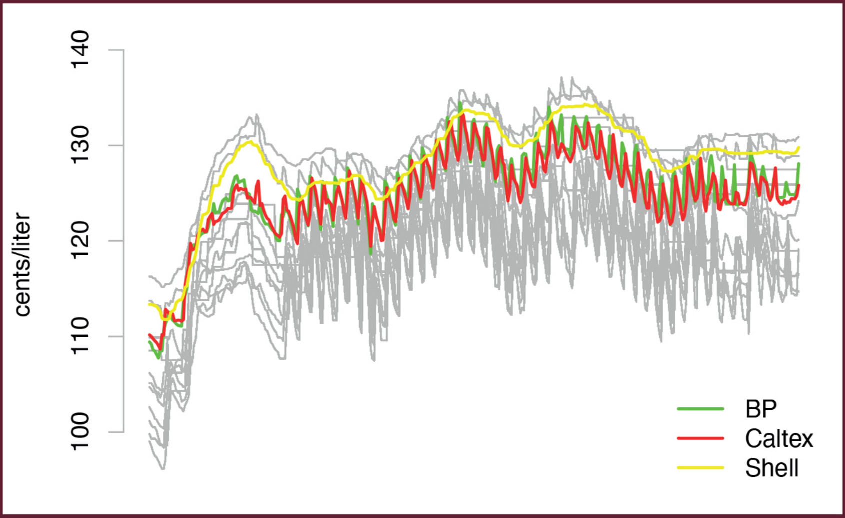
SERIES ON STATISTICS IN BUSINESS AND ECONOMICS

GAME THEORY FOR APPLIED ECONOMETRICIANS

Data Analytics with R



Christopher P. Adams

A Chapman & Hall Book

CRC Press
Taylor & Francis Group

Game Theory for Applied Econometricians

Over the last 30 years the practice and use of game theory has changed dramatically, yet textbooks continue to present game theory with algebraic formalism and toy models. This book, on the other hand, illustrates game theory concepts using real-world data and analyses problems with real policy implications. The focus is on applying current learning to real world problems by providing an introduction to game theory and econometric analysis based on game theoretic principles using the computer language R.

The book covers the standard topics of an introductory game theory course including dominant strategies, Nash equilibrium and Bayes Nash equilibrium. It layers on top of this an approach to statistics and econometrics called Structural Modeling. In this approach, key parameter estimates rely upon game theoretic analysis. The real-world examples used to illustrate these concepts vary in scope and include an analysis of bargaining between hospitals and insurers, equilibrium entry of retail tire stores, bid rigging in timber auctions and contracts in 19th century whaling.

This book is aimed at the general reader with the equivalent of a bachelor's degree in economics, statistics or some more technical field. The book could be used as a text for an upper level undergraduate course or a lower level graduate course in economics or business.

Christopher P. Adams was born and raised in Melbourne Australia and is a lifelong Carlton supporter. Chris received his PhD in economics from the University of Wisconsin - Madison. He has taught at the University of Vermont, Dartmouth College, University Maryland and Johns Hopkins. He has 17 years experience working in merger regulation and antitrust for the US Federal Trade Commission. He is currently a Principal Analyst at the US Congressional Budget Office. Chris' work and research focus on econometrics, empirical industrial organization, pharmaceutical innovation and auctions. His work has been published in various academic journals including *The Econometrics Journal*, *Health Affairs*, *Health Economics*, *Marketing Science* and *Economics Letters*. Chris is a colon cancer survivor and a research patient advocate with ECOG-ACRIN. Most importantly, he is a father to CJ and husband to Deena.

CHAPMAN & HALL/CRC Series on Statistics in Business and Economics

Recently Published Titles

Empirical Research in Accounting
Tools and Methods
Ian D. Gow and Tongqing Ding

Game Theory for Applied Econometricians
Data Analytics with R
Christopher P. Adams

For more information about this series, please visit: https://www.routledge.com/Chapman-and-HallCRC-Series-on-Statistics-in-Business-and-Economics/book-series/CHSBE

Game Theory for Applied Econometricians

Data Analytics with R

Christopher P. Adams

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business
A CHAPMAN & HALL BOOK

Designed cover image: Christopher P. Adams

First edition published 2025
by CRC Press
2385 NW Executive Center Drive, Suite 320, Boca Raton FL 33431

and by CRC Press
4 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

CRC Press is an imprint of Taylor & Francis Group, LLC

ISBN: 978-1-032-39833-4 (hbk)
ISBN: 978-1-032-39835-8 (pbk)
ISBN: 978-1-003-35160-3 (ebk)

DOI: 10.1201/b23262

Typeset in CMR10 font
by KnowledgeWorks Global Ltd.

Publisher's note: This book has been prepared from camera-ready copy provided by the authors.

Contents

Introduction **xiii**

I Static Games of Complete Information **1**

1 Games **3**
1.1 Introduction . . . 3
1.2 Some Strategic Situations . . . 3
1.2.1 Aliens Attack! . . . 4
1.2.2 Can AI Cope with Traffic? . . . 4
1.3 End of Skyjacking . . . 5
1.3.1 More Security? . . . 5
1.3.2 Hijacking Data using **R** . . . 5
1.3.3 United Arlines Flight 93 . . . 6
1.4 A Game . . . 7
1.4.1 Definitions . . . 7
1.4.2 Examples . . . 8
1.5 Prisoner's Dilemma . . . 10
1.5.1 The Game . . . 10
1.5.2 What is the Best Outcome? . . . 11
1.5.3 What should You Do? . . . 11
1.5.4 Does This Make Sense? . . . 12
1.6 Empirical Analysis: *Friend or Foe* using **R** . . . 12
1.6.1 Trust Box . . . 12
1.6.2 Data Analysis in **R** . . . 13
1.6.3 Analysis in **R** . . . 14
1.7 Discussion and Further Reading . . . 17

2 Nash Equilibrium **18**
2.1 Introduction . . . 18
2.2 What is a Nash Equilibrium? . . . 19
2.2.1 Dominance . . . 19
2.2.2 Prisoner's Dilemma . . . 20
2.2.3 Coordination Game . . . 21
2.3 Nash Equilibrium . . . 22
2.3.1 Definition . . . 22

2.3.2 Algorithm for Finding Nash Equilibrium 22
2.3.3 Prisoner's Dilemma Game 23
2.3.4 Coordination Game 23
2.4 Entry Games 24
2.4.1 Bresnahan and Reiss 24
2.4.2 Two Firm Entry Game 25
2.4.3 Many Firm Entry Game 27
2.4.4 Nash Equilibrium 28
2.4.5 Fixed Cost of Entry 28
2.4.6 Equilibrium Number of Firms 28
2.4.7 Simulation of Entry Game in **R** 29
2.5 Empirical Analysis: Tire Markets using **R** 30
2.5.1 Data 31
2.5.2 Structural Model 32
2.5.3 Entry Estimator 33
2.5.4 Estimation in **R** 33
2.5.5 Structural Estimation 35
2.5.6 Policy Simulation in **R** 36
2.6 Discussion and Further Reading 38

