}h_{t-1} + x_{t-1}y_{t-1})/h_t$. Given initial prior $N(m_1, h_1)$ on β, the control problem is to select the x_t sequentially so as to minimize the sum of the expected losses $E\,[\int_{t=1}^{T} q_t\,(y_t)]$, where the q_t are the non-negative losses.

Let $f_t\,(m_t, h_t) = \inf E\,[\int_{t=1}^{T} q_i\,(y_i)/m_t, h_t]$ for $i = t, \ldots, T$. This is the infimum for the sum of the expected losses for periods t through T inclusive given prior $N(m_t, h_t)$ on β at time t. As the initial prior has been assumed normal, the prior at the time of the tth decision will necessarily be normal. By backward induction,

$$f_t\,(m_t, h_t) = \min E\left[q_t\,(y_t) + f_{t+1}\,(m_{t+1}, h_{t+1})/x_t, m_t, h_t\right], \text{ for } t = 1, \ldots, T \text{ with } f_{T+1} = 0.$$

Because x is constrained to a compact set, the infimum is obtained. The first term in the expectation measures the effect of decision x_t upon the loss in the current period, while the second measures the effect upon future losses given optimal future behavior. The larger x_t^2, the more precise will be the future knowledge of β as the precision (variance) of the posteriori will be larger (smaller). Current decisions affect future as well as current losses so there will be a trade-off between stabilization and experimentation. In particular, the larger x_t^2, the more informative is the experiment. This result implies that the optimal decision will be larger in absolute value than the one that minimizes expected loss in the current period, so the optimal policy is to sacrifice some stability in order to gain information. Grossman, Kihlstrom and Mirman (1977) show that the Prescott result holds when the normality restriction on the prior is relaxed.

Trefler (1993) generalizes the work by Prescott (1972) and Grossman et al. (1977) by taking away the normality assumption and by considering an infinite-horizon problem with multiple controls and unknown parameters, in which Bayesian learning provides the link between periods. In each period the firm takes an action (price), observes the outcome (demand) and updates its beliefs about the unknown parameters of the stochastic process generating outcomes. Namely, a dynamic programming problem with a Markovian structure, where the state space is the space of prior beliefs, and the transition from the state in a period to the next one follows Bayes' rule. More specifically, in each period t the decision-maker chooses an action $x_t \in X$ and observes an outcome $y_t \in Y$. The probability of outcome y_t depends both on x_t and on an unknown parameter $\theta \in \Theta$. Assume that x_t and θ are conformable for matrix multiplication and let $\vartheta(./x_t, \theta)$ be a probability measure over outcomes. The decision-maker begins with a prior, $\mu_0 \in P(\Theta)$ about the unknown parameter θ. In period t her beliefs about θ

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

are given by μ_t. At the end of period t, having chosen x_t and observed y_t, she uses Bayes' rule to update her beliefs about θ to μ_{t+1} and receives a reward $r(x_t, y_t, \theta)$. The salient features of the setup are that both x_t and θ enter $\vartheta(./x_t, \theta)$ and that the only intertemporal link is Bayesian learning. An informal discussion of the Markovian structure of the dynamic programming problem is as follows. The state space is the space of prior beliefs, $P(\Theta)$. The transition from the state in period t (the prior μ_t) to the state in period $t+1$ (the posterior, μ_{t+1}) follows Bayes' rule. In choosing x_t, the decision-maker must take into account its effect on future beliefs, μ_{t+1}. At the time x_t is chosen, however, y_t and hence μ_{t+1} are unknown so that the decision-maker must calculate the distribution of future beliefs μ_{t+1}, across the values that y_t may take. This forms the transition probability from μ_t to μ_{t+1}, which is described by the measure $q(d\mu_{t+1}/\mu_t, x_t)$. To illustrate the use of q, consider the value function for this dynamic programming problem, $V(\mu_t)$. The expectation of $V(\mu_{t+1})$, given the information available at t, is

$$\int_{P(\Theta)} V(\mu_{t+1})q(d\mu_{t+1}/\mu_t, x_t),$$

or more simply, $\int_{P(\Theta)} V(\tilde{\mu})\, dq(\mu_t, x_t)$. The author establishes the existence of a value function $V(\mu_0)$ and an optimal sequence of Markov actions. The decision-maker's objective is to maximize the expected discounted reward, $E\left[\sum_{t=0}^{\infty} \delta^t r(x_t, y_t, \theta)\right]$. To make clear the dependence of this expectation on μ_t, define

$$u(x_t, \mu_t) = \int_{\Theta.} \int_Y r(x_t, y_t, \theta)\vartheta(dy_t/x_t, \theta)\, \mu_t(d\theta).$$

Then, the problem is to maximize $E\left[\sum_{t=0}^{\infty} \delta^t u(x_t, \mu_t)\right]$, where μ_t evolves according to Bayes' rule. Using q and μ, the problem may be reduced to a standard dynamic programming problem with action space X and state space $P(\Theta)$. For $\mu \in P(\Theta)$ define

$$V(\mu) = Max_{x \in X} \left\{ u(x, \mu) + \delta \int_{P(\Theta)} V(\tilde{\mu})\, dq(\mu, x) \right\},$$

where $0 < \delta < 1$ is the discount factor. If there is a function V satisfying the above equation, then it gives the value of the problem to a decision-maker with prior beliefs μ who behaves optimally. Define the expected value of information as $I(x; \mu, r) = \int_{P(\Theta)} V(\tilde{\mu})\, dq(\mu, x) - V(\mu)$. Trefler (1993) addresses two questions: When does one action result in larger expected value of information than another? What implications do endogenous information and learning have for the sequence of optimal actions; that is, what is the direction of experimentation? Trefler's results show the dependence of $I(x; \mu, r)$ on x for the reservation price market, posted price market, and normal models. In the former customer purchases follow a binomial process with probability of purchase inversely related to price, while in the post–price market the demand process is Poisson. The main findings of the paper are a characterization of the expected value of information and the direction of experimentation in the above-mentioned markets. These results are related to the literature on the expected value of information and Blackwell's comparison of experiments (see Blackwell, 1953, and DeGroot, 1962).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

The experimentation literature has by and large focused on broadly defined bandit models, where *stage payoffs are also random signals*, so that information gathering incurs an endogenous opportunity cost. Rustichini and Wolinsky (1995) study the problem of a monopoly that is uncertain about the demand it faces and learns about it over time through its pricing experience. The demand for the monopoly's product varies stochastically over time. Time is divided into discrete periods labeled $t = 1, 2, \ldots$. In each period the monopoly faces a unit demand with reservation price d_t. At the beginning of each period, before it knows d_t, the monopoly quotes a price p_t. Thus, if $p_t \leq d_t$, the monopoly will sell a unit, and if $p_t > d_t$, it will sell nothing. Let I_t record whether or not there has been a sale at period t. That is, $I_t = 1$ if $p_t \leq d_t$ and 0 otherwise. The maximum demand price, d_t, can assume two values: 1 and $D > 1$. It follows a Markov process with transition probabilities,

$$Prob\left[d_{t+1} = 1/d_t = D\right] = Prob\left[d_{t+1} = D/d_t = 1\right] = \alpha.$$

Let w_t denote the probability with which the monopoly believes that $d_t = 1$. Thus, w_1 is the prior and subsequently it evolves as follows: $w_{t+1} = \alpha$ if $1 < p_t \leq D$ and $I_t = 1$; $w_{t+1} = (1-\alpha)$ if $1 < p_t \leq D$ and $I_t = 0$, and $w_{t+1} = (1-\alpha)w_t + \alpha(1-w_t)$ if $p_t = 1$. Given a price sequence $\{p_t\}$ and a sequence of demand realizations $\{d_t\}$, the monopoly's discounted profit is $\sum \delta^t p_t I_t$, where $\delta < 1$ is the discount factor. At the beginning of period t the monopoly knows the history $h_t = (p_1, I_1), \ldots, (p_{t-1}, I_{t-1})$. Its problem is to choose a pricing policy $p_t(h_t)$ so as to maximize $E\left[\sum \delta^t p_t I_t\right]$. The authors use a two-armed bandit framework to study monopoly pricing when the buyers' reservation value changes randomly. In particular, the optimal policy is characterized by a cutoff belief W. If $w_t \leq W$, then $p_t = D$. If $w_t > W$, then $p_t = 1$. Thus, given δ and α, there are three types of optimal policy, depending on the relative size of D. When D is sufficiently small or sufficiently large, the monopoly will always quote the same price, 1 or D respectively. When D is in the intermediate range, the optimal policy involves price changes whose frequency depends on N, where N is the smallest non-negative integer such that, if $w_t = (1-\alpha)$ and $p_t = \ldots = p_{t+N-1} = 1$, then $w_{N+t} \leq W$. That is, N is the number of times that the seller quotes $p = 1$ after a price offer $p_{t-1} = D$ was rejected. In implementing the optimal policy, the monopoly will make two kinds of error. In some periods it will charge $p_t = 1$ when $d_t = D$, while in others it will charge $p_t = D$ when $d_t = 1$. The stationary probabilities of these two types of error capture the frequency with which these errors are made in the long run, and hence provide some measure of the extent of learning associated with the optimal policy. Rustichini and Wolinsky (1995) focus on non-negligible pricing errors even when the frequency of change is negligible. For certain parameter combinations, learning will cease completely even though the state keeps changing.

Keller and Rady (1999) study optimal experimentation by a monopolist who faces an unknown demand curve subject to random changes, and who maximizes profits over an infinite horizon in *continuous time*. Their model introduces Brownian noise and relies on filtering techniques for belief updating. The monopolist knows the slope and intercept of each demand curve and the transition probabilities, but it does not know which demand curve it faces at any given time. There are two possible states, $k = 0$ or 1, each characterized by a linear demand curve and the transitions between these states are governed by a Markov process. Specifically, in state k, the expected per-period demand curve (expected price as a function of quantity) is $p = \alpha_k - \beta_k q$, where α_k and β_k are positive constants. At each instant, it chooses from a given interval of feasible quantities, and observes a price that is the "true"

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

price (derived from the prevailing demand curve) plus noise. The state changes according to a continuous time Markov process with some transition probabilities. In particular,

$$Pr\,(k_s = k,\ \forall s \in [t, t + \Delta t]/k_t = k) = \exp(-\alpha_k \Delta t).$$

At each time t, the monopolist chooses an output level q_k, from an exogenously given interval $Q = [q_{min}, q_{max}]$ of feasible quantities. The resulting increment in total revenue is $dR_t = q_t[(\alpha_{kt} - \beta_{kt}q_t)\,dt + \sigma dZ_t]$, where Z is a standard Wiener process independent of the process k, and $\sigma > 0$ is a constant known to the monopolist. Thus, $dR_t = q_t dP_t$, where dP_t is the increment of a cumulative price process P given by $dP_t = (\alpha_{kt} - \beta_{kt}q_t)\,dt + \sigma dZ_t$. The monopolist derives all its information about the state of demand from observing this price process. The monopolist's initial belief about the state of demand is characterized by π, its subjective probability that $k_0 = 1$. Given this belief, its objective is to choose $\boldsymbol{q} = \{q_t\}$ so as to maximize:

$$u^{\boldsymbol{q}}(\pi) = E_\pi \left[\int_0^\infty r\, e^{-rt} dR_t \right] = E_\pi \left[\int_0^\infty r\, e^{-rt} q_t\,(\alpha_{kt} - \beta_{kt}q_t)\,dt \right],$$

where $r > 0$ is the monopolist's discount rate. Following a production strategy and observing the associated price process, the monopolist updates its beliefs about the state of demand in a Bayesian fashion.

Given a strategy q, beliefs evolve according to the (filtering) equation $d\pi_t = \lambda\,(\pi_t)\,dt + \sum(\pi_t,\, q_t)dZ_t^{\boldsymbol{q}}$, where $\lambda\,(\pi_t) = (1 - \pi)\,\lambda_0 + \pi\lambda_1$, $\sum(\pi,\, q) = \sigma^{-1}\pi\,(1 - \pi)\,(\Delta\alpha - \Delta\beta q)$, ($\Delta\alpha$ and a $\Delta\beta$ being the difference in intercepts and in slopes, respectively, between the two expected demand curves) and Z^q is a Wiener process with respect to the agent's information sets. In other words, the change in beliefs $d\pi_t$ has mean $\lambda\,(\pi_t)\,dt$ and variance $\sum^2(\pi_t, q_t)$. The monopolist's value function and the Bellman equation for its decision problem is defined as $u^*(\pi) = sup_q u^{\boldsymbol{q}}(\pi)$ for π in [0, 1]:

$$u^*\,(\pi) = max_{q\epsilon Q} \left\{ rE_\pi \left[q\,(\alpha_k - \beta_k q) \right] dt + e^{-rdt} E_\pi \left[u^*(\pi + d\pi) \right] \right\}.$$

The Bellman equation is the main tool for constructing optimal strategies, which will in fact be stationary Markov strategies. The monopolist can increase the information content of the price signal by moving away from the confounding quantity, that is, the quantity at which the two demand curves intersect; setting the confounding quantity itself leads to a completely uninformative signal. The authors show that there are two qualitatively very different regimes, determined by the discount rate and the intensities of demand curve switching, and the dependence of the optimal policy on these parameters is discontinuous. One regime is characterized by extreme experimentation and good tracking of the prevailing demand curve, the other by moderate experimentation and poor tracking. Moreover, in the latter regime the agent eventually becomes "trapped" into taking actions in a strict subset of the feasible set. The authors build upon several strands of the literature on optimal Bayesian learning. A number of authors have identified situations where it is optimal to experiment, and have characterized the agent's strategy as a function of his beliefs. Examples include Prescott (1972), Grossman et al. (1977), Mirman, Samuelson and Urbano (1993a) and Trefler (1993).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

These papers do not consider confounding actions, so the different experimentation regimes described in Keller and Rady (1999) do not arise.

Moscarini and Smith (2001) consider a Bayesian formulation of sequential analysis. A given decision-maker is uncertain about a payoff-relevant state of the world, and before deciding, can buy multiple independence and independently distributed (i.i.d.) informative signals at constant marginal cost. The decision-maker should then purchase one at a time, and act when sufficiently convinced of one state. Thus, the authors introduce and explore a continuous time model of sequential experimentation with explicit information purchases. Their general finding is that the optimal experimentation level grows in the Bellman value prior to stopping and acting. This monotonicity is critical to the analysis, and admits a concrete economic intuition – there are two decisions at each instant: stop or experiment, and then at what level. The driving features of the model are impatience and an increasing and strictly convex cost function of within-period experimentation. These two assumptions yield some robust predictions: experimentation intensity grows with a project's expected payoff. Among falsifiable implications, the authors establish an upward secular drift in the experimentation level for not too convex cost functions. They explore an experimentation problem, inspired by the statistical literature on sequential hypotheses testing, that is different in two key respects: first, information is explicitly costly, since the state-independent information costs are known and so uninformative; second, there is eventual stopping, so that delay cost drives all the results. The authors characterize the optimal sample size given convex costs and discounting.

1.3 Two-period Models

In two-period models, the monopoly may find it profitable to adjust its first-period quantity away from the myopically optimal level in order to increase the informativeness of its first-period price observation and hence increase second-period profits. The firm can thus collect information in the first period, but only at the cost of sacrificing period-one profits. The amount of information collected is endogenously determined by the firm's first-period action.

Mirman, Samuelson and Urbano (1993a, MSU thereafter), provide non-parametric results on the direction of experimentation for a two-period problem in which the unknown parameter takes on only two values. They examine a quantity-setting firm and focus on the effect of experimentation on first-period output levels. Namely, the inverse monopolist's demand is $P = g(\gamma, Q) + \varepsilon$, where P is price, Q is quantity, $\gamma \in \{\underline{\gamma}, \overline{\gamma}\}$ is a parameter unknown to the monopolist, and ε is a random variable. The prior probability that $\gamma = \overline{\gamma}$ is ρ_0. The random variable ε is characterized by a density $f(\varepsilon)$ that has expected value zero and is continuously differentiable on the real line. Also, $f(\varepsilon)$ satisfies the monotone likelihood ratio property (MLRP): $f'(\varepsilon)/f(\varepsilon)$ is a continuous and non-increasing function. The firm chooses a quantity Q in period one and observes the price $P = g(\gamma, Q) + \varepsilon$. Because ε is random, the firm may be unable to infer the value of γ from its price observation. The firm uses Bayes' rule to construct a posterior probability, denoted ρ, that $\gamma = \overline{\gamma}$. Then, it chooses its period-two quantity Q_2 and receives price $P_2 = g(\gamma, Q_2) + \varepsilon$. Let $V(\rho)$ be the maximized value of period-two expected profits as a function of the (posterior) probability ρ: $V(\rho) = max_{Q2}(\rho Q_2 g(\overline{\gamma}, Q_2) + (1 - \rho) Q_2 g\left(\underline{\gamma}, Q_2\right)$. In the first period, the firm's problem is to find Q^E such that

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

$$Q^E \in argmax_Q \left\{ \rho_0 Qg\left(\overline{\gamma}, Q\right) + \left(1-\rho_0\right) Qg\left(\underline{\gamma}, Q\right) + E_P V(\rho\left(P, Q\right)) \right\},$$

where ρ is calculated as a function of P and Q via Bayes' rule, or

$$\rho\left(P, Q\right) = \frac{f(P - g\left(\overline{\gamma}, Q\right))\rho_0}{f(P - g\left(\overline{\gamma}, Q\right))\rho_0 + f(P - g\left(\underline{\gamma}, Q\right))(1-\rho_0)}.$$

Notice that for a given value of Q, the value of the posterior probability ρ is a random variable. The firm does not experiment if it sets a quantity equal to the myopically optimal output, $Q^{NE} = argmax_Q \left\{ \rho_0 Qg\left(\overline{\gamma}, Q\right) + \left(1-\rho_0\right) Qg\left(\underline{\gamma}, Q\right) \right\}$. In this case the period-one quantity is set so as to maximize period-one (myopic) profits and does not take account of the future. Given the above two maximization problems, the firm will experiment whenever $\{dE_P V(\rho\left(P, Q\right))/dQ\} \neq 0$. Some algebra shows that the conditions for experimentation are that (1) the information must be valuable ($V\left(\rho\right)$ strictly convex in ρ), and (2) adjustments in quantity must be capable of increasing the informational content of price ($dg\left(Q^{NE}, \overline{\gamma}\right)/dQ \neq dg\left(Q^{NE}, \underline{\gamma}\right)/dQ$). The authors also establish sufficient conditions for experimentation to lead to an increase or decrease in period-one quantity. Intuitively, the firm adjusts its period-one quantity to push the mean demand curves further apart. This spreads apart the distributions from which the random variable, price, might be drawn and makes price a more informative signal of the true distribution. MSU's results contrast with those derived for the case of uncertain utility in Grossman et al. (1977) and Kihlstrom et al. (1984). In particular, Grossman et al. (1977) suggest that experimentation does not occur in the case of a linear demand curve with uncertain intercept, because the random variable observed by the firm is unaffected by the firm's quantity, the firm cannot affect the informativeness of the variable it observes and no experimentation occurs. However, MSU's results suggest that this intuition is incomplete in two respects. First, cases arise such as linear demand curves that intersect on the horizontal axis, in which the random variable observed by the firm is affected by the firm's quantity but the firm does not experiment because the resulting information is not valuable. Second, even when information is valuable, the ability of the firm to profitably experiment depends not on a single distribution but on the relative positions of the distributions that correspond to the various possible parameter values. The key to experimentation is the ability to affect these relative positions by altering the firm's quantity. MSU then expand the model to allow the possibility that the firm may set either price or quantity, and can choose which of these to set; and establish conditions under which a firm prefers to be a price or quantity setter and show that quantity-setting and price-setting firms may choose to experiment in quite different ways.