3 Oligopoly 39
3.1 Introduction 39
3.2 Cournot's Model 40
3.2.1 Two Firm Model 40
3.2.2 Best Response 40
3.2.3 Nash Equilibrium 41
3.2.4 Cournot Model with N Firms 42
3.2.5 Cournot Model in **R** 43
3.2.6 Solve for the Nash Equilibrium with **R** 43
3.2.7 Simulation of Cournot Model in **R** 44
3.3 Bertrand's Model 45
3.3.1 Two-Firm Game 45
3.3.2 Nash Equilibrium 46
3.4 Hotelling's Model 47
3.4.1 Hotelling's Line 47
3.4.2 Differentiated Goods Game 48
3.4.3 Best Response 48
3.4.4 Nash Equilibrium 49
3.5 Empirical Analysis: Hamburger Competition with **R** 49
3.5.1 Data 49
3.5.2 Estimation (Part 1) 51
3.5.3 Empirical Equilibrium 51
3.5.4 Estimator 53
3.5.5 Estimator in **R** 54
3.5.6 Distances 55

3.5.7 Ownership . . . 56
3.5.8 Ownership of Outlets . . . 57
3.5.9 Estimation (Part 2) . . . 58
3.5.10 Goodness of Fit . . . 59
3.5.11 A Merger of McDonald's Outlets . . . 59
3.6 Discussion and Further Reading . . . 61

4 Empirical Entry Games **63**
4.1 Introduction . . . 63
4.2 Empirical Model . . . 63
4.2.1 Single Firm Entry Model . . . 64
4.2.2 Multiple Firm-Independent Entry Model . . . 65
4.2.3 Entry Game . . . 67
4.3 Empirical Analysis: Bookstore Entry with **R** . . . 67
4.3.1 Single Firm Entry Model in **R** . . . 68
4.3.2 Multiple Firm-Independent Entry in **R** . . . 70
4.3.3 Entry Game in **R** . . . 71
4.4 Empirical Analysis: Bookstore Entry using **R** . . . 72
4.4.1 Data . . . 73
4.4.2 Estimation of Single Firm Entry . . . 73
4.4.3 Estimation of Two Firm Entry Model . . . 75
4.4.4 Estimation of the Entry Game Model . . . 76
4.4.5 Model Fit . . . 76
4.5 Policy Analysis using **R** . . . 77
4.6 Discussion and Further Reading . . . 79

5 Mixed Strategies **80**
5.1 Introduction . . . 80
5.2 Zero-Sum Games . . . 80
5.2.1 Rock-Paper-Scissors . . . 80
5.2.2 Nash Equilibrium . . . 81
5.2.3 Mixed Strategies . . . 82
5.2.4 Penalty Kicks . . . 82
5.2.5 Algorithim for Mixed Strategy Nash Equilibrium . . . 83
5.3 Rock-Paper-Scissors . . . 84
5.3.1 Mixed Strategy Nash Equilibrium . . . 85
5.3.2 More Complicated Version . . . 85
5.3.3 Using **R** to Determine when Player 1 is Indifferent . . . 86
5.3.4 Solving for MSNE using **R** . . . 87
5.4 Empirical Analysis: Penalty Kicks using **R** . . . 89
5.4.1 Penalty Kick Game . . . 89
5.4.2 Data . . . 90
5.4.3 GMM Estimator . . . 91
5.4.4 GMM Estimator in **R** . . . 92
5.4.5 Difference By Footedness . . . 94

5.5 Multiple Equilibria . . . 95
5.5.1 Coordination Game . . . 96
5.5.2 Bookstore Entry . . . 96
5.5.3 MSNE Estimator in **R** . . . 97
5.6 Empirical Analysis: Bookstore Entry with MSNE using **R** . . . 98
5.7 Discussion and Further Reading . . . 99

II Dynamic Games of Complete Information 101

6 Dynamic Games 103
6.1 Introduction . . . 103
6.2 Extensive Form . . . 104
6.2.1 Coordination Game . . . 104
6.2.2 Strategies Revisited . . . 105
6.2.3 Nash Equilibrium of the Coordination Game . . . 105
6.2.4 Game Tree . . . 106
6.3 Subgame Perfection . . . 107
6.3.1 Definition . . . 108
6.3.2 Coordination Game . . . 108
6.3.3 Empirical Entry Game . . . 109
6.3.4 Equilibrium in Entry Game . . . 109
6.3.5 Empirical Implications . . . 110
6.4 Empirical Analysis: Bookstore Entry with Subgame Perfection in **R** . . . 110
6.4.1 SPNE Estimator . . . 110
6.4.2 SPNE Estimates . . . 111
6.5 Discussion and Further Reading . . . 111

7 Repeated Games 113
7.1 Introduction . . . 113
7.2 Repeated Prisoner's Dilemma . . . 114
7.2.1 Prisoner's Dilemma . . . 114
7.2.2 Normal Form Representation . . . 115
7.2.3 Finitely Repeated Game . . . 115
7.2.4 Infinitely Repeated Prisoner's Dilemma . . . 116
7.2.5 Cooperation as a Nash Equilibrium . . . 117
7.3 Bertrand Competition . . . 118
7.3.1 Two Firm Model . . . 118
7.3.2 J Firm Model . . . 118
7.3.3 Nash Equilibrium in **R** . . . 119
7.4 Empirical Analysis: Retail Gasoline Pricing using **R** . . . 120
7.4.1 Perth Gas Price Data . . . 120
7.4.2 Estimating Parameters . . . 120
7.4.3 Parameter Estimates . . . 122

7.5 Repeated Oligopoly . . . 124
7.5.1 Collusive Equilibrium . . . 124
7.5.2 Identifying Collusion . . . 125
7.5.3 Choosing Super Markups . . . 126
7.5.4 Estimating Collusive Prices in **R** . . . 126
7.6 Empirical Analysis: Collusion in Perth Gas Stations using **R** . . . 127
7.6.1 Super Markups . . . 128
7.6.2 Analyzing Mergers with Collusion . . . 128
7.7 Discussion and Further Reading . . . 130

8 Bargaining **133**
8.1 Introduction . . . 133
8.2 Ultimatum Game . . . 134
8.2.1 The Game . . . 134
8.2.2 Game Tree . . . 135
8.2.3 Normal Form . . . 135
8.2.4 Nash Equilibrium . . . 136
8.2.5 Subgame Perfection . . . 138
8.3 Empirical Analysis: Ultimatum Game in India using **R** . . . 138
8.3.1 Data . . . 138
8.3.2 Equilibrium Play or Fairness? . . . 139
8.3.3 Do Stakes Matter? . . . 140
8.4 Two Period Alternating Offers Game . . . 143
8.4.1 The Game . . . 143
8.4.2 Game Tree . . . 144
8.4.3 Subgame Perfection . . . 144
8.5 Infinite Alternating Offers Game . . . 145
8.5.1 The Game . . . 145
8.5.2 Game Tree . . . 146
8.5.3 Subgame Perfection . . . 147
8.5.4 Game Ends in Period 1 . . . 148
8.5.5 Infinite Alternating Offers Game in **R** . . . 148
8.6 Nash Bargaining Model . . . 150
8.6.1 The Model . . . 150
8.6.2 Nash Solution using **R** . . . 151
8.7 Modeling Hospital Competition and Pricing . . . 151
8.7.1 Bargaining Model . . . 152
8.7.2 Demand for Hospitals . . . 153
8.7.3 Willingness-to-Pay . . . 153
8.8 Empirical Analysis: Hospital Competition with **R** . . . 154
8.8.1 Data . . . 155
8.8.2 Pricing and WTP . . . 156
8.8.3 Mergers and Willingness To Pay . . . 158
8.9 Discussion and Further Reading . . . 159

III Static Games of Incomplete Information 161

9 Bayes Nash Equilibrium 163
9.1 Introduction . . . 163
9.2 A Bayesian Game . . . 164
9.2.1 A Grading Game . . . 164
9.2.2 Grading Game Simulation using **R** . . . 164
9.2.3 Definitions . . . 166
9.2.4 The Game . . . 167
9.2.5 Course Grading Game . . . 167
9.2.6 Equilibrium . . . 168
9.2.7 Equilibrium of Course Grade Game . . . 169
9.3 Empirical Entry Game . . . 169
9.3.1 The Game . . . 169
9.3.2 Equilibrium . . . 170
9.3.3 Estimating Entry Games . . . 171
9.3.4 Estimating Entry Games in **R** . . . 172
9.4 Empirical Analysis: Mega Bookstore Entry (Again) using **R** . . . 173
9.4.1 Estimates . . . 173
9.4.2 Policy . . . 173
9.5 Discussion and Further Reading . . . 175

10 Auctions 176
10.1 Introduction . . . 176
10.2 Sealed Bid Auctions . . . 177
10.2.1 Sealed Bid Model . . . 178
10.2.2 Bayes Nash Equilibrium . . . 178
10.2.3 Sealed Bid Simulation in **R** . . . 179
10.2.4 Sealed Bid Estimator . . . 180
10.2.5 Sealed Bid Estimator in **R** . . . 181
10.3 English Auctions . . . 183
10.3.1 Second Price Auction Game . . . 184
10.3.2 Bayes Nash Equilibrium . . . 184
10.3.3 Order Statistics . . . 185
10.3.4 Identifying the Value Distribution . . . 186
10.3.5 English Auction Estimator . . . 186
10.3.6 English Auction Estimator in **R** . . . 187
10.4 Empirical Analysis: Testing the Rationality of Loggers using **R** 188
10.4.1 Timber Data . . . 188
10.4.2 Sealed Bid Auctions . . . 189
10.4.3 Comparing English Auctions to Sealed Bid Auctions . . . 190
10.5 Empirical Analysis: Testing for Collusion using **R** . . . 192
10.5.1 A Test of Collusion . . . 192
10.5.2 "Large" English Auctions . . . 193
10.5.3 Large English Auction Estimator . . . 193

10.5.4 Large English Auction Estimator in **R** . . . 194
10.5.5 Evidence of Collusion . . . 194
10.6 Discussion and Further Reading . . . 196

11 Auctions with Affiliated Valuations **197**
11.1 Introduction . . . 197
11.2 Auctions with Common Values . . . 197
11.2.1 Simple Model . . . 198
11.2.2 Winner's Curse . . . 198
11.2.3 Bayes Nash Equilibrium . . . 199
11.2.4 Common Value Auction in **R** . . . 200
11.2.5 Simulation of Common Value Auction using **R** . . . 201
11.2.6 Estimator for Common Values Auctions . . . 203
11.2.7 Common Values Estimator in **R** . . . 204
11.3 Empirical Analysis: Signal Distribution from OCS Auctions using **R** . . . 205
11.3.1 Data . . . 206
11.3.2 Estimating the Signal Distribution . . . 207
11.4 Auctions with Coalitions . . . 209
11.4.1 The Benefit of Coalitions . . . 210
11.4.2 Estimating Coalitions in **R** . . . 211
11.4.3 Policy Simulation . . . 212
11.5 Discussion and Further Reading . . . 215

IV Dynamic Games of Incomplete Information **217**

12 Moral Hazard **219**
12.1 Introduction . . . 219
12.2 Principal-Agent Game . . . 220
12.2.1 Simple No Contract Game . . . 220
12.2.2 Bayes Nash Equilibrium . . . 221
12.2.3 Simple Contract Game . . . 222
12.2.4 More Complicated Contract Game . . . 223
12.2.5 Bayes Nash Equilibrium . . . 224
12.2.6 Parameterized Model . . . 224
12.2.7 Simulation with **R** . . . 225
12.3 Empirical Analysis: Whaling Corporations in the 19th Century using **R** . . . 226
12.3.1 Whaling Data . . . 227
12.3.2 Regressions . . . 227
12.3.3 Calibrating the Model . . . 230
12.4 Discussion and Further Reading . . . 232

13 Adverse Selection **233**
13.1 Introduction . . . 233

13.2 Akerlof's Lemons 233
13.2.1 The Lemons Market 234
13.2.2 Lemons Game 234
13.2.3 Bayes Nash Equilibrium 235
13.3 Insurance 235
13.3.1 Model 236
13.3.2 Game with Unknown Types 236
13.3.3 Bayes Nash Equilibrium 237
13.3.4 Simulation using **R** 237
13.3.5 Mandatory Insurance 239
13.4 Empirical Analysis: Health Insurance using **R** 240
13.4.1 Willingness to Pay 240
13.4.2 Premiums 240
13.4.3 Tax Policy 243
13.5 Discussion and Further Reading 245

Bibliography **247**

Subject Index **253**

Author Index **257**

Introduction

The Research Assistant

You are back in DC! Yes. It is as hot and humid as it was for your internship, maybe even hotter and humiditier. You managed to secure a job as an "RA" at the US Federal Trade Commission. You are in the Bureau of Economics in AT1. You have an awesome place in Navy Yard, not far from the Nats stadium, the Wharf and the home for DC United. A mate from college is your room mate and they started at the Fed last week. Your career is about to begin!

First Day

What the heck is he talking about? A colleague, was it John or Jeff, maybe Dave, has been talking 100 mph for last few minutes and you are just catching snippets. "We need to estimate WTP." "Can you believe they are using Elzinga-Hogarty." "I hope Alabama joins the suit." "Bill Town is great." You have been nodding but you quickly lost the thread of the conversation.

You decided to go with a suit, which was definitely a mistake. Even though you were hardly outside, you are sweating buckets and no one else seems to be in a suit, not even the managers. The managers are lovely and seemed very excited that you were starting. They liked that you had coding experience in **R** and had done some cool work in your internship on minimum wages. Apparently, you are going to jump straight into a case.