Some generalizations of MSU's results are the following. Creane (1994) analyzes a version of MSU with noise heteroscedasticity and finds that the direction of information with respect to the choice variable can easily change under a general class of distributions. In particular, suppose that the inverse monopolist's demand is now $P = a + bQ + \varepsilon\phi\left(Q\right)$ with b unknown, ε is a random variable and ϕ is some known function. He shows that the expected value of information is increasing (decreasing) in Q, if $\phi - Q\left(d\phi/dQ\right) > 0 (< 0)$. The implication is that, when information is a function of the choice variable, the result of a model may be an outcome of the structure of the uncertainty and not of the existence of uncertainty. Alepuz and Urbano (1995) analyze the learning behavior of a risk-averse monopolist and find two

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

conflicting effects in the experimental behavior: a stronger preference for the ex post reduction in uncertainty, but ex ante the returns to information are more uncertain. Urbano (1992) generalizes some results of MSU to the case of *n* values of a monopolist's unknown demand parameter. She extends previous results by giving a more formal treatment to concepts such as the value of information and more informative signals. One application of this methodology is Avila-Baltuille, Caballero-Sanz and Urbano (1993), to a monopoly with uncertainty that concerns a new technology to be adopted by the firm. The paper examines a Bayesian learning model with intra-firm diffusion. The authors discover that the firm will not only practice output experimentation but also sampling experimentation, that is, the monopolist will increase both the output level and the rate of adoption of the unknown technology for experimental purposes with respect to the myopic choices. Moreover, there are marginal incentives to experiment created by diffusion that introduces complementarities in the learning behavior of the firm.

2 THE CONTINUATION: TWO SELLERS, ONE-SIDED EXPERIMENTATION

The value of information in oligopoly games has been the subject of intensive research. These studies typically assume either that firms transmit information through "certifiable/verifiable announcements" or that the signals that yield information to the firms are exogenously generated. In contrast, the following models endogenously determine the amount of information available to firms and analyze how learning behavior can modify the outcome of competition in a duopoly industry facing demand uncertainty.

Experimentation by a firm is the use of present actions to vary the amount of information available in the future. However, when other firms observe these actions they also affect their inferences about the same parameter and hence the underlying market competition. It is the purpose of this section to study experimentation under competition.

2.1 The Duopoly Case

Mirman, Samuelson and Schlee (1994) generalize the results of MSU to a duopoly model. Firms can draw inferences concerning the uncertain market demand from observation of their outputs and market price and may adjust their outputs away from myopically optimal levels to affect the informativeness of the market price. As in MSU's model, the two-period duopoly has the inverse market demand $P = g(\gamma, Q) + \varepsilon$, where P is market price, Q is industry output, $\gamma \in \{\underline{\gamma}, \overline{\gamma}\}$ is a parameter unknown to the monopolist, and ε is the realization of a random variable $\tilde{\varepsilon}$. The random variable $\tilde{\varepsilon}$ is characterized by a density $f(\varepsilon)$, with zero mean and continuously differentiable on the real line. Also, $f(\varepsilon)$ satisfies the monotone likelihood ratio property. The common knowledge prior probability that $\gamma = \overline{\gamma}$ is ρ_0. In period one, each duopolist i chooses a quantity Q_i and observes the price $P = g(\gamma, Q) + \varepsilon$, and quantities $Q = Q_1 + Q_2$, but not the realization of $\tilde{\varepsilon}$. Therefore, each firm revises its prior beliefs according to Bayes' rule to obtain the common posterior ρ. Since firms' quantities are observed, firms can manipulate the extent to which beliefs revision occurs rather than the direction in which beliefs are revised – as in signal-jamming models. The authors develop conditions and present examples under which the value of information is positive and negative and under which firms will increase or decrease quantity to manipulate information. Their

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

main contribution is to extend duopoly experimental behavior to mixed strategy equilibrium and to show some cases for which the net value of information for the duopoly is positive. However, they are unable to relate the duopoly experimental behavior with that of the monopoly in their general setting.

To answer the above question, Alepuz and Urbano (1999) analyze learning behavior by a Cournotian duopoly and consider a continuum of possible values for the slope of demand. Consider a two-period duopoly model. The firms produce a homogeneous product over the two periods. Inverse market demand in each period is given by $\tilde{P}_t = a - \tilde{\theta}(q_t^1 + q_t^2) + \tilde{\varepsilon}_t$, where $\tilde{P}_t$ is the price in period t, $t = 1, 2$, q_t^i is firm i's quantity in period t, $t = 1, 2$. Parameter $\tilde{\theta}$ is the fixed random slope and $\tilde{\varepsilon}_t$ is each period's random demand shock. It is assumed that each of these two random variables has full support on R and that they are independently and normally distributed. The value of θ is unknown to firms, but they a priori believe that $\tilde{\theta} \sim N(m, h)$ where $h = 1\sigma_\theta^2$ is $\tilde{\theta}'s$ precision. Given that $\tilde{\varepsilon}_t \sim N(0, \tau)$, $t = 1, 2$, and τ is $\tilde{\varepsilon}'$s precision, after observing P_1 and knowing Q_1, the firms' new belief about θ is that it is a realization of a random variable that is normally distributed according to $N(\hat{\theta}, \hat{h})$, where $\hat{\theta} = (mh + \tau Q_1(a - P_1)/h + \tau Q_1^2)$ and $\hat{h} = h + \tau Q_1^2$. Let $V\left(\hat{\theta}\right)$ be each firm's equilibrium expected profits in the second period. Then, each firm i's two-period expected profits as a function of first period outputs is:

$$\Pi^i\left(q_1^1, q_1^2\right) = \pi_1^i\left(q_1^1, q_1^2\right) + E\left(V\left(\hat{\theta}\right)\right),$$
$$\text{with } E\left(V\left(\hat{\theta}\right)\right) = \int_{-\infty}^{+\infty} V(\hat{\theta}(PQ_{11}, Q_1))f(a - P_1)dP_1,$$

where f is the density function of the random variable $a - \tilde{P}_1$. Strategic interaction adds a new effect into the analysis since the result of the experimenting behavior by one firm is observed by the rival. In such a setting and if information is valuable, is there any unilateral incentive to experiment? Does a monopoly experiment more or less than a duopoly? The main contribution of this paper is twofold. First, to generalize duopoly experimentation to the case of a continuum of possible demand curves, and to prove that the informativeness of the commonly observed signal increases with a firm's output. Second, to relate experimentation and market competition. The authors find out that what is relevant is the a priori uncertainty about the random demand. In particular, they show that if the random slope is sufficiently precise, which, in turn, makes the commonly observed market signal precise as well, then the monopoly will experiment more than the Cournotian duopoly. The intuition behind this result can be understood by noting that, under duopoly, firms face a strategic informational choice in the sense of how informative to make the publicly observed market signal. If initial beliefs about the random slope are precise, so is the market signal and then, posterior beliefs may become too accurate by experimentation. Hence, the rival will be better informed and a tougher competitor: too precise signals discourage experimentation by duopolists. Then, the incentives to learn by overproducing are smoothed down by the harder competitive conditions that may follow. This discourages experimentation by the duopolists and therefore the monopoly will experiment more; that is, it will produce much more away from its myopic output.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Extending Grossman et al.'s (1977) work to an oligopoly setting, it has been shown that the incentives to experiment and the implication of experimentation can be quite different. In a homogeneous-product setting Mirman et al. (1994) and Alepuz and Urbano (1999) found that experimentation can result in either a higher or lower price depending on the demand specification. The analysis of Aghion, Espinosa and Julien (1993), Alepuz and Urbano (1994) and Harrington (1995) finds that uncertainty over the degree of product differentiation impacts the amount of price dispersion. Aghion et al. (1993) assume that the degree of substitution between products is unknown to explain the phenomenon of price dispersion under oligopoly. They study this phenomenon and its dynamic evolution in the context of a Hotelling duopoly model and then extend the analysis to general demand functions and to N-firm oligopolies. The authors' results are driven by the specific way of modeling the error terms that affects both market demands and that, in turn, determine which specific market signal the firms want to make more informative. With symmetric demand this is accomplished by price dispersing. However, this result is not generally true – under both general expected demand and error terms specifications – unless firms' ability to make market signals more informative is the same for both in each market, like in the Hotelling model. Alepuz and Urbano (1994) allow for a more general modeling of the market demand random terms and for the general class of joint distribution function of the noises that satisfies the generalized strict monotone likelihood ratio property. The model clarifies the learning mechanism that operates in general learning duopoly models for the existence of price dispersion. Their results show that when firms experiment in all markets at the same time and have the same ability to make market signals more informative then, provided that products are substitutes, they will price-disperse in an attempt to increase the information content of these market signals. Hence a sampling effect may arise as the global outcome of market learning behavior. Harrington (1995) sets up a simple price-setting duopoly in which firms are uncertain about the degree of product differentiation. Firms learn by observing the difference in the quantities that are demanded in light of the differences in their prices. The author shows that the informativeness of the market experiment is increasing in the amount of price dispersion. Harrington's central finding is that the qualitative effect of experimentation varies with the type of markets. In markets with high substitutable products, firms create more price dispersion than is predicted by static profit maximization so as to create a more informative market experiment. In markets with highly differentiated products, firms compress their prices in order to reduce how much information is generated about product substitutability. Belleflamme and Bloch (2001) compare experimentation about product differentiation in a linear setting under four market structures: quantity-setting and price-setting monopoly, Cournot and Bertrand duopoly. Quantity-setting firms always experiment by raising their quantities and the monopolist experiments relatively more than the duopolist. A price-setting monopolist does not experiment. The value of information to Bertrand duopolists may be positive or negative depending on the degree of product differentiation. When information is valuable, price-setting duopolists experiment by lowering prices. A numerical example indicates that the intensity of experimentation is higher in a Cournot duopoly than in a Bertrand duopoly.

Alepuz and Urbano (2005), explore experimentation and learning in *asymmetric* duopoly markets with product differentiation and demand uncertainty. Market demands are given by $\widetilde{P}_1 = a - b_1Q_1 - cQ_2 + \widetilde{\varepsilon}_1$, and $\widetilde{P}_2 = a - b_2Q_2 - cQ_1 + \widetilde{\varepsilon}_2$, where $\widetilde{\varepsilon}_1$ and $\widetilde{\varepsilon}_2$ are uncorrelated random shocks of demand, whose joint distribution is characterized by a

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

continuous differentiable density $f(\varepsilon_1, \varepsilon_2)$ with zero means, full support in $\boldsymbol{R}^2$. The demand slopes b_1 and b_2 are unknown fixed parameters of the market demand, with $b_i \in \{\underline{b_i}, \overline{b_i}\}$. Let $\alpha_0 = \Pr\{b_1 = \overline{b_1}\}$ and $\beta_0 = \Pr\{b_2 = \overline{b_2}\}$ be the (common knowledge) initial probabilities. Parameter c, the degree of product substitution between the two markets, can be either positive (if products are substitutes) or negative (if products are complements). As usual, in the first period firms choose simultaneously quantities Q_{11}, Q_{21}, observe first-period prices, P_{11}, P_{21}, but not the realization of the random variables, and update (α_0, β_0) to obtain the common posteriors $(\alpha(P_{11}, P_{21}), \beta(P_{11}, P_{21}))$. Let $\bar{\varepsilon_1} = P_{11} - a + \overline{b_1}Q_{11} + cQ_{21}$, $\underline{\varepsilon_1} = P_{11} - a + \underline{b_i}Q_{11} + cQ_{21}$, etc. It is assumed that functions $\beta_0 f(\varepsilon_1, \bar{\varepsilon_2}) + (1-\beta_0) f\left(\varepsilon_1, \underline{\varepsilon_2}\right)$ and $\alpha_0 f(\bar{\varepsilon_1}, \varepsilon_2) + (1-\alpha_0) f\left(\underline{\varepsilon_1}, \varepsilon_2\right)$ satisfies the strict monotone likelihood property with respect to ε_1 and ε_2, respectively

Let $V_i(\alpha(P_{11}, P_{21}), \beta(P_{11}, P_{21}))$ be firm i's second-period equilibrium expected profits as a function of posterior beliefs on market 1 and 2. Then, firm i's expected profits for the two periods is:

$$\Pi_i(Q_{11}, Q_{21}) = \pi_{i1}(Q_{11}, Q_{21}) + E(V_i(\alpha, \beta),$$

where:

$$E(V_i(\alpha, \beta) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} V_i(\alpha(P_{11}, P_{21}), \beta(P_{11}, P_{21}))\, h(P_{11}, P_{21})\, dP_{11}dP_{21}, \text{ and}$$
$$h(P_{11}, P_{21}) = \alpha_0\beta_0 f(\bar{\varepsilon_1}, \bar{\varepsilon_2}) + \alpha_0(1-\beta_0) f\left(\bar{\varepsilon_1}, \underline{\varepsilon_2}\right) + (1-\alpha_0)\beta_0 f\left(\underline{\varepsilon_1}, \bar{\varepsilon_2}\right) + (1-\alpha_0)(1-\beta_0) f\left(\underline{\varepsilon_1}, \underline{\varepsilon_2}\right).$$

Here, in contrast with symmetric models, the value of information for each firm depends on the uncertainty of both markets. Thus, we may find situations where the information about a market may be valuable (for one firm), while the information about the other market may have not value at all. The authors define the concepts of strategic substitutability and strategic complementarity in information and show how both the mode of information competition and the transmission of information across markets affect duopoly experimentation. The information competition – as either information strategic substitutability or information strategic complementarity – prevailing in markets is introduced through the notion of "favorableness" of news in the sense of Milgrom (1981) and it may be related with product substitutability or complementarity. In particular, under Cournot competition and with asymmetric linear stochastic demand, the value of information from both markets is positive when products are market complements. On the other hand, a firm is only interested in information about its own market in case of product substitutability. In addition to the value of information, the feature of partial information transmission across markets will either encourage or discourage the production of information. The partial information transmission will depend on the correlation between the market signals (P_{11}, P_{21}). The authors allow for both positive and negative correlation between the market shocks, which determine, in turn, the respective correlation between the market signals.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

2.2 Experimentation, Information Sharing and Information Manipulation

Considerable research exists concerning the incentives for firms in strategic environments to commit to sharing information in the future. Interest in this issue is motivated by the prevalence of industry sharing agreements and related court and Federal Trade Commission cases (see, for instance, Vives, 1990). In the standard information-sharing model, each firm is to receive private information in the future and knows ex ante the precision of every firm's future information. Each firm can commit itself to reveal the information it will receive (an information-sharing agreement). After information is received and shared according to their prior commitments, the firms compete strategically (e.g., quantity competition). Despite the role information plays, these models do not explore the source of the new information. Instead it is assumed to be exogenous. The effects of information-sharing agreements on information production has been generally ignored in earlier research. When these effects are taken into account, some results can change.

Creane (1995) replaces the exogenous information assumption with an endogenous learning model. Each firm is not only able to enter into an information-sharing agreement, but can also choose how much private information to produce. Endogenizing information production allows an analysis of the interaction between the production and the sharing of information. Using results from the endogenous learning literature, the author models firms as Bayesian agents with unknown cost functions. Firms learn by observing the noisy cost realization from each production run. By varying output, a firm can learn more about its costs. This information production model is then merged with a standard information-sharing model. The author finds that agreements to exchange information affect the value and production of information. With unknown costs, a learning-by-doing-like effect also arises. These effects affect consumer welfare, the incentive to receive information, and the incentive to enter into information-sharing agreements. Information-sharing contracts may have negative future effects on firms through decreased information production. However, the decreased information production has the current benefit of softening competition, which induces information-sharing agreements under conditions contrary to previous results and vice versa. Contrary to previous work, being less informed does not guarantee a learning-by-doing advantage. Further, under quantity competition, the sharing of information increases information production by firms. Current and future consumer welfare increases with increased information production. Even though firms may value the sharing of a fixed amount of information, they may not want to receive information because of the current negative effect on the firm that an increase in information production by the rival can have.

On the other hand, signaling models have examined the transmission of private information through perfectly observed actions. The question is what happens when firms cannot observe their opponents' quantities and cannot draw precise inferences because market prices are a noisy function of quantities. Urbano (1993) examines this issue in a two-period Cournot duopoly, where each firm is uncertain about the value of some demand parameter and engages in experimentation, as in Alepuz and Urbano (1999) and Mirman et al. (1994). In addition, the actions of the firms are not observable by its rivals and, therefore, there is an informational interaction between firms even when the opportunity to experiment is not present. This interaction arises because it is possible for a firm to adjust its (unobserved) quantity in order to manipulate the non-trivial distribution of some variable observed by a rival. This phenomenon is known as *signal jamming* (see Mirman, Samuelson and Urbano, 1993b). In Urbano (1993),

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

firms may thus change their myopically optimal output in order to affect the information flow, which in turn influences both their own future decision and their rival's future decision. The basic result is that three forces shape the firm's information manipulation decision. One arises because of experimentation and the other because of signal jamming. A new force also arises from the interaction between signal jamming and experimentation. In particular, complementarities exist in the level of signal jamming and experimentation undertaken by firms.