John/Jeff/maybe-Dave is explaining the case to you. It seems to have something to do with hospitals in Alabama. Your screen is full of data, there are columns with things you recognize like zip codes, age, and then things like DRG which you have no idea about. Apparently, this data is from a payer, although you are not sure what that is. The rows are claims. There is a column with the price, but they seem wrong. The numbers are enormous, 150008, 25020, 83251. Is this dollars? You were in hospital a few weeks back getting some stitches after an incident playing kick ball. You are pretty sure it was $25.00

Your job is to determine the relationship between price and WTP. You just wish you knew what WTP stood for.

Second Day

Yesterday you were able to get **R** and `RStudio` set up and get access to the payer data. You were able to calculate WTP for the hospitals and you showed Dave the results.

OK. Now we need to calculate WTP post-merger. You responded that you didn't think you had the data, you were pretty sure that there was no indicator for merger in the data. Dave laughed out loud. Then he saw the expression on your face, caught himself, and stated matter-of-factly, no we have to simulate the merger.

Simulate the merger? How would we know what would happen to prices after the hospitals merge? How do we know how prices are determined now? From what you could gather so far there were three hospitals in the city and five insurance companies serving beneficiaries in the area. The prices in the data were determined by bargaining between these hospitals and the insurance companies. You search your memory back to microeconomics, you remember the class on monopoly pricing, when the seller had market power to determine price. Do these hospitals have market power? You remember something in the text book about monopsony, when the buyer has market power to determine price. Do the insurers have market power? Do both the hospitals and the insurers have market power?

The Book

This is an empirical game theory book. Traditionally, game theory is presented as a theoretical subject. Generally, applications are discussed but there are no explicit empirical applications. Yet, since the 1990s, game theory has been at the heart of empirical analysis of competition and markets in the economics sub-field of industrial organization. In economics, this layering of theory on to empirical analysis is called **structural econometrics**. This book is focused on the game theory, not necessarily the econometrics.

What Does it Cover?

The book aims to provide an introduction game theory, a mathematical approach to understanding economic relationships. The goal is for the reader to understand what a game is and how the mathematics works. The reader will be able to create a game that explains behavior of economic actors observed in her data. She will be able to ask *how* questions. She will be able to see how parameters of the game relate to characteristics of the data. She may be able to use the game to ask *what if* questions. What would the economic actors observed in the data do if the world was different? What if the government introduced a new policy? What if technology changed. What if the hospitals merged?

The book covers standard game theory concepts such as **normal form games**, **exentensive form games**, **Nash equilibrium**, **mixed strategy Nash equilibrium**, and **subgame perfection**. It also covers standard empirical methods such as **linear regression**, **two-way fixed effects**, **logit models**, and **maximum likelihood estimation**. But it covers a number of concepts that are important to the intersection of game theory and data analysis such as **generalized method of moments** and **two-step estimators**.

What is the Approach?

The book teaches game theory through code, in particular, it will use the scripting language **R**. It is not primarily aimed at teaching **R**. Rather, it is primarily aimed at teaching game theory. This idea of using computer programming as a tool of instruction goes back to at least Seymour Papert and MIT's AI lab in the 1970s.[1] Papert helped develop a programming language called `Logo`. The goal of `Logo` was to teach geometry by programming how a **turtle** moves around the screen. You may have used one of the offspring of `Logo`, such as `Scratch` or `Lego Mindstorms`.

The book uses Papert's ideas to teach game theory. You will learn the math of the game or estimation method and then how to program that game or estimation method. The book makes particular use of the computer's ability to simulate data. This allows us to experiment with more complicated and realistic games than is possible with pen and paper.

The book is written in `RStudio` using `Sweave`. `Sweave` allows LaTeX to be integrated into **R**. LaTeX is a free type-setting language that is designed for writing math. Much of the code that is used in the book is actually presented in the book. Sometimes it is more practical to create a data set outside the of book. In those cases, the data and the code that created the data are available here `https://github.com/christopherpadams/EmpiricalGameTheory`. In a couple of other cases, the preferred code does not produce nice output for the book, so it is left out.

What is How Analysis?

Often we want to understand patterns we see in the data. How are prices related to the number of firms in the market? Why do firms enter some markets and not others?

For those questions, we can use game theory to guide our thinking and our analysis. A game theoretic model of firm entry can help us understand why some markets had both Borders and Barnes & Noble, other markets had just one, and many more have none.

[1] `https://el.media.mit.edu/logo-foundation/`

What is What If Analysis?

Game theory is also important in *what if* analysis. In those analysis, we want to understand how behavior will change when faced with a situation not observed in the data. What would have happened if Borders and Barnes & Noble would have merged in the mid-1990s? Would there have been more or fewer mega bookstores? Should the federal government allow bidders to collude in oil drilling auctions? We can estimate the parameters of game theoretic model using observed data and then make changes to the model to simulate what would happen in the case that is not observed in the data. We can then run the model and predict the outcome. We can simulate a merger between Borders and Barnes & Noble in the mid-1990s or what happens to bids on oil drilling leases when bidders are allowed to collude.

What About the Real World?

The book presents interesting and important questions. The book presents an analysis of competition between various types of firms and asks what happens if mergers are allowed or not. It considers price regulation policies for retail gasoline and whether these regulations lead to higher prices for consumers. It looks at how the US federal government runs auctions for timber and oil drilling leases. Hopefully, the book points you to new questions and new data to answer existing questions.

The book does not recommend policies. The government economist, Alice Rivlin, argued that it is extremely important to provide policy makers with objective analysis. In a memo to staff of the Congressional Budget Office (CBO), she said the following.[2]

> We are not to be advocates. As private citizens, we are entitled to our own views on the issues of the day, but as members of CBO, we are not to make recommendations, or characterize, even by implication, particular policy questions as good or bad, wise or unwise.

Economists in government, the private sector and the academy, work on important policy questions. Economists are most effective when they do not advocate for policy positions, but present objective analysis of the economics and the data. This book presents an objective analysis of interesting policy questions but doesn't state whether the policy positions are good or bad, wise or unwise.

[2] https://www.cbo.gov/sites/default/files/Public_Policy_Issues_Memo_Rivlin_1976.pdf

The Outline

The book is laid out the same way Robert Gibbons laid out his classic text, *Game Theory for Applied Economists*. There are four parts. **Static games of complete information**, dynamic games of complete information, static games of incomplete information, and dynamic games of incomplete information.

Static Games of Complete Information

This part presents the simplest version of a game.

Chapter 1 introduces the basic mathematical concepts of game theory. It analyzes the most famous game in game theory, the **prisoner's dilemma**. The game is used in a TV game show and the chapter uses data on actual behavior in the game where the players of the game make choices worth thousands of dollars. Do real people on a TV game show play the game as the mathematics predicts?

Chapter 2 introduces two important equilibrium concepts, **dominant strategy equilibrium** and **Nash equilibrium**. The chapter uses Nash equilibrium to understand how the number of tire retailers varies from city to city.

Chapter 3 studies **oligopoly**, markets with a small number of competitors, and three models of how these markets work, **Cournot**, **Bertrand**, and **Hotelling**. The most general model allows competing firms to be similar but not the same. This model is used to understand pricing competition between McDonald's outlets in late 1990s Santa Clara county.

Chapter 4 considers the implications of multiple Nash equilibria in a game where two firms are choosing whether or not to enter the same market. This game is used to analyze the entry decisions by the mega bookstores in the 1990s, Borders and Barnes & Noble. The chapter analyzes problems where the game does not always make a single prediction.

Chapter 5 analyzes **mixed strategies** in the context of both **coordination games** and **zero-sum games**. It uses mixed strategies to model entry by the mega bookstores, Borders, and Barnes & Noble. Zero-sum games were the first types of games analyzed using game theory. Many parlor games like chess, droughts, and poker are zero-sum games. The chapter uses mixed strategies and zero-sum games to understand the choices made by soccer players when kicking and defending penalty kicks in the English Premier League.

Dynamic Games of Complete Information

These games are substantially more complicated than the games presented in the first part of the book. In response, we need to make a number of simplifying assumptions that allow us to use the richness of the dynamics without being overwhelmed by the complexities.

Chapter 6 introduces **subgame perfection** and uses the concept to analyze the entry dynamics of the mega bookstores, Borders and Barnes & Noble.

Chapter 7 presents three different models of **bargaining**. It asks whether the simplest, the **ultimatum game**, makes predictions that are consistent with actual behavior of actual people when making decisions involving large sums of money. The answer is no, not really. A more complicated game makes more reasonable predictions. Luckily the **Rubenstein alternating offers game** makes predictions similar to a much simpler analytical tool, the **Nash bargaining solution**. This tool is used to analyze mergers between hospitals in Palm Beach County, the home county for the publisher of this book.

Chapter 8 returns to **oligopoly markets** but allows more complicated interactions between the firms. The chapter analyzes the pricing behavior of gasoline retailers in Perth Australia. The chapter presents a model to explain the weird saw tooth pattern in retail gas prices. It considers the extent of the prediction error when a merger model assumes firms choose prices more independently than they actually do.

Static Games of Incomplete Information

In the first two parts of the book, the players of the game are assumed to know everything. In the second two parts of the book, that assumption is relaxed.

Chapter 9 revisits analysis of entry by mega bookstores, Borders and Barnes & Noble. The difference is that each firm observes its own entry costs but not the entry costs of the other firm. In this model equilibrium, the firms do not know their competitors unobserved costs of entry. The firms know their own costs of entering a new market, but not their competitors costs of entering that same market. While the game is more complicated than the game presented in Chapter 4, it is sometimes simpler to use in data analysis.

Chapter 10 and Chapter 11 analyze auctions. Chapter 10 uses two of the main auction types, **sealed bid auctions** and **English auctions** to analyze bidding behavior and collusion in US Forestry auctions conducted in the 1970s.

Chapter 11 considers auctions where the bidders don't know exactly how much to value the item they are bidding on. The classic example is oil drilling auctions on the US Outer Continental Shelf (OCS). The chapter asks whether the government should allow firms to collude in those auctions.

Dynamic Games of Incomplete Information

The fourth part of the book analyzes the problems of moral hazard and adverse selection.

Chapter 12 considers the **principal-agent problem** and uses it to analyze the use of corporate financing in 19th-Century whaling in New England. What sort of contracts were used by firms and families to finance whaling expeditions, where the whaling ships were literally sailing around the world?

Chapter 13 brings game theory to health insurance markets. The chapter presents a model of the used car market, which suggests that information problems will cause the market to fail. Similarly, information problems in health insurance markets suggest that government interventions such as large tax subsidies for people insured through their work are necessary in order to have people insured.

Notation

As you have seen above, the book uses particular fonts and symbols for various important things. It uses the symbol **R** to refer to the scripting language. It uses `typewriter font` to represent code in **R**. Initial mentions of an important term are in **bold face font**.

When discussing actual data, it uses x_i to refer to the observed characteristic for some individual i. It uses $\mathbf{x}$ to denote a vector of the x_i's. For matrices, it uses $\mathbf{X}$ for a matrix and $\mathbf{X}'$ for the matrix transpose. A row of that matrix is $\mathbf{X}_i$ or $\mathbf{X}_i'$ to highlight that it is a row vector. Lastly, for parameters of interest it uses Greek letters. For example, β generally refers to a vector of parameters, although in some cases it is a single parameter of interest, while $\hat{\beta}$ refers to the estimate of the parameter.

Hello R World

To use this book you need to download **R** and `RStudio` on your computer. Both are free.

Download R and RStudio

First, download the appropriate version of `RStudio` here: `https://www.rstudio.com/products/rstudio/download/#download`. Then you can download the appropriate version of **R** here: `https://cran.rstudio.com/`.

Once you have the two programs downloaded and installed, open up `RStudio`. To open up a *script* go to "File > New File > R Script." You should have 4 windows, a script window, a console window, a global environment window, and a window with help, plots, and other things.

Using the Console

Go to the console window and click on the `>`. Then type `print("Hello R World")` and hit enter. Remember to use the quotes. In general, **R** functions have the same basic syntax, `functionname` with parentheses, and some input inside the parentheses. Inputs in quotes are treated as text while inputs without quotes are treated as variables.

```
> print("Hello R World")
  [1] "Hello R World"
```

Try something a little more complicated.

```
> a = "Chris"  # or write your own name
> print(paste("Welcome",a,"to R World"))
  [1] "Welcome Chris to R World"
```

Here we are creating a variable called `a`. The `#` is used in **R** to "comment out" lines in codes. **R** does not read the line following the hash.

In **R** we can place one function inside another function. The function `paste` is used to join text and variables together. The function `paste()` defaults to placing a space between the inputs. When placing one function inside another make sure to keep track of all of the parentheses. A common error is to have more or less closing parentheses than opening parentheses.

```
> paste(
  "Welcome",
   a,
   "to R World"
) |>
  print(
  )
  [1] "Welcome Chris to R World"
```

R can also accept code that looks like above. Using spaces and new lines helps a human reader understand what the code is saying. The symbol `|>` says take the result from above and use it in the following function. For some reason, there are two different symbols used `%>%` or `|>`. You can select which type under global options. Here `|>` is used.

Thanks

I'm grateful to my wife, Deena Ackerman, for allowing me time to work on this project. Thank you to numerous friends and colleagues for providing feedback and suggestions on the book. I am particularly grateful to Devesh Raval and Emek Basker who gave extensive feedback on early drafts. Thanks also to a number of researchers who have been willing to provide me with the data used in the book including Raph Thomadsen, Eric Hilt and David Byrne. All errors are my own.