Bernhardt and Taub (2015) characterize a duopoly buffeted by demand and cost shocks. Firms learn about shocks from common observation, private observation, and noisy price signals. Firms internalize how outputs affect a rival's signal, and hence output. They distinguish how the nature of information – public versus private – and of what firms learn about – common versus private values – affect equilibrium outcomes. Firm outputs weigh private information about private values by more than common values. Thus, prices contain more information about private-value shocks.

2.3 Learning and Strategic Pricing

Much of the existing literature on dynamic choice under uncertainty has focused on the case where a single decision-maker chooses sequentially among a fixed set of alternatives. In many economic situations the alternatives are supplied by a separated economic agent or group of agents. Bergemann and Välimäki (1996) and Felli and Harris (1996) are two early papers analyzing the impact of price competition on experimentation. The authors show that if there is only individual learning, the dynamic duopoly competition with vertically differentiated products can achieve efficiency.

Bergemann and Välimäki (1996) set up a simple dynamic equilibrium model of price formation under learning and uncertainty. In an infinite horizon model with price competition, a buyer chooses sequentially between products whose qualities are initially unknown to all parties in the model; the buyer does not know the underlying characteristic of the products, while the producers are uncertain about the buyers' tastes. Price competition between two firms, indexed by $i = 1, 2$, takes place in discrete time with an infinite horizon, $t = 0, 1, 2, \ldots$. The firms announce in each period their prices, p_i, simultaneously. The goods produced by the two firms differ only with respect to their (expected) quality. Firms have the same unit costs normalized to zero. The buyer has unit demand in each period. At time t, the buyer's expected valuation of a purchase is a linear function of the expected quality and the price: $E_t X_t^i - p_t^i = x_t^i - p_t^i$, where the random realization of the quality of product i in period t is denoted by X_t^i. The expected value of the quality realization, X_t^i, conditional on the history until period t, is given by $x_t^i = E_t X_t^i$. For simplicity, the attention is concentrated in *sampling processes*. A sampling process is a sequence $X^i = \{X_t^i\}_{t=0}^{\infty}$ of independent, identically distributed random variables $X_0^i, X_1^i, \ldots$, drawn from a distribution with an unknown (vector-valued) parameter θ^i belonging to a family of distributions. The associated density functions are denoted by $f^i(./\theta^i)$. The prior density for the parameter $\theta^i \in \boldsymbol{R}^n$ is given by $\pi_0^i(\cdot)$. The posterior beliefs are represented by $\pi_t = (\pi_t^1, \pi_t^2)$. After observing the random variable X_t^i in period t, π_t^i is converted by Bayes' rule into π_{t+1}^i:

$$\pi_{t+1}^i\left(\theta^i/X_t^i\right) = \frac{\pi_t^i(\theta^i)f^i(X_t^i\theta^i)}{\int \pi_t^i(\phi)f^i(X_t^i/\phi)d\phi}.$$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Starting with prior beliefs and applying the above equation recursively, a sequence of beliefs $\{\pi_t\}_{t=0}^{\infty}$ is obtained. The consumer and the firms discount the future with the same discount factor, β, with $0 < \beta < 1$. Past quality realizations together with past prices and past consumer decisions constitute the history of the game. Denote with H_t the set of all possible histories up to, but not including period t. An element $h \in H_t$ includes all past prices, $p_s = (p_s^1, p_s^2)$, $0 < s < t$, the consumers decision variable, $d_s = (d_s^1, d_s^2)$, where $d_s^i = 1$ if the consumer accepts the offer of firm i in period s, and $d_s^i = 0$ otherwise, and the random realizations X_s^i of the purchased product i, $0 < s < t$. Hence $h_t = (p_0, d_0, X_0^i, \ldots, p_{t-1}, d_{t-1}, X_{t-1}^i)$, where the upper index $i = 1, 2$ indicates the identity of the selling firm. A pricing strategy of seller i at any time t is a function from the history into a distribution on the real numbers: $p_t^i \colon H_t \to \Delta\boldsymbol{R}$. The buyer makes her purchase decision knowing the past play and the prices currently offered. Her acceptance strategy is a function from the history and the current prices into her decision space. Denote by $\boldsymbol{d_s} = \{d_t\}_{t=0}^{\infty}$ the sequence of decision functions starting in period s. Similarly $\boldsymbol{p_s^i} = \{p_t^i\}_{t=0}^{\infty}$ is the sequence of future pricing strategies of firm i starting in period s. The discounted expected profit for firm i under a given strategy triple $(\boldsymbol{d_s}, \boldsymbol{p_s^1}, \boldsymbol{p_s^2})$ at time s is $E_s\left[\sum_t^{\infty} \beta^{t-s} d_t^i p_t^i\right]$, and the expected present value for the consumer in period s is,

$$E_s\left[\sum_t^{\infty} \beta^{t-s}\left[d_t^1\left(X_t^1 - p_t^1\right) + d_t^2\left(X_t^2 - p_t^2\right)\right]\right].$$

Each player acts so as to maximize the expected discounted return given the beliefs over the return processes and the strategies of the other players. A solution for this decision problem is available in the statistical literature on multi-armed bandits. For $n = 2$, the maximization problem of the decision-maker is to find an allocation strategy d^* that solves:

$$Max_d E_t\left[\sum_t^{\infty} \beta^t d_t^1 X_t^1 + \sum_t^{\infty} \beta^t d_t^2 X_t^2\right].$$

The solution to the above equation is the celebrated index policy of Gittins and Jones (1974), who showed that it is possible to assign to each alternative an index function $M^i(\pi_t^i)$ that depends only on the state π_t^i of project i. The index $M^i\left(\pi_t^i\right)$ of alternative i in state π_t^i is the supremum over all terminal rewards, such that the decision-maker still prefers to continue with the random stream; or alternatively, it is the infimum over all terminal rewards such that the decision-maker is indifferent between continuing with the random sequence and retiring with the stopping reward $M^i(\pi_t^i)$.

Each purchase yields additional information about the true product quality to all parties in the model. In some periods the buyer is willing to sacrifice some of her current payoffs in order to gain additional information, which is valuable for future decisions. This temporal separation of costs and benefits causes no ex ante efficiency losses in the single-player case since the cost of experimentation has to be born by the same agent who enjoys any gains from successful experiments. The innovation of this paper is to endogenize the cost of experimentation to the consumer by allowing for price competition between the sellers. The role of prices is then to allocate intertemporally the costs and benefits of learning between buyers and sellers. Bergemann and Välimäki (1996) characterize the set of Markov perfect equilibria. The central result of the paper states that in spite of future rent seeking by the

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

firms, all Markov perfect equilibria in the model are efficient. In particular, an efficient amount of experimentation is undertaken on any Markov perfect equilibrium path. Using this fact, the authors can deduce the sequencing of consumer purchases immediately, since the efficient path coincides with the solution paths of the buyer's decision problem when prices are fixed to be identically zero. The remaining task is thus to calculate the prices that support efficient experimentation in equilibrium and determine the division of surplus between the buyer and the sellers along the efficient path. In the *cautious equilibrium*, current prices provide the buyer with insurance against future rent seeking resulting from successful experiments. The equilibrium pricing rule is quite simple. In each period, the selling price is equal to the difference in expected qualities and hence is similar to the equilibrium price in the myopic Bertrand game. The identity of the seller does not, however, coincide with the myopic game since the efficient path involves experimentation at some nodes. Bergemann and Välimäki (1996) identify this equilibrium, which besides its unique robustness properties has a strikingly simple, seemingly myopic pricing rule. Prices below marginal cost emerge naturally to sustain experimentation. Intertemporal exchange of the gains of learning is necessary to support efficient experimentation.

Felli and Harris (1996) consider a continuous model with uncertainty represented by a Brownian motion and interpret the (matching) model in the context of the labor market with wage renegotiation, and where human capital is seen as information. They use a variant of the continuous-time bandit framework to study equilibrium wage dynamics in a setting where two firms and a worker learn about the worker's aptitude to perform firm-specific tasks.

3 SOCIAL LEARNING: EXPERIMENTATION AND EXTERNALITIES

In multi-agent learning situations, informational externalities may reduce the number of experiments undertaken below the socially efficient level. As buyers choose among, for instance, new experience goods or firms decide whether to adopt a new technology, the availability of information from others' decisions gives rise to a free-rider problem. Rather than performing a costly experiment herself, a buyer may opt to wait to see how the market evaluates the new product. Thus, given that experimentation typically entails an opportunity cost, and that information obtained from an experiment is valuable to all players, individual players attempt to free ride on the experiments of others. This informational externality drives a wedge between equilibrium experimentation and socially optimal experimentation. On the other hand, an individual player may be encouraged to experiment more if, by doing so, she can bring forward the time at which the information generated by the experimentation of others becomes available. This encouragement effect mitigates the free-rider effect.

The idea of an informational externality arising in a sequential learning model is already central to Rob (1991), who studies a dynamic model of entry when the size of the market is uncertain. The model develops Bayesian learning by building insights but the novelty is that the author considers the multi-agent case. As it turns out, new features come to light when many agents interact. Most importantly, a learning externality emerges because present generations of firms confer informational benefits upon as yet non-existent firms, and these externalities result in an inefficiently slow pace of capacity expansion. These externalities exist because information is a public good – namely, observation of existing firms' profit-and-loss

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

data is informationally beneficial to potential firms, yet potential firms do not pay for these benefits.

Bolton and Harris (1999) extend the classic two-armed bandit problem to a many-agent setting in which N players each face the same experimentation problem. The main change from the single-agent problem is that an agent can now learn from the current experimentation of other agents. Information is therefore a public good, and a free-rider problem in experimentation naturally arises. More interestingly, the prospect of future experimentation by others encourages agents to increase current experimentation. There are N identical infinitely lived risk-neutral players. At each time t, these players simultaneously and independently choose the proportion of the current period $[t,\ t+dt]$ to devote to each of the two actions available to them, namely 0 (the safe action) and 1 the risky action. If player i chooses to devote a proportion α_i of the current period to the risky action, then she receives the total payoff

$$d\pi_i^0\,(t) = (1-\alpha_i)\,s\,dt + (1-\alpha_i)^{1/2}\sigma\,dZ_i^0(t)$$

from the safe action and the total payoff $d\pi_i^1\,(t) = \alpha_i\,\mu\,dt + \pi_i^{1/2}\sigma\,dZ_i^1\,(t)$ from the risky action. All players then observe all the proportions chosen and all the resulting payoffs. The underlying payoff of the risky action is unknown but common to all players, and it can be either higher or lower than that of the safe action. The actual payoff obtained by a player from an action is the underlying payoff of that action plus noise. More explicitly, all players observe α_i, $d\pi_i^0$, and $d\pi_i^1$ for all $1 \le i \le N$. Once players have chosen how to allocate their time and the payoffs have been realized, all players observe all choices and all payoffs. They therefore obtain information about the underlying payoff of the risky action by observing the payoffs derived from the risky action. Player i's objective is to maximize the expectation of the present discounted value of her payoff stream, namely

$$E\left[\int_0^\infty re^{-rt}\left(d\pi_i^0 + d\pi_i^1\right)(t)\right],$$

where $r > 0$ is the discount rate.

Several features of this model are worthy of comment. First, $d\pi_i^0\,(t)$ is composed of the deterministic contribution $s\,dt$ and the stochastic shock $\sigma\,dZ_i^0(t)$. Since the contribution $s\,dt$ is known, it follows that $d\pi_i^0$ conveys no information about $\mu \in \{l,\ h\}$. Similarly, $d\pi_i^1\,(t)$ is composed of the deterministic contribution $\mu\,dt$ and the stochastic shock $\sigma\,dZ_i^1(t)$. The first contribution ensures that $d\pi_i^1\,(t)$ conveys some information about μ. The second ensures that this information is noisy. Second, if player i devotes a proportion α_i of the current period $[t,\ t+dt]$ to the risky action, then her total payoff $d\pi_i^0\,(t)$ from the safe action is distributed normally with mean $(1-\alpha_i)\,s\,dt$ and variance $(1-\alpha_i)\,\sigma^2 dt$, and her total payoff from the risky action is distributed normally with mean $\alpha_i\,\mu\,dt$ and variance $\alpha_i\sigma^2 dt$. These means and variances can be compared with the means and variances obtained when she devotes a proportion α_i of the periods $[t,\ t+dt]$ in the interval of time $[T,\ T+\Delta T)$ to the risky action, but devotes each period exclusively either to the safe action or to the risky action. Bolton and Harris (1999) are concerned primarily with perfect equilibria in stationary Markov strategies.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Such strategies depend only on the natural state variable for the problem at hand, namely the players' common belief p that μ is high. In order to formulate the Bellman equations for equilibrium strategies then it is necessary to determine how p evolves. Let $p(t)$ denote the prior belief that μ is high at time t, suppose that player i devotes a proportion α_i of the period $[t,\ t + dt]$ to the risky action, let $p(t + dt)$ denote the posterior belief that μ is high at time $t + dt$ and let $dp(t) = p(t + dt) - p(t)$ denote the change in beliefs over the period $[t,\ t + dt]$. Finally, let $\varphi\,(p) = (p(1-)((h - l)/\sigma))^2$. Then, conditional on the information available to players at time t, the change in beliefs dpt is distributed normally with mean 0 and variance $\left(\sum_{i=1}^{N} \alpha_i\right) \varphi\,(p\,(t))\,dt$. Let $m\,(p) = (1 - p)\,l + ph$ be the expectation of the flow payoff from the risky arm when μ is believed to be h with probability p. The Bellman equation states that the current payoff is the maximum over the control variable α_i of the expectation of the current flow payoff, $(1 - \alpha_i)\,s + \alpha_i\,m\,(p)$, plus the discounted value of the *rate of change* of the continuation payoff.

Since in any given period of this game, each player must divide her time between the "safe" action and the "risky" action, then they obtain information about the underlying payoff of the risky action by observing the payoffs derived from the risky action. The authors provide an analysis of the set of stationary Markov equilibria. The equilibrium experimentation is analyzed in terms of the free-rider and the encouragement effects. The free-rider effect is easy to explain: extra current experimentation by the other players provides player i with information at no cost, and this information is used as a substitute for information that she would otherwise have had to supply for herself at some opportunity cost. As for the encouragement effect: extra future experimentation by the other players encourages player i to increase her current experimentation in order to bring forward the time at which the extra information generated by the other players becomes available. In summary, then, current experimentation by one player is a strategic substitute for current experimentation by another, but future experimentation by one player is a strategic complement for current experimentation by another.

The many-agent continuous-time strategic experimentation literature also includes Keller and Rady (2010) and Keller, Rady and Cripps (2005). The former paper studies a game of strategic experimentation with a two-armed bandit where the risky arm distributes lump-sum payoffs according to a Poisson process. The intensity of this process is either high or low, and unknown to the players. They consider Markov perfect equilibria with beliefs as the state variable. There is no equilibrium where all players use cut-off strategies, and all equilibria exhibit an encouragement effect relative to the single-agent optimum. Keller et al. (2005) analyze a game of strategic experimentation with a two-armed bandit whose risky arm might yield payoffs after exponentially distributed random times. Free riding causes an inefficiently low level of experimentation in any equilibrium where the players use stationary Markovian strategies with beliefs as the state variable. Décamps and Mariotti (2004) study a specific duopoly model where each player learns about the quality of a common value project by observing some public information plus the experience of her rival. Investment costs are private information, and the background signal takes the form of a Poisson process conditional on the quality of the project being low. The resulting attrition game has a unique, symmetric equilibrium, which depends on initial public beliefs. The authors determine the impact of changes in the cost and signal distributions on investment timing, and how equilibrium is affected when a first-mover advantage is introduced. Rosenberg, Solan and Vieille (2007) set

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

up a model where each of two players operates a one-arm bandit machine in discrete time and must decide when to stop operating the machine. The innovation of their model is that they drop the assumption that payoffs are publicly observed. As a consequence, there is no commonly observed state variable, such as a common posterior belief, on which to condition one's actions. Modeling dynamic pricing under social learning can also be found in Bose et al. (2006, 2008), in which buyers take actions sequentially, based on the history of previous purchases, prices, as well as their private information about a common value component. Their models are closer to the herding literature: a short-lived buyer makes a purchasing decision in a predetermined sequence.

3.1 Learning and Rational Expectation Models

Mirman, Salgueiro and Santugini (2014) study learning in perfect competition. A representative price-taking firm, which has complete information about the market, sells a good whose quality is unknown to some buyers. Demand is composed both of informed and uninformed buyers. The uninformed buyers use the price to infer information about quality. On the supply side, the representative, price-taking firm produces and sells the good. The cost of production is assumed to be increasing in quality and quantity. There is also a demand shock, which is known to the firm but unknown to buyers that prevents the market price from being perfectly informative about quality. Even though the firm is a price-taker, information is disseminated through the price. It is the shape of the supply curve that influences the amount of information contained in the price, which, in turn, affects the competitive equilibrium through the learning process of the uninformed buyers. Information flows and market outcomes are entwined because the uninformed buyers, who learn from prices, also participate in trading. In fact, the presence of uninformed buyers and their learning activity influence the informational content of the price. There is thus a two-way relation between trading and learning. Not only does learning from prices have an effect on decisions, but also the agents' decisions impact the market price, thus influencing the informational content of the price and the learning process. The uninformed buyers make decisions on the basis both of prior beliefs and the price – learning can be decomposed into two effects: a beliefs effect and a price effect. The beliefs effect reflects the change in behavior due to the asymmetry of information and the use of prior beliefs. The direction of the beliefs effect depends only on the bias of the prior beliefs. The price effect reflects the change in behavior due to updating beliefs. Unlike the beliefs effect, the sign of the price effect depends on the bias of the prior beliefs and the demand shock. Mirman, Salgueiro and Santugini (2015) extend the previous model by addressing the issue of risk aversion in a competitive equilibrium when some buyers engage in learning and information is conveyed through the price system. Specifically, since the learning process yields uncertainty, the authors study the effect of risk aversion on the equilibrium outcomes of the model, including the amount of information released by the market. They show that risk aversion has an effect on the market outcomes but not on the flow of information. In particular, an increase in risk aversion lowers the competitive price and quantity. However, an increase in risk aversion does not change the amount of information embedded in the equilibrium price.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

4 LEARNING MODELS OF INTRODUCTORY AND DYNAMIC PRICING

In markets for new products and services, sellers face uncertainty over the product's fit to consumers' needs. In these markets, heterogeneity in consumers' willingness to pay for the product creates the opportunity for firms to profitably adopt price discrimination techniques, such as menu pricing. In addition, information about a product's performance is widely and publicly accessible through an increasing number of channels. The availability of such aggregate information in a dynamic environment enables firms to modify their menu prices on the basis of the opinion of their customers. In this scenario, a forward-looking firm must screen consumers in order to maximize revenues, while taking into account the informational value of sales. By selling additional units of the product (for example, by offering introductory discounts), the firm accelerates the buyers' learning process, thereby trading off (1) the long-run profits that accrue due to the diffusion of information against (2) the maximization of current revenue.

Several models of introductory and dynamic pricing under uncertainty about product quality have been developed. The main work in this area is due to Bergemann and Välimäki (1997, 2000, 2002, 2006), Villas-Boas (2004, 2006), Bonatti (2011), Weng (2015), Papanastasiou and Savva (2016), among others. In particular, Bergemann and Välimäki (1997, 2000) analyze a duopoly model of price competition where market participants are uncertain about the degree of horizontal or vertical differentiation of the two firm's products, while Bergemann and Välimäki (2002) consider the entry of new products.

Bergemann and Välimäki (1997) model dynamic competition in a duopolistic market for experience goods. An established firm and a firm with a new product compete in prices in an infinite-horizon, continuous-time model. Buyers have heterogeneous preferences over the products and sellers compete in prices. Thus, the authors analyze the diffusion of a new product of uncertain value in a duopolistic market. Both sides of the market, buyers and sellers, learn the true value of the new product from experiments with it. The authors assume that the product incorporates both a common- and a private-value component to the buyers. To keep the model analytically tractable, it is also assumed that the private-value component of every buyer is common knowledge and may reflect idiosyncratic taste, location, or the like. In contrast, the common component is learned gradually over time as more experience is accumulated. The information obtained in any single trial with the new product is a noisy signal of product quality. More formally, consider a dynamic duopoly, where firms with differentiated products compete in prices in an infinite-horizon, continuous-time setting. The first firm is well established in the market, and its product characteristics are common knowledge at the beginning of the game. The second firm has a new product whose value has to be learned over time. The preferences of the buyers are described by a Hotelling location model. The buyers are uniformly distributed on the interval [0, 1], and they have unit demand at each instant of time. The value of the certain product for individual n is given by s_n with $s_n = s + nh$, $n \in [0, 1]$. The parameter $h > 0$ represents the horizontal differentiation between the products, and as such, h is a measure of the heterogeneity among the buyers. Symmetrically, the value of the uncertain product for individual n is given by, μ_n with $\mu_n = \mu + (1 - n)\, h$, $n \in [0, 1]$. The value, $\mu \in \{\mu_L, \mu_H\}$ of the new product is initially

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

unknown to all parties. Also, $0 < s-h < \mu_L < s < \mu_H < s+h$. The inner inequalities imply that the new product can be of either lower or higher value than the established one. The outer inequalities assert that in either case, the efficient allocation would assign a positive measure of buyers to both products. The size of h determines how much the value of the product to the buyer and ultimately the choice behavior of the buyer are influenced by her location.

The model is one of horizontal and vertical differentiation, where the horizontal differentiation is common knowledge at the outset but the extent of vertical differentiation is uncertain. The uncertainty about the value of the second product can be resolved only by experimentation – that is, through purchases of the new product. The performance of the new product is, however, subject to random disturbances, and any single experiment with the new product provides only a noisy signal about the true underlying value. The information conveyed by an experiment depends on the size of the experiment. As each buyer is of measure zero, the size of her purchase is negligible and hence the information generated by an individual experiment is also negligible. In consequence, all relevant information is contained in the aggregate outcome. The aggregate or market outcome is the performance of the product over all buyers, which is assumed to be publicly observable. The market players extract the information provided by the noisy market outcome to improve their common prior beliefs over time.

The learning process of the market represents a *signal-extraction problem*, which reduces to the description of the law of motion of the posterior belief. It is immediately verified that the instantaneous mean and the variance of the market outcome are linear in the market share of the new seller. As the value of μ can only be μ_L or μ_H, posterior beliefs about the quality are completely characterized by $\alpha\,(t) = pr(\mu = \mu_H / F(t))$, where $F(t)$ is the history generated by $X\,(n\,(t))$. An experimentation policy prescribes for every posterior belief α the shares of buyers allocated to the sellers. Denote by $n(\alpha)$ the market share of the new product and by $1 - n(\alpha)$ the share of the established product. The strong long-run average is defined by the following optimization problem:

$$V(\alpha_0) = sup_{'n(\alpha)} lim_{T\to\infty} E\left[\int_0^T (n\,(\alpha)\,\mu\,(n\,(\alpha)) + (1-n\,(\alpha))\,s\,(n\,(\alpha)) - \nu(\alpha))\,dt/\alpha_0\right],$$

where $\nu(\alpha)$ is the long-run average values. The pricing policies and market shares in the unique Markov perfect equilibrium are obtained explicitly. The dynamics of the equilibrium market shares display excessive sales of the new product relative to the social optimum in early stages and too low sales later on. The diffusion path of a successful product is S-shaped.

Bergemann and Välimäki (2000) present a model of entry and exit with Bayesian learning and price competition. A new product of initially unknown quality is introduced in the market, and purchases of the product yield information on its true quality. It is assumed that the performance of the new product is publicly observable. The value of the product is determined by the quality of the match between consumer preferences and product characteristics. Additional information is acquired only through repeat purchases. In each period, buyers observe a noisy signal of the true value of the product. With this assumption, all buyers and sellers condition their behavior on the same information and we can abstract from individual differences in past observations. As agents learn from the experiments of others, informational externalities arise. The authors determine the paths of sales and prices in a Markov perfect equilibrium and compare them to the Pareto-optimal paths. In contrast to the

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

one-sided experimentation problems, Bergemann and Välimäki (2000) find that equilibrium experimentation often exceeds the Pareto-optimal level in two-sided models. The paper shows that the conventional wisdom that informational externalities lead to inefficiently low levels of experimentation may be reversed in a two-sided learning model. The introduction of sellers into the multi-agent learning model creates a market where experiments are priced. The new seller sponsors the uncertain alternative and rewards buyers for experiments through low prices. In contrast to one-sided learning models, the seller provides direct incentives for the buyers to experiment. The ownership of the product allows the seller to extract the future benefits of current experimentation that would have evaporated without the assignment of property rights. The main theme of the paper is the importance of the market structure for efficiency conclusions in a model of informational externalities. Bergemann and Välimäki (2002) analyze the entry of new products into vertically differentiated markets where an entrant and an incumbent compete in quantities. The qualitative features of the optimal entry strategy are shown to depend exclusively on the relative ranking of established and new products based on current beliefs. Superior products are launched relatively slowly and at high initial prices, whereas substitutes for existing products are launched aggressively at low initial prices. Bergemann and Välimäki (2006), in contrast, study a dynamic monopoly pricing problem, in a market for new products and services, with a continuum of buyers and independent valuations. The framework of a continuum of buyers makes it impossible to discuss the impact of a single good news signal on price. Instead, they are more concerned about whether price would go down or eventually go up in equilibrium.

Villas-Boas (2004) considers a model in which consumers learn in the first period about the product they buy and then make choices in the second period about the competing products, given what they learned in the first period. The paper finds that if the distribution of valuations for each product is negatively (positively) skewed, a firm benefit (is hurt) in the future from having a greater market share today – the brand loyalty characteristic. With negative skewness, two effects are identified: on one hand, marginal forward-looking consumers are less price sensitive than myopic consumers, and this is a force toward higher prices. On the other hand, forward-looking firms realize that they gain in the future from having a higher market share in the current period and compete more aggressively in prices. For similar discount factors for consumers and firms, the latter effect dominates. Villas-Boas (2006) analyzes dynamic competition in the case in which consumers are only able to learn about their preferences for a certain product after experiencing it. The idea is that after trying a product and understanding its valuation, a consumer may prefer the product whose valuation she knows better than the product whose valuation remains mostly uncertain. With risk neutrality this can be obtained with products offering a better-than-expected fit with greater likelihood. In this sense, firms may compete fiercely for consumers to first try their products. This paper examines the competitive effects of these informational advantages in an infinite horizon model with overlapping generations of consumers. In an infinite horizon firms have to trade off exploiting any informational advantages today with having lower informational advantages in the future. Similarly, the marginal consumers realize that by purchasing a product today they will be charged a higher expected price in the future. That is, forward-looking consumers become less price sensitive. The paper finds that steady-state prices, for similar discount factors for firms and consumers, are higher the greater the informational differentiation effects. In other words, the effect of increased prices because of lower price sensitivity of the forward-looking consumers dominates the effect of lower prices caused by firm competition for market share

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

for future gains. The intuition is that in an infinite horizon firms realize that they should take advantage of any informational advantages when they have them, because in the future they also have to compete for a new generation of consumers. Furthermore, information advantages may also lead to some consumers finding that the product they tried first is not very valuable for them, and therefore they may be more likely to try another product. Villas-Boas (2006) looks at heterogeneous consumers (also without private information by the firms) but moves away from the consumer experimentation issues by limiting consumers to be in the market for only two periods.

Bonatti (2011) addresses the issue of designing dynamic menus to sell experience goods. The author characterizes the evolution of menu prices as information about product quality is gradually revealed, and examines the interaction of the screening and learning problems. More specifically, Bonatti (2011) develops a dynamic model with a monopolist firm and a continuum of small consumers. Consumers purchase repeatedly and have multi-unit demands in each period. Each consumer's valuation of the firm's product depends on both a private value and a common value component. Denote by θ an idiosyncratic, private value component, representing the buyer's personal willingness to pay for the product. For each buyer, θ belongs to the interval $\Phi = [\theta_L, \theta_H]$. The idiosyncratic component θ is the consumer's private information. It is distributed in the population according to a continuously differentiable distribution $F(\theta)$. Denote by μ a common value component that represents the quality of the match between the product and the needs of the market. This parameter may only take one of two values, $\mu \in \{\mu_L, \mu_H\}$ with $0 < \mu_L < \mu_H$. Each consumer's valuation for q units of a product is a separable function of the product's quality μ and of the consumer's willingness to pay θ. The complete information utility of a consumer with willingness to pay θ, who purchases q units of a product of quality μ, for a total charge of p, is given by $U(\mu, \theta, qp) = \mu\theta u(q) - p$. Product quality μ is unknown initially to both the firm and the consumers, and all market participants share the common prior belief $\alpha_0 = \Pr(\mu = \mu_H)$. In each period, a monopolist posts a menu of price–quantity pairs. The firm prices anonymously, and prices and quantities are adjusted flexibly. In a direct mechanism, the firm's strategy is a pair of piecewise differentiable functions $q_t : \Phi \to R_+$ and $p_t : \Phi \to R_+$ in each period. These functions determine the quantity and the total charges assigned to each buyer θ. Suppose each buyer purchases quantity $q_t(\theta)$ and pays total charges of $p_t(\theta)$. The firm then obtains flow profits of

$$\Pi\,(q_t, p_t) = \int_{\theta_L}^{\theta_H} p_t\,(\theta) - c(q_t(\theta))f(\theta)d\theta.$$

These buyers have private information about their willingness to pay, providing the firm with an incentive to price differentially. The quality of the product is unknown initially; more information is generated through experimentation. As purchases are made, both the firm and the consumers observe signals about the product's quality and, as a result, revise their beliefs. The amount of information in the market is increasing in the total quantity sold in each period. As a result, the firm can control the information flow to the market. The aggregate market experience and the associated law of motion of belief is an adaptation of Bergemann and Välimäki (1997) to allow for multi-unit demand. Bonatti (2011) characterizes the incentive-compatible menus of contracts market by adjusting the level of sales. Learning occurs through consumption, and each unit sold provides additional information. Thus the firm wants to sell

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

additional units to gain more information when uncertainty about quality is high and beliefs are more responsive to news. The second component is related to efficiency. As consumers grow more optimistic about the quality of the product, their willingness to pay increases, thereby creating the opportunity for the firm to realize larger gains from trade. Therefore, the firm offers larger quantities in this case. The third component is adverse selection. Positive signals about quality increase the spread in buyers' valuations for the product. This makes the incentive compatibility constraints more difficult to satisfy and induces the firm to offer fewer units to buyers who have a lower willingness to pay. The firm pursues the dual objectives of generating information and screening consumers simultaneously.

Weng (2015) investigates how a monopolist sells a new experience good to many buyers over time in the presence of individual learning. The monopolist and the buyers initially are equally unsure about the effectiveness of the product. Without having seen the effectiveness of the product, potential purchasers become increasingly pessimistic and in order to keep buyers purchasing the product, the price must be reduced. Dynamic monopoly pricing is modeled as an infinite-horizon, continuous-time process. The monopolist sells a perishable experience good. She can neither price-discriminate across buyers nor commit to a price rule. At each instant of time, the monopolist first posts a spot price, which is contingent on the available public information about the experiences of the buyers. Each buyer then decides to either buy one unit of the experience good or take an outside option (modeled as another good of known characteristics). The experience good generates random lump-sum payoffs according to independent Poisson processes. The arrival rate of the lump-sum payoffs depends on an unknown individual attribute, which is binary and uncorrelated across buyers. A key feature of the model is that buyers can become ex post heterogeneous in two ways: heterogeneity can be induced by either different outcomes or different actions. The author considers two different cases. In the good news case, the experience good generates positive lump-sum payoffs; in the bad news case, it generates negative lump-sum damages (e.g., side-effects of new drugs). This paper fully characterizes the symmetric Markov perfect equilibrium for both cases. The equilibrium purchasing behavior in the good news case is characterized for an arbitrary number of buyers. It turns out that the equilibrium experimentation level is always lower than the socially efficient one when at least one buyer has received a lump-sum payoff. This is due to the existence of ex post heterogeneity: known buyers are willing to pay more than unknown buyers. Without price discrimination, the trade-off between exploitation and exploration leads to inefficient early termination of experimentation. The symmetric Markov perfect equilibrium for the bad news case is always efficient.

Papanastasiou and Savva (2016) analyze how the presence of social learning interacts with the adoption decisions of strategic consumers and the dynamic pricing decisions of a monopolist firm, within a simple two-period model. There is a monopolist firm selling a new product to a fixed population of strategic consumers, over two periods. Two alternative classes of dynamic-pricing policies may be employed: the firm may either (1) announce the full price path from the beginning of the selling horizon (pre-announced pricing) or (2) announce only the first-period price, and delay the second-period price announcement until the beginning of the second period (responsive pricing). Consumers are heterogeneous in their preferences for the product and make adoption decisions to maximize their expected utility. The authors' addition to this simple model, and the focal point of their analysis, is the introduction of ex ante quality uncertainty (faced by both the firm and consumers), which may be partially resolved in the second period by observing the product reviews of first period buyers (social

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

learning, SL). Because in the presence of SL the product's quality is partially revealed in the second period, the interaction between the firm and consumers is transformed from a game whose outcome can be perfectly anticipated from the onset (in the absence of SL), to one whose outcome is of a probabilistic nature (i.e., a stochastic game). When the firm commits to a price path ex ante (pre-announced pricing), the presence of social learning increases the firm's ex ante expected profit, despite the fact that it exacerbates consumers' tendency to strategically delay their purchase. As opposed to following a price-skimming policy that is always optimal in the absence of social learning, the authors find that, for most model parameters, the firm will announce an increasing price plan. When the firm does not commit to a price path ex ante (responsive pricing), interestingly the presence of social learning has no effect on strategic purchasing delays. Under this pricing regime, social learning remains beneficial for the firm and prices may either rise or decline over time, with the latter being ex ante more likely. Furthermore, contrary to results reported in existing literature, in settings characterized by social learning, price commitment is generally not beneficial for a firm facing strategic consumers.

5 EXPERIMENTATION AND SIGNAL DEPENDENCE

In models of active learning or experimentation, agents modify their actions to affect the distribution of a signal that provides information about future payoffs. A standard result in the experimentation literature is that agents experiment, if at all, to increase their information. This finding is a direct consequence of Blackwell's theorem: one experiment is more informative than another if and only if all expected utility maximizers prefer to observe the first. Blackwell's theorem presupposes, however, that most analyses of the effect of experimentation on short-run decisions assume that the information structure is noisy; that is, for each value of the uncertain parameter the observed signal only conveys information and does not directly affect future payoffs. Often, however, signals are directly payoff relevant, a phenomenon called *signal dependence*. For example, if a firm is uncertain about its demand and uses today's sales as a signal of tomorrow's demand, then that signal may also directly affect tomorrow's profit if the good is durable or if consumers form consumption habits. Signal dependence arises very naturally in many models. With signal dependence, the analysis of optimal experimentation becomes much more complex. First, Blackwell's theorem is no longer applicable: if the signal enters the payoff function directly, some agents may prefer a less informative experiment, since the less informative experiment may be associated with better signal realizations. From this fact alone, one suspects that experimentation might reduce information. Second, the precise definition of experimentation or active learning becomes more complicated. Datta, Mirman and Schlee (2000) analyze optimal learning in models with signal dependence and noiseless information and show that, if the signal is payoff relevant, experimentation may indeed reduce information. They show that, despite the inapplicability of Blackwell's theorem, agents always experiment to increase information if the information structure is *noiseless*: given the true value of the unknown parameter, the signal realization is deterministic. Thus, they re-establish the classic experimentation theorem, even though Blackwell's theorem does not apply. Datta, Mirman and Schlee (2002) re-examine the issue of optimal experimentation in the context of dynamic problems in which today's signal or action is directly payoff relevant for the future. They show that experimentation may

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

reduce information. They also provide sufficient conditions on the primitives for information-increasing experimentation.

There are a few papers analyzing experimentation in signal dependent problems. For instance, El-Gamal and Sundaram (1993) focus on learning in the long run and show that signal dependence helps mitigate incomplete learning results that are common in standard repeated experimentation models. Intuitively, signal dependence prevents an agent from getting "stuck" in an uninformative action, ensuring that the agent continues to learn through time. Chade and Vera de Serio (2002) consider an infinite horizon single-sale model where the demand and the seller's valuation for the good depend on the state of the world that is unknown to her, and she meets potential buyers over time who behave strategically. In this environment with asymmetric information, active learning, and strategic behavior, the authors characterize the perfect Bayesian equilibria of the game and obtain two main results. First, they prove that, even if the buyers do not attempt to manipulate the seller's learning process, her incentives to post a high price and to experiment are not necessarily monotonic in the information conveyed by a buyer's rejection to buy at a high price. Second, they show that, as the discount factors of the seller and the buyers tend to one, there are equilibria where the seller *never* trades at the "wrong" price in the limit. The main forces that drive this result are (1) the existence of heterogeneous buyers in each state of the world, and (2) the assumption that the number of buyers is greater than the number of sellers.