Discussion and Further Reading

The book is laid out the same way as Gibbons (1992). If you are looking for a more detailed or technical description of the various types of games then see Fudenberg and Tirole (1991). A lot of the applications presented in this book are from the sub-field of economics called industrial organization. The classic theory text for that field is Tirole (1988). Recently more empirical-oriented industrial organization books have come out including Aguirregabiria (2021) and Hortaçsu and Joo (2023). Paarsch and Hong (2006) have a similar orientation with a focus on auctions. If you are interesting in some of the economic experiments presented in the book, then Camerer (2003) is a good but somewhat dated overview of experimental game theory.

The book uses the coding language **R** to illustrate models and empirical problems. It generally uses the `tidyverse` and `data.table` flavors. The best introduction to the language is the books by Hadley Wickham and coauthors, in particular *R for Data Science* which is here https://r4ds.had.co.nz/.

Part I

Static Games of Complete Information

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

1

Games

1.1 Introduction

Game theory provides answers to life's many questions. Which side of the street should you walk on when there is no sidewalk? Will we ever talk to aliens? When should you turn across traffic at an intersection? Which mergers should be allowed? Should government intervene to stop bank runs? Who rides shotgun? During the COVID-19 emergency how many potential vaccines should the US government promise to purchase? Should NFL teams punt on 4th down more than they do?[1]

When there is no side walk, you should walk on side facing the oncoming traffic. The reason is that you can see the car coming toward you, but more importantly, you can see that the car coming toward you, can see you. You know that they know you are there. And, importantly, they know that you know, that you are there. Of course, you know that they know that you know that you are there. And they even know that. In game theory, we call this infinite circle of knowledge, **common knowledge**, and it is a fundamental assumption of most games. Players in the games analyzed in this book know everything, they even know that.

The chapter provides an introduction to games and game theory. It informally introduces a couple of strategic situations. It asks why there is no hijackings of airplanes any more. It presents data showing a significant drop off in hijackings early in the 2000s. The chapter presents the most famous game in game theory, the **prisoner's dilemma**. The chapter ends with an empirical analysis of how actual people play the prisoner's dilemma game. It uses data from the TV game show, *Friend or Foe*.

1.2 Some Strategic Situations

The section introduces two strategic situations: communicating with aliens and turning into traffic.

[1]The answers are: The opposite side of which cars drive. No. It depends where you live. How close competitors are the firms? Yes. Winner of Rock-Paper-Scissors. Many more than they actually did. No.

DOI: 10.1201/9781003351603-1

1.2.1 Aliens Attack!

Many years ago you could sign up to have your computer used for something called the Search for Extraterrestrial Intelligence (SETI). The idea is that the program would use all the spare computing power around the world to work through a huge data set of radio signals from the world's radio telescopes. It was like Bitcoin mining, but for aliens, this search would look for evidence that the signal was generated by an intelligent life form from another planet.

To determine the reasonableness of this exercise, consider a game. This game has two players, an Earthling and an Alien. Each player has two choices. They can listen for a signal from the other player or they could send a signal to the other player. Because of physics we will assume that for both players listening for the signal is substantially cheaper than sending the signal. Consider the Earthling's choices. If they listen for a signal, it is pretty cheap and if they hear a signal then jackpot. We learn that we are no longer alone in the Universe! If the Earthling sends a signal, then it is expensive and there is a possibility that the Alien may hear the signal and they will learn that we exist. For the Alien, the outcomes are similar for their two choices.

What is an Earthling to do? If they listen, it is cheap and there is some possibility of finding out something amazing. If they send a signal, it is expensive, and while the Alien may learn something, the Earthling may not. The Earthling is probably better off listening for the signal. The Alien is also better off just listing for the signal. Given that both the Earthling and the Alien are listening, they are never going to hear anything and the search for extraterrestrial intelligence may not in fact be very intelligent.

1.2.2 Can AI Cope with Traffic?

It is not clear that we can or should have humans and AI-driven cars interact on the road. Consider the problem of turning across traffic at a traffic light. Most Americans, when confronted with the question of when to turn across traffic, will generally wait for the green arrow. While the through traffic often speeds up on the yellow. In Australia in the 1980s, there were few lights with green arrows and so drivers were pretty aggressive at turning across traffic. Australians tend to slow and stop at a yellow light, lest you T-bone the person aggressively turning across traffic. In Vermont, you may be surprised to learn that it is customary to allow drivers to turn across traffic on the green light. It is a bit of a shock to hit the gas only to look up and see a car turning across you. As in Australia, Vermont drivers learn to go easy when the light goes green in order to allow the turning traffic through. There are three different equilibrium outcomes in three different places. Would an AI-driven car know these rules and know how they change from place to place?

1.3 End of Skyjacking

Airplane hijacking used to be a thing. The high point was the late 1960s and early 1970s with over 50 a year, but even as late as 2000, there were over 20 hijackings a year. By the Twenty teens, there were some years with zero. To see what happened? See Figure 1.1

1.3.1 More Security?

Was there an increase in security? Yes. Security increased in the 1970s and 1980s. Obviously, there was a huge increase in security in the 2000s after the terrorist attacks on September 11, 2001. But that is not the reason, at least not the reason for the difference in the number hijackings before and after September 11, 2001. What changed was the game between passengers and hijackers.

When hijackers take over a plane, they are way outnumbered by the passengers. Hijacking relies on an implicit understanding between hijackers and passengers. Passengers agree to sit quietly in order to be safely released. Hijackers agree to release passengers so that they can get the plane or the ransom or whatever it is that hijackers want. You might think that hijackers have weapons while passengers don't, but neither may be true. The September 11 hijackers used box cutters, not exactly an AR-15.

1.3.2 Hijacking Data using R

We can see what happened by importing and plotting data on airline hijackings.[2] Here we use `read.csv()` which creates a `data.frame` object called `df`. This object is basically the stock standard data set we use in Stata or any other statistical programming language. If you go over to the Environment window and click on `df`, you will see an excel-like sheet with variable names as the column names and years on the rows. To call out variables in the data set `df`, we use the symbol `$` and then the name of the variable, for example `df$Year`.

```
> # dir = "ENTER LOCATION OF DATA FILES"
> file = paste0(dir, "Airline_Hijackings.csv")
> df = read.csv(file)
```

```
> plot(df$Year, df$Hijackings, type = "l")
> abline(v = 2001)
```

[2] Data from the Aviation Safety Network, https://aviation-safety.net/statistics/period/stats.php?cat=A1.

FIGURE 1.1
Annual count of airline hijackings around the world from 1940 to 2021. The plot shows the increase up to the 1970s and then a slow decline into the 90s, then a dramatic dropoff after 2001.

Figure 1.1 presents the annual count of airline hijackings around the world. Hijacking really spiked in the late 60s early 70s. It then steadily declined to 2001. The figure shows the dramatic change in the number of hijackings before and after 2001.

Figure 1.1 is made with the base **R** plotting system. The function `plot()` creates a scatter plot with years on the x-axis and the number of hijacking in the y-axis. In order to get the line, use the option `type = "l"`. In addition, a vertical line is added using the function `abline()` setting `v = 2001`.

1.3.3 United Arlines Flight 93

Figure 1.1 suggests that something changed in the early 2000s. We actually know exactly when and where the game between hijackers and passengers changed. It changed on United Airlines Flight 93 over central Pennsylvania on the morning of September 11, 2001. There had been three flights hijacked earlier in the morning, the hijackers took over the plane and the passengers sat quietly. The passengers and crew assumed that they would be safe as long as they obeyed the instructions. The hijackers were playing a different game. They didn't want to take the plane to somewhere or get the ransom money.

They wanted to fly the plane into New York's World Trade Center or the Pentagon in Arlington, VA. Passengers and crew on Flight 93 learned that the game had changed through text messages and phone calls. They knew that sitting quietly was not going to get them home to their families safely. Passengers and crew overwhelmed the hijackers and brought the plane down in a field killing all on board.

The game had changed. Hijackers can no longer rely on the compliance of crew and passengers. We see this even with people being rowdy on planes. The passengers and crew overwhelm them and the rowdy passenger ends up tied down to their seat with duct tape.

1.4 A Game

A game is a formal mathematical object. This formalism is super important. New students to game theory will often substitute their own intuition for the formal mathematics. This is a big error. Keep your intuition in check. Intuition is useful for understanding the result you get or suggesting that you may have made an error in your calculation. To find the solution to a game, use math.

The section presents the formal mathematical objects used in game theory.

1.4.1 Definitions

A basic game has three formal parts.

- A set of **players**
- A set of **strategies**
- A set of **payoffs**

Definition 1. *A* player *is a strategic actor in the game. They have choices and make their choices based on what other players in the game do.*

What makes game theory different from other mathematical models of economic activity is the assumption that players in the game are strategic actors. They account for the reaction of other players in the game when making their choice. In Econ 101, we assumed that the players of the game, the consumers or businesses, are not strategic. These player may optimize utility or profits, but they take the actions of the other players as fixed. Players in a game account for the strategic behavior of other players in the game.

Definition 2. *A* strategy *is a complete plan of actions for every possible circumstance faced by the player in the game.*

When we start in this book, the strategies will be very simple. They will just be actions. Each player will have a limited set of choices and their strategies correspond to those choices. In the second part of the book things get a whole lot more complicated. A **strategy** is a complete plan. Think about how complicated this object can be. Think about the game chess. There are millions and millions of different configurations. It is estimated that there are 10^{44} possible legal positions in a game of chess.[3] That is 1 followed by 44 zeros. A strategy for chess must state what the player will do in each possible case!

Definition 3. Payoffs *are the outcomes of interest to the players of the game, these may be winning or losing, monetary values or utilities.*

The third object of a game is the payoffs. The **payoffs** to players depend directly upon the combination of actions taken by players in the game. It is this dependence that makes it a strategic situation.

Often in games the absolute magnitude of the payoffs is less important than the relative magnitude. Students new to game theory may be put off by the weird precision that seems to exist in the games. In reality, the numbers are just stand ins, what really matters is the relative ordering of the outcomes.

1.4.2 Examples

To get a bit more of a feel for what a game is, consider the examples mentioned above.

First up is the SETI game. There are three formal parts of the game, players, strategies and payoffs. Assume that when the player chooses to signal they also listen for a signal.

- Players: Earthling, Alien
- Strategies:
  - Earthling: Listen, Signal
  - Alien: Listen, Signal
- Payoffs:
  - Earthling: Listen, Alien: Listen
    - Earthling: Nada
    - Alien: Nada
  - Earthling: Listen, Alien: Signal
    - Earthling: Earthling learns there are Aliens and it changes the course of human history!
    - Alien: Nada plus a large cost for sending the signal.

[3] https://github.com/tromp/ChessPositionRanking accessed 4/14/23.

  - Earthling: Signal, Alien: Listen
    * Earthling: Earthling pays a large cost for nothing
    * Alien: Aliens learn there are Earthlings changing the whole course of Alien history!
  - Earthling: Signal, Alien: Signal
    * Earthling: Earthling pays a large cost but learn their are Aliens!
    * Alien: Aliens pay a large cost but learn there are Earthlings!

    Earthlings and Aliens pay a large cost but both learn each other exists.

We can do the same thing and formally present the hijacking game.
Hijacking game:

- Players: Hijacker, Passenger
- Strategies:
  - Hijacker: Kill passengers, Don't kill passengers
  - Passenger: Fight hijacker, Don't Fight hijacker
- Payoffs:
  - Hijacker: Kill, Passenger: Fight
    * Hijacker: Dead
    * Passenger: Dead
  - Hijacker: Kill, Passenger: Don't Fight
    * Hijacker: Dead or arrested
    * Passenger: Dead
  - Hijacker: Don't Kill, Passenger: Fight
    * Hijacker: Dead or injured
    * Passenger: Dead or injured
  - Hijacker: Don't Kill, Passenger: Don't Fight
    * Hijacker: Get money or where they are going
    * Passenger: Go home to their families

Prior to Flight 93, everyone thought that game was such that the Don't Kill, Don't Fight outcome was the most likely to occur. In the first three flights of September 11, 2001, the hijackers took advantage of the fact that crew and passengers thought that by choosing Don't Fight, they would have a reasonable possibility of going home to their families. During Flight 93, passengers and crew realized that hijackers were playing the Kill strategy and adjusted their behavior accordingly. Today passengers and crew will almost certainly play Fight. Given that strategy, Hijackers are better off not hijacking in the first place.

1.5 Prisoner's Dilemma

It is probably the most famous game in game theory. If anyone has heard all about game theory, then they have heard of the **prisoner's dilemma** game. The classic version of the tale is that there are two suspects to a crime. The suspects are taken into two different cells. The detectives visit each suspect in turn. They offer the same deal to both. The deal is that if the suspect rolls over and cops to the crime implicating the other suspect then they get to walk. While if they stay quiet and their mate rolls over they go away for a long time. Each suspect is presented with this deal and made aware that their mate also got the deal.

The suspects also know that if they stay mum, they will probably get charged with a less serious crime. The police don't have anything to tie them to the more serious crime except each other. Finally, if they both squeal, then they both get an intermediate sentence.

The section presents a formal analysis of the prisoner's dilemma game.

1.5.1 The Game

Table 1.1 presents the **normal form** representation of the prisoner's dilemma game. Each player has two choices, stay mum or squeal, with Suspect 1's choice on the rows. The payoffs for each outcome are in the cells, with Suspect 1 listed first. The payoffs are in number of years of prison, with bigger numbers being worse, and thus negative.

TABLE 1.1
Normal form representation prisoner's dilemma game, with two players S_1 and S_2. For S_1 their choices are the rows and their payoffs are listed first in each cell.