6 APPLICATIONS OF LEARNING

In static agency models (with two types of agents) the main issue is how the principal can design a separating contract in which agents of different types self-select. In the process, these agents reveal their information. The procedure involves paying an informational rent to the "good" agent. The question then naturally arises if such a procedure also works in a repeated or dynamic context. Jeitschko, Mirman and Salgueiro (2002) study the dynamics of a stochastic, two-period principal–agent relationship. The agent's type remains the same over time. Contracts are short term. The principal designs the second-period contract, taking into account the information available about the agent after the first period. Compared to deterministic environments significant changes emerge: first, fully separating contracts are optimal. Second, the principal has two opposing incentives when designing contracts: the principal "experiments", making signals more informative, yet dampens signals, thereby reducing up-front payments. As a result, "good" agents' targets are ratcheted over time.

Uncertainty and informational asymmetries are also present when firms decide to enter a foreign market. Moner-Colonques, Orts and Sempere-Monerris (2008) examine a firm's internationalization decision – between foreign direct investment and exports – that competes against a host-informed rival. Each entry mode entails different costs and has different informational implications. The authors show that the incumbent host firm will produce below its first-period monopoly output to encourage entry via foreign direct investment. The incumbent prefers facing a stronger competitor in period two because then strategic uncertainty is reduced. Akhmetova (2010) provides a new way of looking at the dynamics of new exporters. The author points out that there may be a state of the firm in between full market access and non-exporting, where the firm is granted the chance to learn about demand before making a final decision. The duration of this learning stage is moreover determined

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

endogenously – by the firm and market characteristics, and is a random variable, affected by the draws of demand signals that the firm obtains. Similarly, the total entry cost, which here would be the sum of total testing costs and the (one-time) sunk cost of entry, is endogenous and random. The model makes it possible to predict the dynamics of exports by new exporters. One specific application is the study of the response to trade liberalization – once tariffs fall, new firms will be willing to export, and how they do so can be determined within this model. This dynamic depends on the firm and market variables, as well as features of uncertainty.

Willems (2017) investigates the extent to which the active learning motive of a seller who faces uncertainty on the slope of his demand curve can reconcile the volatile, discrete pattern followed by individual prices, with the sluggishness observed in the aggregate price level. The author develops a different explanation for the micro–macro conflict in pricing behavior – one that is also able to generate individual price series that show a lot of discreteness. The way in which this learning process is modeled differs from the passive approach used to analyze the Fed's learning process on the Phillips curve, namely, the learning process resulting from the dynamic outcome of an adaptive process that may converge to a self-confirming equilibrium. In the Fed's model, a reduced-form long-run Phillips curve is derived from having private agents adaptively forecasting inflation and the government knowing the "true" Phillips curve and how agents forecast. The government takes into account that a constant inflation policy will eventually be learned by private agents and the resulting outcome is close to the Ramsey outcome (see, for example, Sargent, 1999). Although the class of learning rules can be fairly general, it does not include the government's optimal policy and the resulting equilibrium has mis-specified beliefs. In contrast, this paper shows that modeling the seller as an optimal experimenter allows the model to replicate at least three aspects of the data: the experimentation motive generates volatile pricing patterns showing a lot of discreteness at the individual level, while the fact that there still is some learning going on makes the aggregate price level respond sluggishly to shocks. So once one considers a model that optimizes the learning process itself, it turns out that learning is not only able to match the aggregate dimension of the data, but it is also capable of reproducing important micro-elements of it. This result is not driven by some form of irrationality: sellers in the model are just responding optimally to the fact that they face a demand curve with an unobserved, time-varying slope.

The strategic experimentation framework is also used as a building block to investigate broader issues. For example, Strulovici (2010) investigates voting in a strategic experimentation environment. In particular, the incentives for collective experimentation when individual interests may be in conflict and are revealed gradually and at times that are random and may also vary across individuals. The analysis is conducted in a two-armed bandit model in which a safe alternative yields a constant, homogeneous payoff to everyone, while the risky alternative yields payoffs whose unknown distribution, or type, may vary across individuals. At each instant, society elects one of the two alternatives according to some voting rule. Individuals learn their type only through experimentation with the risky alternative. In the benchmark setting the risky action is, for each individual, either good or bad, and these types are independently distributed across individuals. Moreover, any news shock fully reveals to its recipient that the risky action is good for him, that is, he is a sure winner. By contrast, unsure voters are those individuals who have not yet received any positive news about their type, and who become increasingly more pessimistic as experimentation goes on. The benchmark setting focuses on simple majority voting.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Bergemann and Hege (2005) examine a stylized model of the funding of a research project where the merit of an idea and the time and money needed for completion are uncertain. They specifically investigate how stopping decisions are taken in the presence of agency conflicts in the form of entrepreneurial opportunism. The project will succeed with a positive probability in every period in proportion to the volume of funds provided, so that uncertainty is represented by a simple stochastic process. As continued research efforts are undertaken and no success is forthcoming, Bayesian learning will lead to a gradual downgrading of the belief in the project's prospects. The project either ends with success or will eventually be abandoned in the light of persistent negative news. Time horizon itself is infinite to address the essence of the uncertainty about the time to completion, but abandonment will occur in finite time.

Moscarini, Ottaviani and Smith (1998) set up a social learning model where individuals take actions sequentially, after observing the history of actions taken by the predecessors, and an informative private signal. If the state of the world is changing stochastically over time during the learning process, only temporary informational cascades – situations where socially valuable information is wasted – can arise. Furthermore, no cascade ever arises when the environment changes in a sufficiently unpredictable way.

Hörner and Samuelson (2009) examine a repeated interaction between an agent who undertakes experiments and a principal who provides the requisite funding for these experiments. The agent's actions are hidden, and the principal cannot commit to future actions. The repeated interaction gives rise to a dynamic agency cost – the more lucrative the agent's stream of future rents following a failure, the costlier the current incentives for the agent. As a result, the principal may deliberately delay experimental funding, reducing the continuation value of the project and hence the agent's current incentive costs. The authors characterize the set of recursive Markov equilibria and find that there are non-Markov equilibria that make the principal better off than the recursive Markov equilibrium, and that may make both agents better off. Efficient equilibria front-load the agent's effort, inducing as much experimentation as possible over an initial period, until making a switch to the worst possible continuation equilibrium. The initial phase concentrates the agent's effort near the beginning of the project, where it is most valuable, while the eventual switch to the worst continuation equilibrium attenuates the dynamic agency cost.

Bonatti and Hörner (2009) consider moral hazard problems when effort affects the speed of learning. They examine moral hazard in teams over time. Agents are collectively engaged in an uncertain project, and their individual efforts are unobserved. Free riding leads not only to a reduction in effort, but also to procrastination. The collaboration dwindles over time, but never ceases as long as the project has not succeeded. In fact, the delay until the project succeeds, if it ever does, increases with the number of agents. The authors show why deadlines, but not necessarily better monitoring, help to mitigate moral hazard.

Cripps et al. (2008) examine two agents who learn the value of an unknown parameter by observing a sequence of *private* signals. The signals are independent and identically distributed across time but not necessarily across agents. The authors show that when each agent's signal space is finite, the agents will commonly learn the value of the parameter; that is, that the true value of the parameter will become approximate common knowledge. The essential step in this argument is to express the expectation of one agent's signals, conditional on those of the other agent, in terms of a Markov chain. This allows them to invoke a contraction mapping principle ensuring that if one agent's signals are close to those expected

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

under a particular value of the parameter, then that agent expects the other agent's signals to be even closer to those expected under the parameter value. In contrast, if the agents' observations come from a countably infinite signal space, then this contraction mapping property fails.

Eeckhout and Weng (2015) analyze a general setup of experimentation with common values, and show that in addition to the well-known conditions of value matching (level) and smooth pasting (first derivative), this implies that the second derivatives of the value function must be equal whenever the agent switches action. This condition holds generally whenever the stochastic process has continuous increments. The main appeal of their approach is its applicability, which is demonstrated with two applications featuring common value experimentation: strategic pricing, and job search with switching costs. In the canonical experimentation problem in discrete time and with independent arms, the value of pulling each arm itself is not a function of the cutoff. In contrast, when there is common value experimentation, the underlying states are no longer independent and pulling any given arm affects the value of the other arms. The immediate implication is that the decision to pull any given arm affects the value of pulling all other arms. As a result, the authors have to solve for the value of each of the arms and the cutoffs simultaneously and they can no longer apply the canonical model logic.

In addition to the theoretical body of work, there are a few empirical studies attempting to quantify the importance of learning considerations on consumers' dynamic purchasing behavior. In these studies, consumers learn from their individual experience, revise their beliefs about product quality, and consequently modify their choices. A non-exhaustive list of empirical papers on learning and dynamic consumer choice includes Ackerberg (2003) on advertising, learning, and consumer choice in experience good markets; Akçura, Gonul and Petrova (2004) on learning and brand valuation; Crawford and Shum (2005) on learning in pharmaceutical demand; Erdem and Keane (1996) on brand choice processes; Göttler and Clay (2011) on tariff choices and Israel (2005) on automobile insurance. Both, Hitsch (2006) and Song and Chintagunta (2003) analyze learning about the demand on the firm's side, but focus on investment decisions, such as product adoption or exit. Wieland (2000) utilizes numerical methods to characterize the optimal policy function for a learning-by-doing problem that is general enough for practical economic applications. Dynamic simulations indicate that optimal experimentation dramatically improves the speed of learning and the stream of future payoffs. Ching, Erdem and Keane (2013) survey the basic Bayesian learning model of brand choice, pioneered by Eckstein, Horsky and Raban (1988), Roberts and Urban (1988) and Erdem and Keane (1996). The authors describe how subsequent work has extended the model in important ways. For instance, we now have models where consumers learn about multiple product attributes, and/or use multiple information sources, and even learn from others via social networks. And the model has also been applied to many interesting topics well beyond the case of brand choice, such as how consumers learn about different services, tariffs, forms of entertainment, medical treatments and drugs. Ching et al. (2013) also identified some limitations of the existing literature. These include (1) that it has been difficult to distinguish forward-looking and myopic behavior, and (2) that it has not been technically feasible to build inventories into dynamic learning models. The latter is important, because the dynamics generated by inventories can be quite similar to those generated by learning.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

7 CONCLUDING REMARKS

To conclude this chapter, it is fair to mention some references not included in the various sections above. For instance, this survey has not contemplated models of observational learning, where agents do not choose how much information they receive. Observational learning is learning that occurs through observing the behavior of others. It is a form of social learning that takes various forms, based on various processes. Since agents make only one decision, no one gains by experimentation in the observational learning models. The interested reader may consult Smith and Sørensen (2011) and Smith, Sørensen and Tian (2015), where a formal analogy is made between the observational learning model and a model of an impatient experimenter. Models of word-of-mouth transmission of information parallel the observational learning literature: Banerjee and Fudenberg (2004) analyze a model in which agents make one binary decision, while Ellison and Fudenberg (1993) study a word-of-mouth model in which agents base decisions solely on the payoff information contained in their sample.

A different issue is learning and convergence to full-information equilibrium with uninformed firms or how to attain full-information equilibrium with competitive dynamics. Repeated interaction in the market place provides an answer to both the lack of information aggregation with simple market mechanisms and the formation of (fully revealing) rational expectations. Agents process repeated observations of public market data and learn about the relevant uncertainty (adjusting their beliefs in response to observations). The study of convergence to full-information equilibrium with repeated interaction is important since it provides a foundation to competitive equilibria under private information. Indeed, the competitive model with full information will be approximately right even in private-information environments if repeated interaction in the market place resolves the uncertainty. In the Bayesian setting, agents have priors over possible sequences of market prices and update at each date. If the prior does not coincide with the objective distribution on price sequences generated by the behavior of agents, then there is learning about rational expectations equilibria. In general, a crucial element to obtain convergence to a limit equilibrium with correct beliefs with respect to the underlying true economy is the a priori assumed coordination of expectations of agents. Jun and Vives (1996) investigate learning and convergence to a full-information equilibrium with uninformed firms and Vives (1993) considers the implications of uncertainty about costs and studies the market dynamics with asymmetric information. In all cases the speed of learning and the rate of convergence to full-information equilibria are characterized.

Evolutionary game models are close to models of social learning. While these models focus on strategic situations, the players are decidedly non-strategic. Evolutionary models have been used to make selections in general games with multiple, but non-strict, equilibria. Kandori, Mailath and Rob (1993) and Foster and Young (1990) study strategic evolutionary models in which there is perpetual randomness.

Some of the literature on learning in games can be seen as an effort to revive the fictitious play model, or at least to identify more precisely when play converges to equilibrium. The research program of Fudenberg and Kreps (1993, 1995) is dedicated to finding conditions on learning models that guarantee local stability of Nash equilibria. Fudenberg and Levine (1998) discuss some work on learning in games. Models in which learning is described by Bayes' rule lead to surprisingly powerful results. Kalai and Lehrer (1993) present one of the few analyses of repeated games in which players are not myopic.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

One of the main questions in the literature on Bayesian learning is whether learning leads, asymptotically, to good outcomes. Results supporting the statement that agents make optimal decisions asymptotically tell us little about the kind of decisions people normally make. If rational learning leads to good outcomes in a realistic amount of time, then an effort should be made to obtain the faster rates of convergence that are observed in experiments. However, rational learning also provides broad scope for the assertion that learning does not lead to optimal decisions in a realistic length of time. As Sobel (2000, p. 259) points out, "Intelligently designed institutions perform well even if individual participants are poorly informed or boundedly rational" then "the literature on learning could identify institutions that lead to good outcomes either because learning is easier or faster in those settings or because outcomes are not sensitive to poor decisions that agents may make". This suggests that, although the literature in mechanism design tells us a lot about how to design markets, rational learning models may add a new dimension to market design.

REFERENCES

Ackerberg, D.A. 2003. Advertising, learning, and consumer choice in experience good markets: An empirical examination. *International Economic Review*, 44(3), 1007–1040.

Aghion, P., Espinosa, M.P. and Julien, B. 1993. Dynamic duopoly games with learning through market experimentation. *Economic Theory*, 3, 517–539.

Akçura, M.T., Gonul, F.F. and Petrova, E. 2004. Consumer learning and brand valuation: An application on over-the-counter drugs. *Marketing Science*, 23(1), 156–169.

Akhmetova, Z. 2010. Firm experimentation in new markets. Mimeo.

Alepuz, D. and Urbano, A. 1994. Market learning and price dispersion. *Working Paper* AD 94-14. Instituto Valenciano de Investigaciones Económicas (IVIE).

Alepuz, D. and Urbano, A. 1995. A note on risk aversion and learning behavior. *Economics Letters*, 49(1), 51–58.

Alepuz, D. and Urbano, A. 1999. Duopoly experimentation: Cournot competition. *Mathematical Social Sciences*, 37, 165–188.

Alepuz, D. and Urbano, A. 2005. Learning in asymmetric duopoly markets: Competition in information and market correlation. *Spanish Economic Review*, 7(3), 209–243.

Avila-Baltuille, C., Caballero-Sanz, F. and Urbano, A. 1993. A model of intra-firm diffusion with Bayesian learning: The monopoly case. *Revista Española de Economía. Monográfico: "Investigación y Desarrollo"*, 83–107.

Banerjee, A. and Fudenberg, D. 2004. Word-of-mouth learning. *Games and Economic Behavior*, 46, 1–22.

Belleflamme, P. and Bloch, F. 2001. Price and quantity experimentation: A synthesis. *International Journal of Industrial Organization*, 19, 1563–1582.

Bergemann, D. and Hege, U. 2005. The financing of information: Learning and stopping. *The RAND Journal of Economics*, 36(4), 719–752.

Bergemann, D. and Välimäki, J. 1996. Learning and strategic pricing. *Econometrica*, 64(5), 1125–1149.

Bergemann, D. and Välimäki, J. 1997. Market diffusion with two-sided learning. *The RAND Journal of Economics*, 28(4), 773–795.

Bergemann, D. and Välimäki, J. 2000. Experimentation in markets. *Review of Economic Studies*, 67(2), 213–234.

Bergemann, D. and Välimäki, J. 2002. Entry and vertical differentiation. *Journal of Economic Theory*, 106(1), 91–125.

Bergemann, D. and Välimäki, J. 2006. Dynamic pricing of new experience goods. *Journal of Political Economy*, 114(4), 713–743.

Bernhardt, D. and Taub, B. 2015. Learning about common and private values in oligopoly. *RAND Journal of Economics*, 46(1), 66–85.

Blackwell, D. 1953. Equivalent comparison of experiments. *Annals of Mathematical Statistics*, 24, 265–272.

Bolton, P. and Harris, C. 1999. Strategic experimentation. *Econometrica*, 67(2), 349–374.

Bonatti, A. 2011. Menu pricing and learning. *American Economic Journal: Microeconomics*, 3(3), 124–163.

Bonatti, A. and Hörner, J. 2009. Collaborating. *American Economic Review*, 101(2), 632–663.

Bose, S., Orosel, G., Ottaviani, M. and Versterlund, L. 2006. Dynamic monopoly pricing and herding. *The RAND Journal of Economics*, 37(4), 910–928.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Bose, S., Orosel, G., Ottaviani, M. and Versterlund, L. 2008. Monopoly pricing in the binary herding model. *Economic Theory*, 37(2), 203–241.

Chade, H. and Vera de Serio, V. 2002. Pricing, learning, and strategic behavior in a single-sale model. *Economic Theory*, 19, 333–353.

Ching, A.T., Erdem, T. and Keane, M.P. 2013. Learning models: An assessment of progress, challenges, and new developments. *Marketing Sciences*, 32(6), 913–938.

Crawford, G.S. and Shum, M. 2005. Uncertainty and learning in pharmaceutical demand. *Econometrica*, 73(4), 1137–1173.

Creane, A. 1994. Experimentation with heteroskedastic noise. *Economic Theory*, 4(2), 275–286.

Creane, A. 1995. Endogenous learning by doing and information sharing. *International Economic Review*, 36(4), 985–1002.

Cripps, M.W., Ely, J.C., Mailath, G.J. and Samuelson, L. 2008. Common learning. *Econometrica*, 76(4), 909–933.

Datta, M., Mirman, L.J. and Schlee, E.E. 2000. Learning with noiseless information and payoff-relevant signals. *Economic Theory*, 16, 63–75.

Datta, M., Mirman, L.J. and Schlee, E.E. 2002. Optimal experimentation in signal-dependent decision problems. *International Economic Review*, 43(2), 577–607.

Décamps, J.P. and Mariotti, T. 2004. Investment timing and learning externalities. *Journal of Economic Theory*, 118, 80–102.

DeGroot, M.H. 1962. Uncertainty, information and sequential experiments. *Annals of Mathematical Statistics*, 33, 404–419.

Easley, D. and Kiefer, N.M. 1988. Controlling a stochastic process with unknown parameters. *Econometrica*, 56, 1045–1066.

Eckstein, Z., Horsky, D. and Raban, Y. 1988. An empirical dynamic model of brand choice, *Working Paper* 88, University of Rochester.

Eeckhout, J. and Weng, X. 2015. Common value experimentation. *Journal of Economic Theory*, 160, 317–339.

El Gamal, M.A. and Sundaram, K. 1993. Bayesian economists. . .Bayesian agents: An alternative approach to optimal learning. *Journal of Economic Dynamics and Control*, 17, 355–383.

Ellison, G. and Fudenberg, D. 1993. Rules of thumb for social learning. *Journal of Political Economy*, 101, 612–643.

Erdem, T. and Keane, M.P. 1996. Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Science*, 15, 1–20.

Felli, L. and Harris, C. 1996. Learning, wage dynamics and firm-specific human capital. *Journal of Political Economy*, 104(4), 838–868.

Foster, D. and Young, P. 1990. Stochastic evolutionary game dynamics. *Theoretical Population Biology*, 38, 219–232.

Fudenberg, D. and Kreps, D. 1993. Learning and mixed equilibria. *Games and Economic Behavior*, 5, 320–367.

Fudenberg, D. and Kreps, D. 1995. Learning in extensive-form games, I: Self-confirming equilibria. *Games and Economic Behavior*, 8, 20–55.

Fudenberg, D. and Levine, D. 1998. Learning in games. *European Economic Review*, 42, 631–639.

Gittins, J.C. and Jones, D.M. 1974. A dynamic allocation index for the sequential design of experiments. In J. Gani et al. (eds), *Progress in Statistics*. Amsterdam: North-Holland, 241–266.

Göttler, R.L. and Clay, K. 2011. Tariff choice with consumer learning and switching costs. *Journal of Marketing Research*, 48(4), 633–652.

Grossman, S.S., Kihlstrom, R.E. and Mirman, L.J., 1977. A Bayesian approach to the production of information and learning by doing. *Review of Economic Studies*, 44, 533–547.

Harrington, J.E. 1995. Experimentation and learning in a differentiated-products duopoly. *Journal of Economic Theory*, 66, 275–288.

Hitsch, G.J. 2006. An empirical model of optimal dynamic product launch and exit under demand uncertainty. *Marketing Science*, 25(1), 1–25.

Hörner, J. and Samuelson L. 2013. Incentives for experimenting agents. *The RAND Journal of Economics*, 44(4), 632–663.

Israel, M. 2005. Services as experience goods: An empirical examination of consumer learning in automobile insurance. *American Economic Review*, 95(5), 1444–1463.

Jeitschko, T., Mirman, L.J. and Salgueiro, M. 2002. The simple analytics of information and experimentation in dynamic agency. *Economic Theory*, 19, 549–570.

Jun, B. and X. Vives. 1996. Learning and convergence to a full-information equilibrium are not equivalent. *Review of Economic Studies*, 63, 653–674.

Kalai, E. and Lehrer, E. 1993. Rational learning leads to Nash equilibrium. *Econometrica*, 61, 1019–1045.

Kandori, M., Mailath, G. and Rob, R. 1993. Learning to play equilibria in games with stochastic perturbations. *Econometrica*, 61, 29–56.

Keller, G. and Rady, S. 1999. Optimal experimentation in a changing environment. *Review of Economic Studies*, 66(3), 475–507.

Keller, G. and Rady, S. 2010. Strategic experimentation with Poisson bandits. *Theoretical Economics*, 5(2), 275–311.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Keller, G., Rady, S. and Cripps, M. 2005. Strategic experimentation with exponential bandits. *Econometrica*, 73(1), 39–68.

Kihlstrom, R., Mirman, L. and Postlewaite, A. 1984. Experimental consumption and the "Rothschild effect". In M. Boyer and R.E. Kihlstrom (eds) *Bayesian Models of Economic Theory*. Amsterdam: North-Holland.

Milgrom, P.R. 1981. Good news and bad news: Representation theorems and applications. *Bell Journal of Economics*, 12, 380–391.

Mirman, L.J., Salgueiro, E.M. and Santugini, M. 2014. Learning in a perfectly competitive market. *CIRPEE Working Paper* 14–23.

Mirman, L.J., Salgueiro, E.M. and Santugini, M. 2015. Noisy learning in a perfectly competitive market. *CIRPEE Working Paper* 15–02.

Mirman, L.J., Samuelson, L. and Schlee, E. 1994. Strategic information manipulation in duopolies. *Journal of Economic Theory*, 62, 363–384.

Mirman, L.J., Samuelson, L. and Urbano, A. 1993a. Monopoly experimentation. *International Economic Review*, 34 (3), 549–563.

Mirman, L.J., Samuelson, L. and Urbano, A. 1993b. Duopoly signal-jamming. *Economic Theory*, 3, 129–149.

Moner-Colonques, R., Orts, V. and Sempere-Monerris, J.J. 2008. Entry in foreign markets under asymmetric information and demand uncertainty. *Southern Economic Journal*, 74(4), 1105–1122.

Moscarini, G. and Smith, L. 2001. The optimal level of experimentation. *Econometrica*, 69(6), 1629–1644.

Moscarini, G., Ottaviani, M. and Smith, L. 1998. Social learning in a changing world. *Economic Theory*, 11, 657–665.

Papanastasiou, P. and Savva, N. 2016. Dynamic pricing in the presence of social learning and strategic consumers. *Management Science*, 63(4), 919–939.

Prescott, E.C. 1972. The multi-period control problem under uncertainty. *Econometrica*, 40, 1043–1058.

Rob, R. 1991. Learning and capacity expansion under demand uncertainty. *Review of Economic Studies*. 58, 665–675.

Roberts, J. and Urban, G. 1988. Modeling multiattribute utility, risk, and belief dynamics for new consumer durable brand choice. *Management Science*, 34(2) 167–185.

Rosenberg, D., Solan, E. and Vieille, N. 2007. Social learning in one-arm bandit problems. *Econometrica*, 75(6), 1591–1611.

Rothschild, M. 1974. A two-armed bandit theory of market pricing. *Journal of Economic Theory*, 9, 185–202.

Rustichini, A. and Wolinsky, A. 1995. Learning about variable demand in the long run. *Journal of Economic Dynamics and Control*, 19, 1283–1292.

Sargent, T.J. 1999. *The Conquest of American Inflation*, Princeton, NJ: Princeton University Press.

Smith, L. and Sørensen, P.N. 2011. Observational learning. In S.N. Durlauf and L.E. Blume (eds), *The New Palgrave Dictionary of Economics*. Online Edition.

Smith, L., Sørensen, P.N. and Tian, J. 2017. Informational herding, optimal experimentation, and contrarianism. *Review of Economic Studies*. Under Revision.

Sobel, J. 2000. Economists' models of learning. *Journal of Economic Theory*, 94, 241–261.

Song, I. and Chintagunta, P.K. 2003. A micromodel of new product adoption with heterogeneous and forward-looking consumers: Application to the digital camera category. *Quantitative Marketing and Economics*, 1(4), 371–407.

Strulovici, B. 2010. Learning while voting: Determinants of collective experimentation. *Econometrica*, 78(3), 933–971.

Trefler, D. 1993. The ignorant monopolist: Optimal learning with endogenous information. *International Economic Review*, 34(3), 565–581.

Urbano, A. 1992. Monopoly experimentation: A generalization. *Revista Española de Economía*, 9, 275–305.

Urbano, A. 1993. Information gathering and manipulation in a duopoly. *Revista Española de Economía,* 10(1), 19–43.

Villas-Boas, J.M. 2004. Consumer learning, brand loyalty, and competition. *Marketing Science*, 23(1), 134–145.

Villas-Boas, J.M. 2006. Dynamic competition with experience goods. *Journal of Economics and Management Strategy*, 15(1), 37–66.

Vives, X. 1990. Trade association disclosure rules, incentives to share information and welfare. *RAND Journal of Economics*, 21, 409–430.

Vives, X. 1993. How fast do rational agents learn? *Review of Economic Studies*, 60, 329–347.

Weng, X. 2015. Dynamic pricing in the presence of individual learning. *Journal of Economic Theory*, 155, 262–299.

Wieland, V. 2000. Learning by doing and the value of optimal experimentation. *Journal of Economic Dynamics and Control*, 24(4), 501–534.

Willems, T. 2017. Actively learning by pricing: A model of an experimenting seller. *The Economic Journal*, DOI: 10.1111/ecoj.12360.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

18. Information sharing in oligopoly

Sergio Currarini and Francesco Feri*

1 INTRODUCTION

Oligopolistic firms face obvious incentives to coordinate their output and price strategies, in order to collude on otherwise contested markets. A seemingly related question is whether oligopolists face incentives to disclose or even share their private information on either market or technological conditions *before* engaging in market competition. These incentives to "collaborate" with rival firms do not stem from the softening (or even avoidance) of competition, but rather from the modification of the informational structure under which the upcoming competition will take place. It has been argued that understanding such incentives has strong policy relevance, as it can guide regulative intervention by suggesting whether evidence of information sharing should or should not be interpreted as evidence of market collusion (see Kuhn and Vives's 1995 report on the EU industry).

These considerations have motivated vast attention in the theoretical industrial organization (IO) literature, where game-theoretic models of incomplete information have been employed to disentangle the forces that finally result in the incentives to disclose or share one's private information. Most papers have dealt with situations where information is shared prior to the realization of uncertainty (the *ex ante* case), so that the decision to disclose does not signal anything about one's own private information. There have also been a few attempts to deal with the *interim* case, where firms receive their private information prior to taking action. The *interim* case provides firms' strategies with a signalling content, and is therefore more complex.

In this survey we discuss the main insights from the vast body of research on the subject, with special attention to what is now understood about the role of the various aspects of the oligopolistic model and of the informational structure in generating incentives to share. While the existing literature has studied the forces behind multilateral sharing (disclosure of private information to all other firms in the market), we devote a considerable part of this survey to recent developments of the model that encompass targeted and bilateral sharing agreements, where pairs of firms decide to exclusively share their private information. Using economics terminology, shared information is here a "club" good, compared to the "public good" property of shared information in the traditional multilateral model.

1.1 The Basic Oligopoly Model

This section is based on Raith's (1996) general model of information sharing in oligopoly. Consider a stochastic oligopoly model with n firms; each firm i's profit is affected by a random variable τ_i, distributed normally with zero mean and variance t_i. The covariance between τ_i

* Sergio Currarini wishes to acknowledge the support of the Ministry of Education and Science of the Russian Federation, Grant No. 14.U04.31.0002, administered through the NES CSDSI.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

and τ_j is $t_n \geq 0$ for all i, j. Depending on the specific application, this may represent deviations from the mean of either marginal costs or of the intercept of the demand function. We refer to the vector $\tau = (\tau_1, \tau_2, \ldots, \tau_n)$'s as the "state of the world". Each firm i plays a strategy $s_i \in R_+$ (a quantity in Cournot competition and a price in Bertrand competition). The following expression describes the relation between firm i's profit and the ith component τ_i of the state of the world, firm i's strategy and the strategies of all other firms j:

$$\pi_i = \alpha_i(\tau_i) - \sum_{j \neq i} \varepsilon s_i s_j + (\beta + \gamma_s \tau_i - \delta s_i) s_i. \tag{18.1}$$

In the above expression, the term $\alpha_i(\tau_i)$ is a function of τ_i, δ is assumed to be positive, and $\varepsilon \in \left(-\frac{\delta}{n-1}, \delta\right]$. Expression (18.1) fits a large set of oligopolistic models. Uncertainty on a common demand intercept corresponds to the case of perfectly correlated states of the world (the τ_i's) and $\gamma_s = 1$. Uncertainty about costs corresponds to the case where $\gamma_s = -1$ (in which case the demand intercept is given by β alone). A positive and small ε (relative to δ) expresses a high degree of product differentiation (or a quickly increasing marginal cost as in Kirby, 1988); a negative ε expresses strategic complementarity in firms' strategies.

Firms do not observe the state of the world τ. However, each firm i privately observes a noisy signal y_i about τ_i, with $y_i = \tau_i + \eta_i$, where the noise η_i is normally distributed with zero mean, variance u_{ii} and covariance $u_n \geq 0$. We assume that $t_i = t$ and $u_{ii} =$ for all i, and denote by $p_s = (t + u)$ the variance of signals and by $p_n = (t + u_n)$ the covariance.

The following classification of informational structures have been shown by Raith (1996) to be key in determining the incentives of firms to disclose and share information (we will discuss Raith's work in the next sections):

- **Common value (CV)**: $t_n = t$. In this case, all τ_i's are perfectly correlated. This is the case, for instance, of firms facing a common demand intercept, on which each firm receives a private noisy signal, or of firms producing with perfectly correlated costs.
- **Independent values (IV)**: $t_n = u_n = 0$. In this case, each firm i's profit is affected by a state of the world τ_i whose distribution is independent of the distribution of all other τ_js. In addition, this condition requires that firms' signals are conditionally independent, that is, that the noise of each firm's signal is independent of the noise of the other firms' signals. So, correlation is ruled out both in the market or technological conditions faced by firms, and in the informational channels that firms use to acquire information of their own τ.
- **Perfect signals (PS)**: $u = 0$. This assumption requires that each firm i gets to know with infinite precision its own state τ_i.

1.2 Modelling Information Disclosure and Sharing

In addition to observing their own private signals, firms are allowed to modify the market information structure by disclosing and/or sharing private information with other firms. We will first discuss the two prevailing models used in the literature to represent the technology of information disclosure and sharing.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

In the *strategic* model, each firm decides whether to unilaterally disclose its own information to other firms, and receives the information of all other disclosing firms irrespective of its own disclosure decision. This model is well described as a game in which each firm's strategy is whether to disclose or not its information to either all or a subset of firms, and firms' expected payoffs depend on the disclosure strategy of all firms in the market.

In the *contractual* model, firms share information with competitors on a *quid pro quo* basis: by refusing to disclose its own information, a firm also loses the information of the other disclosing firms. Almost all papers in the literature (with the exception of Kirby, 1988 and Malueg and Tsutsui, 1996) have focused on the comparison between the total absence of sharing and the universal sharing of information (an industry-wide agreement), interpreting the difference in expected payoffs as the incentives to form a trade association for the industry.

The *contractual* model naturally leads itself to a more extensive analysis, based on games of coalition formation and of network formation, where firms can form information-sharing coalitions or bilateral agreements, and exclude other firms from their private information. Malueg and Tsutsui (1996) have focused on the formation of small coalitions of sharing firms, adopting the concept of "coalition-proof equilibrium", based on the robustness of a coalition to "credible" deviations of sub coalitions (see Bernheim and Whinston, 1987).

In a recent paper, Currarini and Feri (2015) have studied the incentives of firms to form bilateral sharing agreements. In the spirit of the contractual model, they have maintained the assumption of *quid pro quo* exchanges: firm i is not allowed to observe firm j's signal unless it reveals its own signal to firm j. Differently from the multilateral case, transitivity of sharing agreements may fail, in the sense that information sharing between firms i and j and between firms j and k need not imply information sharing between firms i and k. An "information structure" is therefore given by a non-directed *network* g, in which each link ij denotes a bilateral information-sharing agreement between firms i and j. We denote by $N_i \equiv \{j : ij \in g\} \cup \{i\}$ the set of neighbours of i in g (including i) and we denote by $n_i = |N_i|$ the number of such neighbours. The information available to firm i in the information structure g is therefore $I_i(g) \equiv \left\{y_j : j \in N_i\right\}$, that is, the set of signals observed by the neighbours of i. We will use the notation $g + ij$ to denote the network obtained by adding to g the link $ij \notin g$, and $g - ij$ to denote the network obtained by severing the link $ij \in g$ from g.

The basic and least stringent notion of equilibrium that is in line with this approach is that of "pairwise stability", first introduced by Jackson and Wolinsky (1996). A *pairwise stable* network g satisfies two conditions: no firm has an incentive to sever any of its links in g and no pair of firms have an incentive to add a new link to g. More formally, let $E\pi_i(g)$ denote the expected profit of firm i if the information structure underlying market competition is described by the network g.

Definition 1 *The information structure g is* ***pairwise stable*** *if:*

(1) $E\pi_i(g) \geq E\pi_i(g - ij)$ *for all* $ij \in g$*;*
(2) $E\pi_i(g + ij) > E\pi_i(g) \rightarrow E\pi_j(g + ij) < E\pi_j(g)$ *for all* $ij \notin g$.

The above definition implicitly rules out the possibility of side payments between firms that are contingent on the sharing of information. In the presence of such transfers, the

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

two conditions of definition 1 would be replaced by the following (see Jackson and Wolinsky, 1996):

Definition 2 *The information structure g is* ***pairwise stable with transfers*** *if:*

(1′) $E\pi_i(g) + E\pi_j(g) \geq E\pi_i(g - ij) + E\pi_j(g - ij)$ *for all* $ij \in g$;
(2′) $E\pi_i(g + ij) + E\pi_j(g + ij) \leq E\pi_i(g) + E\pi_j(g)$ *for all* $ij \notin g$.

A stronger notion of stability allows each firm to revise any subset of its links (instead of only one link), and any pair of firms to form a new one. This notion of *pairwise Nash stability* has been sometimes used in the literature (see Bloch and Jackson, 2006). Formally, point (1) in definition 2 is replaced as follows: *(1″)* $E\pi_i(g) \geq E\pi_i(g - L)$ for all i and subsets L of links maintained by i in g.

In the next sections we discuss the incentives of firms to share information in the various models and approaches discussed above. We start in Section 2 with the traditional multilateral model, to then report in greater detail the more recent contributions on the bilateral model in Section 3.

2 MULTILATERAL INFORMATION SHARING

In this section we discuss the incentives to either disclose or share information with all other firms in a common market. Multilateral information sharing has been the object of a large body of literature, pioneered by Novshek and Sonnenschein (1982), Clarke (1983), Vives (1985), Fried (1984), Gal-Or (1985; 1986), Li (1985), Sakai (1985), Shapiro (1986), Kirby (1988), Sakai and Yamato (1989). More recently, Raith (1996) has provided a general and insightful analysis, encompassing all previous models and shedding light on apparent weaknesses of the theory. The merit of Raith's work is that is has uncovered the primitive forces that are behind all results in the literature, independently of the details of the model of market competition and of technological assumptions.

In a nutshell, the effect of information sharing on competition and profits is the result of: (1) a finer information on market and/or technological conditions (one's own profit function); and (2) a change in the correlation of firms' market strategies, due to the "more similar" information sets available to firms. Early contributions have suggested that information sharing prior to market competition is profitable when it concerns private cost parameters, and when it concerns market demand parameters as long as firms' strategies are complements. When firms, strategies are substitutes, the increased correlation of preferences, due to a widespread better knowledge of demand conditions, makes unilateral disclosure of information unprofitable; moreover, pooling of information in an industry-wide agreement becomes profitable only when products' differentiation is high. In his 1996 paper, Raith shed further light on these early results by stressing the role played by the precision of signals, and the complex interplay between strategic structure of the oligopoly game and the induced correlation of strategies in equilibrium.[1]

[1] All results discussed here refer to the *ex ante* model of information sharing, in which firms set their disclosure and sharing rules prior to being informed via a private signal. There have been a few contributions considering the

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

2.1 Incentives to Disclose Information

We start by considering the incentives of firms to unilaterally disclose their private information. We have referred to this case as the "strategic model". By disclosing private information, a firm is refining the knowledge of rival firms about its own profit function; when states and/or signals are correlated, it also refines rival firms' knowledge about their respective states of the world (the τ_js). In any case, after disclosure there is more "shared" information in the system, and this affects the correlation of strategies in equilibrium. Intuition would suggest that when states are positively correlated, firms take advantage of the increased correlation of strategies when these are complements. The following result in Raith (1996) gives the full account of firms' incentives to disclose for all possible scenarios:

Proposition 1 *Under "independent values", "perfect signals" and "common value" with strategic complements, disclosing information is a dominant strategy. Under "common value" with strategic substitutes, concealing information is a dominant strategy.*

This result is best understood by considering the effects of disclosing private information on one's expected profits. Raith (1996) has shown how:

1. Letting rival firms refine their knowledge about their own respective profit functions has a positive effect on expected profits under strategic complements, and a negative effect on profits under strategic substitutes.
2. Letting rival firms refine their knowledge about one's own profit function always has a positive effect on one's own expected profits.

By disclosing one's own private information, a firm affects its rivals' knowledge about their respective payoff functions only if signals are correlated and imperfectly observed. This implies that point (1) above does not apply under IV and PS. Only point (2) applies in those cases, and disclosure is always profitable. Under CV and complements, both (1) and (2) work in favour of disclosure. Under CV and substitutes, the incentives to disclose result from the trade-off between points (1) and (2). The results by Fried (1984), Li (1985) and Shapiro (1986) follow as corollaries of the above results for "perfect signals". Also, results by Vives (1985) and Gal-Or (1985) on CV situations come as special cases of the result above. Interestingly, the key categories driving the incentives to disclose are not whether uncertainty is about demand or costs, or (not only) whether there are private or common values, but rather about how precise and correlated the signals are, since these aspects of information will determine to what extent disclosure improves rivals, knowledge about their own profit functions, and, ultimately, to what extent correlation of strategies will increase as a result.

interim model, in which disclosure and sharing occur after the signals are privately observed. When information is verifiable, the decisions of whether to share takes on a signalling power. If uncertainty is about costs, each firm would like to be perceived as low cost; it follows that low-cost firms reveal their type, and information revelation unravels to the whole market (Okuno-Fujiwara, Postlewaite and Suzumura, 1990; Van Zandt and Vives, 2006). When there is uncertainty about whether the firm is indeed informed about its cost, then unravelling may fail even if information is verifiable (Jansen, 2005).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

2.2 Incentives to Share Information

We then turn to the richer case of exclusive information-sharing contracts, which we referred to as the "contractual model". Here, firms disclose their private information to all rival firms and receive in return all private information held by rival firms. The main exercise consists therefore in comparing the expected payoff of firms in two opposite scenarios: no information sharing and universal (or complete) pooling of information:

Proposition 2 *Under "independent values" and "perfect signals", complete pooling is always profitable. In the "common value" case, pooling is profitable if:*

$$\frac{\epsilon}{\delta} < \frac{2}{n+1}.$$

The fact that under independent values and perfect signals firms prefer to pool information comes almost as a corollary of Proposition 1. Here, in addition to disclosing one's own private information (which has a positive effect on profits), firms receive additional information about rivals' signals and, therefore, behaviour that, under these conditions of IV and PS, is beneficial. Note that here receiving information from rivals does not improve a firm's information about its own state of the world (in Raith's terminology, there is no "direct adjustment" of strategies after sharing.

The more interesting result here is about common value situations. Here, the final effect of sharing on profits comes as a result of the two effects discussed for the case of disclosure, plus the positive effect of refining one's own information about other firms' behaviour and about one's state of the world (through the correlation induced by the common state of the world). The main insight here is that pooling becomes profitable when the effect of the increased correlation of strategies in equilibrium is weak enough, and this happens when market competition is not too harsh – that is, when the level of product differentiation (here measured by the inverse of ϵ) is strong enough (Kirby, 1988, has shown that the same effect as product differentiation is replicated by steeply increasing marginal costs of production).

Malueg and Tsutsui (1996) have raised the issue of smaller-scale agreements. They show that not only industry-wide agreements can be profitable and immune to individual defections (when products are differentiated), but also that a coalitional agreement by a subset of firms can be stable to defections (more precisely, can be a coalition-proof Nash equilibrium). Their result is obtained in the framework of a three-firm model, and fails to predict information sharing of any kind when products are strongly homogeneous. Smaller-scale agreements are not therefore conducive to information sharing when goods are homogeneous.

As we will discuss in some detail in the next section, a recent contribution by Currarini and Feri (2015) can be used to show that small-scale sharing agreements between firms (bilateral agreements) can generate positive amounts of information sharing in equilibrium even when products are perfectly homogeneous and strategies are substitutes. This result rests on the effect of the conditional correlation of private signals on firms' incentives to share information in small coalitions ($u_n > 0$). The basic intuition behind this result goes as follows. When firms' private information is (conditionally) correlated, the exchange of information within a small coalition of firms has the effect of refining these firms' expectation about all outside firms' signals (and behaviour). This refinement results from the assumed

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

conditional correlation of signals, and comes at "no cost", since it does not imply any additional correlation of strategies with the outside competitors who do not receive the information about coalitional members' signals. The result is therefore due to the "strategic adjustment" mentioned in Raith (1996). The magnitude of the resulting increase in expected profits is larger the larger the number of firms outside the sharing coalition.

3 BILATERAL CONTRACTS AND INFORMATION-SHARING NETWORKS

We now turn to the bilateral model, in which firms agree to share information in pairs. We restrict our attention to the case of uncertainty on a common demand intercept. This is therefore the case of common value: $\tau_i = \tau$ for all i. For a more general analysis that covers the whole class of quasilinear games, see Currarini and Feri (2015). This section is based on published and unpublished results of the authors. In particular, all results in Sections 3.2 and 3.4 are unpublished, and can be found in working paper versions of the paper "Bilateral Information Sharing in Oligopoly" by the authors (Currarini and Feri, 2007). We omit formal proofs of the propositions, some of which involve long algebraic expression and use of computation software. All proofs are available from the authors on request.

3.1 Equilibrium Use of Information

With each possible information structure g we associate the Bayesian Nash equilibrium of the game in which each firm i sets its strategy s_i in order to maximize its profit, given its available information determined by i's links in g, and given the optimal strategies of other firms. Formally, a Bayesian Nash equilibrium associated with g is a vector $s^*(g)$ of function mapping, for each $i \in N$, the available information $I_i(g)$ into a choice s_i, and such that for each firm i, the function $s_i^*(g)$ solves the following problem for all $I_i(g)$:

$$s_i^*(g)(I_i(g)) = \arg\max_{s_i \in R_+} E_{\tau,\eta}\left[\pi_i\left(s_i, s_{-i}^*(g)\right)|I_i(g)\right]. \tag{18.2}$$

The reaction function of firm i as a function of i's information structure is:

$$s_i^*(g)\,(I_i(g)) = \frac{1}{2\delta}\left(\beta + \gamma_s E\left[\tau_i|I_i(g)\right] - \varepsilon\sum_{j\neq i} E\left[s_j|I_i(g)\right]\right). \tag{18.3}$$

Firms' equilibrium strategies are affine in the observed signals:[2]

$$s_i^*(g)\,(I_i(g)) = a_i^g + \sum_{j\in I_i(g)} b_{ij}^g y_j, \qquad i = 1, 2, \ldots n. \tag{18.4}$$

[2] Standard results (see Radner, 1962 and Proposition 3.1 in Raith, 1996).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

The a_i^g and b_{ij}^g coefficients can be computed by solving the following system, which immediately points to the main forces at work within a given information structure:

$$a_i^g = \frac{1}{2\delta}\left(\beta - \varepsilon \sum_{j \neq i} a_j^g\right); \tag{18.5}$$

$$b_{ih}^g = \frac{1}{2\delta}\left(\gamma_s k_1^{ig} - \varepsilon\left(\sum_{j \in N_h \setminus \{i\}} b_{jh}^g + \sum_{z \notin N_i}\sum_{j \in N_z} k_2^{ig} b_{jz}^g\right)\right),\ \forall h \in N_i$$

The coefficients $k_1^{ig} = \frac{t}{p_s + (n_i^g - 1)p_n}$ and $k_2^{ig} = \frac{p_n}{p_s + (n_i^g - 1)p_n}$ describe the way in which a firm $i \in N$ in a network g uses its observed signals to update its beliefs. In particular, k_1^{ig} is applied to all $y_j \in I(g_i)$ to take the expectation of τ, while k_2^{ig} is applied to all $y_j \in I(g_i)$ to take the expectation of the signals y_h, for all $h \notin N_i$.

The β coefficients measure the sensitivity of equilibrium actions to the information received from a given source. For the above expression, we learn that under strategic substitutes (complements) the reaction of firm i to signal h is stronger (weaker) the less signal h is used by other firms. In the case of demand uncertainty, this can be understood as a local congestion effect: the more other firms use a signal, the less a firm wishes to use it. This echoes results from Morris and Shin's (2002) study of the use of information when both private signals and public signals are available to players. In our case, a signal is public only to agents in the neighbourhood of the firm that acts as the source of that signal. From the expression from the β coefficient we also learn that a firm reacts less to a signal that is used by other firms to infer something about the information held by firms they are not linked with. This effect goes through the correlation of signals, and is stronger the larger the k_2 coefficient.

3.2 Exchange of Information and Equilibrium Networks

The incentives of firms to form and maintain links are measured by the *ex ante* value of the equilibrium profits in the various networks that may form as a consequence. For a given network g and firm i, these are given by the expectation $E\pi_i(g)$ of the interim profits taken over all possible realization of the information $I_i(g)$ observed by i in g. Following Proposition 3.4 in Raith (1996), we can write the following:

$$E\pi_i(g) = E(\alpha_i(\tau)) + \delta\left(a_i^g\right)^2 + \beta_n \sum_{j \neq i} a_j^g + \delta Var\left(s_i^*(g)\right) \tag{18.6}$$

It can be shown that the difference in firm i's expected profit in the two information structures g and g' can be expressed as:

$$E\pi_i(g) - E\pi_i(g') = \delta\left[var(s_i^*(g) - var(s_i^*(g'))\right] \tag{18.7}$$

So, a firm's incentive to move from network g to network g' is measured by the change in the variability of its own equilibrium strategy. Since network structures can be highly asymmetric and complex in nature, the analysis of the incentives to form and delete links is conceptually

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

and computationally very complex. In the next section we report on some results that can be obtained in the simplified framework of independent signals.

3.2.1 Independent signals

The case of independent signals was studied in Gal-Or (1985), where each firm receives an imperfect signal of one piece of the demand intercept. This case can still be viewed as a special case of Raith's model, in which the states of the world τ_is are perfectly correlated but the (conditional) correlation of signals exactly compensate the natural correlation of signals through the state of the world.

The equilibrium parameters of firm i (see (18.5)) take the following simple form:

$$a_i^g = \frac{\beta}{2\delta + \varepsilon(n-1)}; \tag{18.8}$$

$$b_{ih}^g = \frac{\gamma_s t}{p_s(2\delta + n_h^g - 1)}, \ \forall h \in N_i$$

From (18.8) we note that for each signal j we have $b_{ij} = b$ for all i and h in N_j^g. Also, from (18.8) $a_i = a_h$ for all $i, h \in N$.

The next proposition provides necessary and sufficient conditions for a network to be pairwise stable when signals are independent:

Proposition 3 *A network g is pairwise stable if and only if both of the following conditions are verified:*

For all $ij \in g$:

$$\frac{1}{(2\delta + \varepsilon(n_j^g - 1))^2} \geq \frac{1}{(2\delta + \varepsilon(n_i^g - 2))^2} - \frac{1}{(2\delta + \varepsilon(n_i^g - 1))^2}; \tag{18.9}$$

$$\frac{1}{(2\delta + \varepsilon(n_i^g - 1))^2} \geq \frac{1}{(2\delta + \varepsilon(n_j^g - 2))^2} - \frac{1}{(2\delta + \varepsilon(n_j^g - 1))^2}; \tag{18.10}$$

For all $ij \notin g$:

$$\frac{1}{(2\delta + \varepsilon n_j^g)^2} > \frac{1}{(2\delta + \varepsilon(n_i^g - 1))^2} - \frac{1}{(2\delta + \varepsilon n_i^g)^2} \tag{18.11}$$

$$\textit{implies} \ \frac{1}{(2\delta + \varepsilon n_i^g)^2} < \frac{1}{(2\delta + \varepsilon(n_j^g - 1))^2} - \frac{1}{(2\delta + \varepsilon n_j^g)^2} \tag{18.12}$$

Note that the congestion effect discussed at the end of Section 3.1 translates here into simple and stark incentives to form a link ij, which only depends on the degrees of the nodes i and j, and on no other features of the network. In particular, when $\epsilon > 0$ (strategic substitutes), the gain in profit due to a link with node j decreases with the degree of j. It is indeed possible to determine two thresholds in the degree of a node j: the value $F(n_i)$ above which a node i of degree n_i would not maintain the link ij, i.e. if $n_j > F(n_i)$ inequality (18.9) is not satisfied; the

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

value $f(n_i)$ above which a node i of degree n_i would not form the new link ij; i.e. if $n_j > f(n_i)$ inequality (18.11) is not satisfied. It can also be shown that F and f are increasing in n_i, meaning that the incentives of node i to link with a given node j increase with the degree of i (see Lemma 1 in Currarini and Feri, 2015).

Define now $\mu = \frac{\varepsilon}{\delta}$ the degree of products' differentiation; the next proposition fully characterises the set of pairwise stable networks:

Proposition 4 *Let $n \geq 3$. If $\mu < 0$ the unique stable network is the complete one. If $\mu > 0$ the set of pairwise stable networks contains: the empty network, the complete network, and all networks made of $s \leq n - 3$ isolated nodes and $p \geq 1$ completely connected components of size $n_1 \geq 3, n_2, \ldots, n_p$ such that $n_i > f(n_{i-1})$ for all $i = 2, \ldots, p$.*

Remark 1 *The set of pairwise stable networks characterized in Proposition 4 is very large. However, Proposition 4 provides two precise qualitative predictions on how information is shared in equilibrium. First, information sharing is essentially organized in groups (the completely connected components), within which the transmission of information is equivalent to one in which firms publicly disclose their signal to all other firms in the group. This type of public disclosure, characterizing the traditional "contractual approach", is here obtained endogenously as a result of private and bilateral arrangements. Second, information-sharing groups must be of different size, to make sure that firms in different groups do not form links (in fact, it can be shown that firms with similar degrees link together) (Figure 18.1).*

We obtain a more narrow prediction for the case in which firms can agree on side payments that are contingent on information sharing. In this case, the formation of links that bridge two components is made easier by the sharing of individual gains, and at most one component of information-sharing firms can be compatible with stability:

Proposition 5 *If side payments are possible, the set of pairwise stable networks contains: the empty network, the complete network, all networks g made of one completely connected component h of size $n(h) \geq 3$ and $n - n(h)$ isolated nodes.*

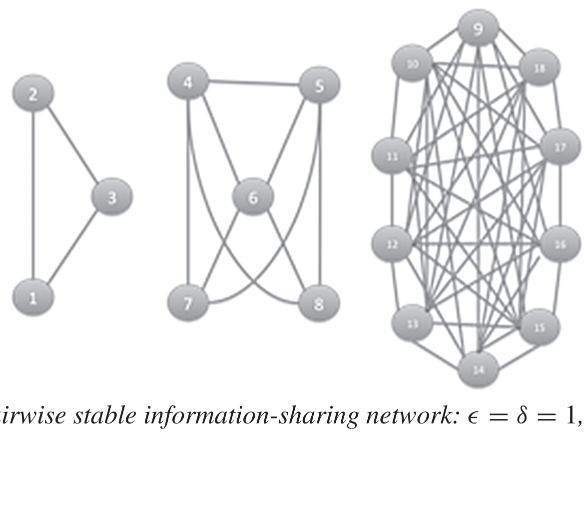


Figure 18.1 A pairwise stable information-sharing network: $\epsilon = \delta = 1$, $n = 18$

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

3.3 The Role of Signals' Correlation

One of the conclusions from the case of independent signals is that the empty network, characterized by no sharing of information, is always a stable outcome in Cournot competition with homogeneous goods and demand uncertainty. In this proposition we discuss the role of signals correlation in generating incentives to share information and, at the same time, to exclude some of the rivals from sharing.

The next proposition shows that in all common value situations (that is, independently of the degree of products differentiation), the empty network is not a pairwise stable structure (and, therefore, not a strongly pairwise stable structure), provided the number of firms in the market and the (conditional) correlation of signals is not too small:

Proposition 6 *Consider Raith's model of oligopolistic competition:*

(i) *If $\mu < \frac{2}{3}$ the empty network is not pairwise stable.*
(ii) *If $\frac{2}{3} < \mu < \frac{2}{1+\sqrt{2}}$ then there exists a p_n^{**} such that for all $p_n < p_n^{**}$ the empty network is not pairwise stable; otherwise (when $p_n > p_n^{**}$) there exists a finite number of firms $n^*(p_n)$ such that for all $n > n^*(p_n)$ the empty network is not pairwise stable.*
(iii) *If $\mu > \frac{2}{1+\sqrt{2}}$ there exists p_n^* and a finite value $n^*(p_n)$ such that for all $p_n > p_n^*$ and $n > n^*$ the empty network is not pairwise stable.*

Let us compare Proposition 6 with Raith's (1996) results for the contractual model (note here that when only two firms are in the market our model and Raith's model are equivalent). In point *(i),* values of $\mu < \frac{2}{3}$ are such that two duopolists would always pool their private information ($\frac{2}{3}$ is in fact $\frac{2}{n+1}$ for $n = 2$). Our result shows that these incentives remain when more firms are in the market. Points *(ii)* and *(iii)* cover situations in which two duopolists may or may not have the incentive to share information, depending on the level of the covariance of signals p_n. Point *(ii)* shows that when these incentives exist (low p_n), they do not vanish as we add firms to the market. More interestingly, when such incentives to bilaterally share information in a duopoly are absent (high p_n), they appear as we add more firms in the market. Finally, point *(iii)* refers to the range of parameters for which two duopolists would never share information, for any value of p_n. Here, it is shown that by adding firms in the market we can generate incentives for bilateral information sharing, provided the covariance p_n is large enough.[3]

To understand the forces at work in Proposition 6, consider again the incentives of two Cournot duopolists to share information. These are determined by two opposite effects on expected profits: the increased accuracy of firms' expectations (a positive effect) and the increased correlation of equilibrium strategies (a negative effect since strategies are substitutes). Unless products are very differentiated (ε positive but small), the second effect dominates the first. The crucial new element of the present proposition is that as we increase the number of firms, the bilateral exchange of information between firms i and j has the additional positive effect of improving the accuracy of the expectation of these two firms on the signal observed by the other firms in the market (and thereby on

[3] Note that the threshold levels of μ in Raith's paper are decreasing in n. Therefore, it is not possible that by adding firms in the market we pass from a situation where the empty network dominates the complete graph to a situation where the opposite is true.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

their equilibrium behaviour). This improved accuracy comes without the disclosure of any additional information to any of these other firms and, in this sense, at no cost. Moreover, this positive effect on profits is larger the larger the number of other firms in the market (from which the requirement on n in Proposition 6).

The result of Proposition 6 rules out the complete absence of information sharing in equilibrium (at least under certain conditions on p_n and n), but leaves open the question of whether stable networks exist in general. Proposition 7 below shows that the complete network is always pairwise stable, for all values of the parameters:

Proposition 7 *Let* $n \geq 3$. *The complete network is always a pairwise stable information structure.*

We conclude that some positive amount of information sharing is always compatible with pairwise stability (Proposition 7), and is always a feature of pairwise stable networks when p_n and n are large enough (in the sense made clear in Proposition 6). The result of Proposition 7 does not fully extend to the notion of strong pairwise stability. The explicit expression of expected profits when multiple links are severed from the complete network is quite complex and does not allow for a closed-form result for all parameters' values. However, numerical simulations suggest that there exist a threshold level of signals' correlation above which the complete network is strongly pairwise stable, and below which it is not. This result is obtained algebraically in the two polar cases of the common value: perfect substitutes ($\mu = 1$) and high differentiation ($\mu = 0$):

Proposition 8 *Let* $n \geq 3$. *If* $\mu = 0$ *the complete network is always strongly pairwise stable. If* $\mu = 1$ *and* p_n *is large enough the complete network is strongly pairwise stable.*

One final issue we wish to address is whether stable networks can be incomplete, with some, but not all, private information being shared. Example 1 presents a common value problem with four firms and homogeneous goods where, for a certain range of parameters, in a strongly pairwise stable network three firms exchange information, and a fourth firm is excluded (again, the algebraic derivations behind Example 1 are omitted and are available from the authors upon request):

Example 1 *Let* $n = 4$ *and* $\varepsilon = \delta$. *For* $p_n > 0.53 \cdot p_s$ *the complete network is strongly stable. For* $p_n < 0.62 \cdot p_s$ *and for* $p_n > 0.71 \cdot p_s$ *the network consisting of a fully connected component of three nodes and an isolated node is pairwise stable, and it is Nash pairwise stable for* $0.58 \cdot p_s < p_n < 0.62 \cdot p_s$ *and for* $p_n > 0.71 \cdot p_s$ *(Figure 18.2).*

We end this section by comparing the insight from Example 1 with Malueg and Tsutsui's (1996) results on stable sharing coalitions in the case of three firms. With only three firms on the market, the empty network would be a pairwise stable architecture and, for all $p_n < 0.68 \cdot p_s$, there are no empty strongly pairwise stable architecture.[4] This is in line with the results by Malueg and Tsutsui (1996), where no information sharing ever occurs when products are homogeneous. By adding a fourth firm in the market, we increase the profitability of bilateral

[4] For greater values of p_n the complete network is strongly stable.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

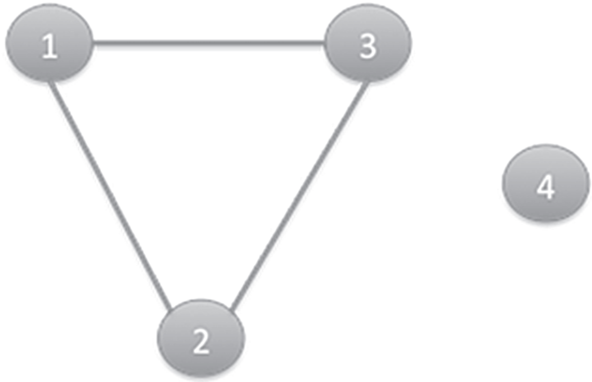


Figure 18.2 *A pairwise stable information-sharing network:* $n = 4$*,* $\varepsilon = \delta, p_n < 0.62 \cdot p_s$ *and* $p_n > 0.71 \cdot p_s$

agreements through the externality effect discussed after Proposition 6, so that the empty network becomes unstable for $p_n > 0.75 \cdot p_s$. Consider then the network consisting of a three-firm fully connected component and an isolated node; within the fully connected component, no firm has an incentive to sever one of its links and, for $p_n > 0.58 \cdot p_s$, no firm has an incentive to sever both its links.[5] Moreover, these firms have an incentive to link to the fourth firm if and only if $p_n < 0.71 \cdot p_s$; otherwise none of them has an incentive to link, having acquired enough information on the private signal of the fourth firm through the existing bilateral agreements. The fourth firm, instead, has an incentive to acquire additional information by forming a link if and only if $p_n > 0.62 \cdot p_s$. Therefore for high values of p_n the fourth firm is excluded from the information-sharing group.

3.4 Asymmetric Firms and the Emergence of Core–Periphery Structures

In this section we wish to discuss the role of asymmetry in the information-structure on the incentives to share and on the equilibrium information-sharing networks. To keep things simple, we limit the analysis to the case of independent signals, as we did in Section 3.2.1 for the symmetric case, and work with homegeneous goods, setting δ to 1.

We relax the assumption that signals are identically distributed, and allow the variances of signals p_s^i to differ across firms. The stability conditions of Proposition 3 are modified to account for this new source of heterogeneity: the network g is pairwise stable if and only if:

for all $ij \in g$:

$$\frac{p_s^j}{(n_j^g + 1)^2} \geq \frac{p_s^i}{(n_i^g)^2} - \frac{p_s^i}{(n_i^g + 1)^2} \tag{18.13}$$

$$\frac{p_s^i}{(n_i^g + 1)^2} \geq \frac{p_s^j}{(n_j^g)^2} - \frac{p_s^j}{(n_j^g + 1)^2} \tag{18.14}$$

[5] Note, that as shown above, with only three firms in the market the complete network is strongly stable for $p_n > 0.68 \cdot p_s$. By adding a fourth firm this threshold becomes smaller because it is more convenient to share information (or more costly to defect fron the information-sharing group).

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

for all $ij \notin g$:

$$\frac{p_s^j}{(2+n_j^g)^2} > \frac{p_s^i}{(1+n_i^g)^2} - \frac{p_s^i}{(2+n_i^g)^2} \tag{18.15}$$

$$\text{implies } \frac{p_s^i}{(2+n_i^g)^2} < \frac{p_s^j}{(1+n_j^g)^2} - \frac{p_s^j}{(2+n_j^g)^2} \tag{18.16}$$

We see that, given the degrees n_i^g and n_j^g, the incentive of i to sever the link ij increases with the ratio of variances $\frac{p_s^i}{p_s^j}$ (conditions (18.13)–(18.14)) and the incentive of i to form the link ij decreases with $\frac{p_s^i}{p_s^j}$ (condition (18.15)). This effect can be understood in terms of the additional variability of i's equilibrium quantity coming from the link ij, remembering that (expected) profits are correlated with the variability of own equilibrium strategy. The higher the term p_s^j, the higher the additional variability of i's quantity due to the link ij, and the higher the informational "value" of j's signal for firm i. Similarly, the higher the term p_s^i, the lower the incentive of firm i to form the link ij; this is because it is more costly to share a signal with higher variance with one additional firm. Again, a high value of p_s^i therefore reflects a high informational value of i's signal.

Note that in this setting of heterogeneous variance, a firm with high variance may not wish to maintain a link (or to form a new one) with another firm with same degree but lower variance. As a consequence, while the empty network is always a pairwise stable information structure (as was proved for the case of independent and identically distributed [i.i.d.] signals), the complete network may fail to be stable when firms have significant heterogeneity in variances. However, as the next proposition shows, this can only happen when the number of firms is small:

Proposition 9 *(i) The empty network is pairwise stable for all distributions of variances, even if side payments are possible; (ii) there exist configurations of variances for which the complete network is not pairwise stable; (iii) for every configuration of variances, there exists a finite number of firms $\bar{n}$ such that for all $n \geq \bar{n}$ the complete network is pairwise stable.*

The intuition of this result is clear: when the degree of two nodes increase, their difference in variances becomes less and less relevant in the stability conditions (18.13)–(18.14).

Since signals with large variance possess higher informational value, the incentive to link to firms observing such signals may remain high even when these firms have already a large degree. We can therefore envisage stable architectures in which firms with large variance have larger degrees than firms with low variance. Among such architectures, we will focus on two classes: networks made of a collection of completely connected components (as in the case of i.i.d. signals) and core–periphery networks.

We next turn to the existence of pairwise stable networks with non–completely connected components. We show that a special class of incomplete architectures, usually referred to as "core–periphery networks", can be pairwise stable for suitable distributions of variances. In more detail, core–periphery networks present a dense set of interconnected nodes – the

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Figure 18.3 *A core–periphery pairwise stable information-sharing network:* $n = 5$, $p_s^1 = 1$, $p_s^2 = p_s^3 = \frac{1}{5}$, $p_s^4 = p_s^5 = \frac{1}{2}$.

core – each linked with all nodes in the network, and sets of peripheral nodes that are internally connected and are linked with the core nodes (see Figure 18.3). Formally, a core–periphery network g consists of a set $\{g_1, g_2, \ldots .g_H\}$ of fully connected subnetworks, such that $i \in g_k$ and $j \in g_m$ implies that $ij \notin g$ for all $k \neq m$ such that $k \in \{2, 3, \ldots, H\}$ and $m \in \{2, 3, \ldots, H\}$, and such that $i \in g_1$ and $j \in g_k$ implies $ij \in g$ for all $k = 1, 2, 3, \ldots, H$. We call the subnetwork g_1 *core (*with size n_c), and the subnetworks $\{g_2, \ldots .g_H\}$ *peripheral planets*. We define a *symmetric* core–periphery network as one in which all peripheral planets have the same size $n_p \geq 1$. Also, we say that planets are *consecutive* in variances if planets can be obtained as a consecutive partition of the set of peripheral nodes ordered with respect to variance.

Proposition 10 provides two qualitative features of pairwise stable symmetric core–periphery networks: peripheral firms are organized in groups that are consecutive in variance, and core firms have larger variance than peripheral firms:

Proposition 10 *Every symmetric pairwise stable core–periphery network is such that peripheral planets are consecutive in variances. Moreover, for each given size n_p, there exists a finite n' such that if $n > n'$ then every symmetric pairwise stable core–periphery network is such that* $\min_{i \in g_1} p_s^i > \max_{j \in g \setminus g_1} p_s^j$.

Intuitively, core firms observe signals that are publicly observed, and therefore have lower informational value. These signals are "desirable" only if they have large variance, from which the second result in Proposition 10. An example of pairwise stable core–periphery network is the following:

Example 2 *Consider a network with five nodes: node 1 is the "core" node, while the two peripheral components are* $\{23\}$ *and* $\{45\}$*. Variances are* $p_s^1 = 1$, $p_s^2 = p_s^3 = \frac{1}{5}$, $p_s^4 = p_s^5 = \frac{1}{2}$*. Relevant stability conditions (18.15)–(18.16) (for links 12, 15 and 34, respectively) are satisfied:*

$$\frac{1}{36} + \frac{1}{5}\frac{1}{16} \geq \frac{1}{25};\ \frac{1}{36} + \frac{1}{2}\frac{1}{16} \geq \frac{1}{25};\ \left(\frac{1}{2} + \frac{1}{5}\right)\frac{1}{25} \leq \frac{1}{5}\frac{1}{36}.$$

4 CONCLUSIONS

In this chapter we have reviewed the main results and insights coming from theoretical research on the incentives to share information in oligopolistic markets. While most of the

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

literature has focused on the formation of industry-wide agreements to either disclose or share information, we have devoted a substantial space to recent contributions that consider the possibility that firms establish sharing agreements with selected partners, excluding other competitors from their private information. All the contributions covered in this survey assume that firms agree to share information at the *ex ante* stage, that is, before getting to know their private signals. This approach aims at describing the incentives to establish long-term agreements, in which information is repeatedly shared and in which the decision to share cannot be made contingent on what the realized information is. In reality, however, firms may decide to share information conditionally to the private signal they receive, that is, in certain situations. This case goes under the name of *interim* model, and has not received much attention in the literature. We believe that more research is needed here, that formally studies the signalling role of non-disclosure decisions, to determine under which condition unravelling of information occurs and full disclosure is guaranteed.

Despite the considerable amount of theoretical work on the subject, there is little empirical evidence on information sharing in real world oligopolistic markets. If anything, the theoretical insights have been used as evidence of collusion in quantity and/or prices where direct evidence of collusion was missing. The lack of empirical research is due to the difficulty in obtaining data that serve as good proxy of information sharing in a world where sharing is itself illegal. This motivates future effort in experimental research in a laboratory-controlled environment. While there exists some experimental research on the use of information in games with complementarities and signal with different degrees of publicness (see Cornand and Heinemann, 2014), experiments on information sharing seem like a very promising avenue of research.

REFERENCES

Bernheim, D. and M. Whinston (1987), Coalition-proof Nash equilibria. II. Applications, *Journal of Economic Theory*, **42**, 13–29.

Bloch, F. and M. Jackson (2006), Equilibrium definitions in network formation games, *International Journal of Game Theory*, **34**, 305–318.

Clarke, R. (1983), Collusion and the incentives for information sharing, *Bell Journal of Economics*, **14**, 383–394.

Cornand, C. and F. Heinemann (2014), Measuring agents' reaction to private and public information in games with strategic complementarities, *Economics*, **17**, 61–77.

Currarini, S. and F. Feri (2007), Bilateral information sharing in oligopoly, *University Ca' Foscari of Venice, Department of Economics Research Paper* 21/07.

Currarini, S. and F. Feri (2015), Information-sharing networks in linear quadratic games, *International Journal of Game Theory*, **44**, 701–732.

Fried, D. (1984), Incentives for information production and disclosure in a duopolistic environment. *Quarterly Journal of Economics*, **99**, 367–381.

Gal-Or, E. (1985), Information sharing in oligopoly, *Econometrica*, **53**, 329–343.

Gal-Or, E. (1986), Information transmission – Cournot and Bertrand equilibria, *Review of Economic Studies*, **53**, 85–92.

Jackson, M.O. and A. Wolinsky (1996), A strategic model of social and economic networks, *Journal of Economic Theory* **71**, 44–74.

Jansen, J. (2005), Information acquisition and strategic disclosure in oligopoly, mimeo, Social Science Research Center Berlin (WZB) and Humboldt University, Berlin.

Kirby, A.J. (1988), Trade associations as information exchange mechanisms, *RAND Journal of Economics*, **19**, 138–146.

Kühn, K.U. and X. Vives (1995), *Information Exchanges among Firms and their Impact on Competition*, Luxembourg: Office for Official Publications for the European Community.

Li, L. (1985), Cournot oligopoly with information sharing, *Rand Journal of Economics*, **16**, 521–536.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Malueg, D. and S.O. Tsutsui (1996), Coalition-proof information exchanges, *Journal of Economics*, **63**, 259–278.
Morris, S. and H.S. Shin (2002), Social value of public information, *The American Economic Review*, **92**, 1521–1534.
Novshek, W. and H. Sonnenschein (1982), Fulfilled expectations Cournot duopoly with information acquisition and release, *Bell Journal of Economics*, **13**, 214–218.
Okuno-Fujiwara, M., A. Postlewaite and K. Suzumura (1990), Strategic information revelation, *Review of Economic Studies*, **57**, 24–47.
Radner, R. (1962), Team decision problems, *Annals of Mathematical Statistics*, **33**, 857–881.
Raith, M. (1996), A general model of information sharing in oligopoly, *Journal of Economic Theory*, **71**, 260–288.
Sakai, Y. (1985), The value of information in a simple duopoly model, *Journal of Economic Theory*, **36**, 36–54.
Sakai, Y. and Yamato, T. (1989), Oligopoly, information and welfare, *Journal of Economics*, **49**, 3–24.
Shapiro, C. (1986), Exchange of cost information in oligopoly, *Review of Economic Studies*, **53**, 433–446.
Van Zandt, T. and X. Vives (2006), Monotone equilibria in Bayesian games of strategic complementarities, *Journal of Economic Theory*, **134**, 330–360.
Vives, X. (1985), Duopoly Information Equilibrium: Cournot and Bertrand, *Journal of Economic Theory*, **34**, 71–94.

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Index

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham

Luis C. Corchón and Marco A. Marini - 9781785363276
Downloaded from Elgar Online at 02/24/2018 12:47:17PM
via University of Durham