S_1, S_2	Mum	Squeal
Mum	$-1, -1$	$-5, 0$
Squeal	$0, -5$	$-3, -3$

We can also write the game out like we did above.

- Players: Suspect 1, Suspect 2
- Strategies:
  - Suspect 1: Mum or Squeal
  - Suspect 2: Mum or Squeal

- Payoffs:
  - {Mum, Mum}: {−1, −1}
  - {Mum, Squeal}: {−5, 0}
  - {Squeal, Mum}: {0, −5}
  - {Squeal, Squeal}: {−3, −3}

There are four outcomes, both choose Mum, both choose Squeal, Suspect 1 chooses Mum and Suspect 2 chooses Squeal, and Suspect 1 chooses Squeal and Suspect 2 chooses Mum.

1.5.2 What is the Best Outcome?

The outcome with the highest payoff for both players is {Mum, Mum}. NOTE, the outcome is not payoffs, but strategies.

This outcome gives a total of −2 points, while other outcomes give both players less in total (−5 and −6). This outcome is **Pareto optimal**. The only way to make one player better off is by making the other player worse off. The only outcome that is not Pareto optimal is {Squeal, Squeal}. Moving from {Squeal, Squeal} to {Mum, Mum} makes both players better off.

1.5.3 What should You Do?

Assume you are Suspect 1. What should you do?

- Assume Suspect 2 chooses Mum. The payoffs from your choices are as follows:
  - Mum: −1
  - Squeal: 0
- Assume Suspect 2 chooses Squeal. The payoffs from your choices are as follows:
  - Mum: −5
  - Squeal: −3

If your mate is going to stay mum, then you get a big payoff from squealing. You get to walk out a free person. So in that case you are better off squealing. What if your mate squeals? That rat bastard! Well, if you stay mum then you go down for a long time. If you also squeal you at least reduce your sentence. So in that case you should squeal. In fact, in both cases you should squeal.

We say that squealing is a **dominant strategy** because for every possible strategy of the other player, you are always better off squealing.

1.5.4 Does This Make Sense?

The reason that the **prisoner's dilemma** game is so famous is that it suggests an outcome that some find unintuitive. It suggests that both players would squeal when both players would be better off if they both agreed to stay mum. By squealing they both end up spending three years in prison, while if they had stayed mum they would have both only spent one year in prison. Why would they end up at the outcome that makes them worse off?

1.6 Empirical Analysis: *Friend or Foe* using R

What happens when real people play games like the prisoner's dilemma presented above? One way economists test this is by running experiments. These are often done on college students for small payoffs, say money or a Starbucks gift card. What about the real world?

In the early 2000s, a game show aired on the Game Show Network called *Friend or Foe*. The show had players pair up and compete against other teams to answer trivia questions. The teams earned cash by answering the questions correctly. However, before taking their winnings home, the two players in the winning team had to play a game. Not a game show game, well sure a game show game, but a game theory game as well. A prisoner's dilemma game to be precise, the game was called *Trust Box*.

The cool thing is that someone watched these games and created a data set with the payoffs, the outcomes, the strategies chosen and some demographic information on the players. Note also that amounts here are in thousands of dollars. So real people playing for real money in a completely unreal situation.

The section presents the game and uses **R** to analyze data collected on how actual contestants played the game.

1.6.1 Trust Box

TABLE 1.2
Normal form representation *Trust Box* game, with two players P_1 and P_2. For P_1 their choices are the rows and their payoffs are listed first in each cell. The team's winnings is x. It is amount which will vary based on how well the team did in the trivia part of the show.

P_1, P_2	Friend	Foe
Friend	$0.5x, 0.5x$	$0, x$
Foe	$x, 0$	$0, 0$

Table 1.2 presents the normal form representation of the *Trust Box* game.

Each player is in a box with a button to press, *Friend or Foe*. If both players press Friend then they share the proceeds evenly, $0.5x$ each, where x is the team's winnings from the trivia part of the game. If the first player chooses Foe and the second player chooses Friend, then the first takes home all the winnings and the second gets nothing, x and 0. If both players choose Foe, then they take home nothing, 0.

While the strategies have different names, this game has the same ranking of payoffs as the prisoner's dilemma game presented in the previous section. Remember it is the ranking of the payoffs that matters.

The choice Foe is a **weakly dominant strategy** for Player 1. Let Player 2 play Friend, for Player 1 Foe is better as $x > 0.5x$. Let Player 2 play Foe, for Player 1 Foe and Friend give the same payoff (0). If a strategy gives the highest payoff against some strategies and the equal highest against all other strategies, we say it is **weakly dominant**. Player 1 is better off choosing Foe. Player 2 is also better off choosing Foe.

1.6.2 Data Analysis in R

We can bring in the data from the game show that is available from the `Ecdat` package.[4] You can see that using the function `data()` you create an object called `FriendFoe` which is the data set. Each row is a different game and the two players are differentiated by whether or not there is a `1` after the variable name. These are the data described in Kalist (2004).

We can ask how actual people play the prisoner's dilemma game and look at which cells of the game described in Table 1.2 the players end up in. The syntax uses the **pipe operator** (`|>`). The result of the code on one line is inserted into the function on the next line. The syntax creates an object called `game_mat` which is eventually going to end up as a 2×2 matrix. The columns and rows of `game_mat` are named with the names of the action in the game. The table is printed using the `xtable` package.

```
> library(Ecdat)
> library(tidyverse)
> data("FriendFoe")
> game_mat = FriendFoe |>
+   summarize(
+     mean(play == "friend" & play1 == "friend"),
+ # this creates a vector of 0 and 1s depending whether
+ # the statement inside the parenthesis is true
+ # note the use of "==" to ask whether something is true
+     mean(play == "foe" & play1 == "friend"),
+     mean(play == "friend" & play1 == "foe"),
+     mean(play == "foe" & play1 == "foe")
+     ) |>
```

[4]To use this package you need to install it using the menu or `install.packages()`.

```
+       matrix(, nrow = 2)
> colnames(game_mat) = rownames(game_mat) = c("Friend","Foe")
```

```
> library(xtable)
> print(xtable(game_mat, digits = 2), floating = FALSE)
```

TABLE 1.3
Fraction of games that finish in each cell. Almost a third of the time the players end up at {Foe, Foe} and take home nothing.

	Friend	Foe
Friend	0.23	0.26
Foe	0.19	0.32

Table 1.3 presents the results of the game. It shows which cells everyone ends up in. Was this what you were expecting? The largest group of people end up playing Foe and Foe and getting nothing. Why is this? Do you think the game show designed it this way? Was the game show hoping that even more people would choose this option? What change to the game could you make to get more people to choose Foe?

1.6.3 Analysis in R

Table 1.3 raises lots of questions. Does the outcome of the game depend on the stakes of the game, that is, how much prize money the two players have won? Does it depend on the gender of the players or how similar the two players are or how much experience they had together.

The next bit of code creates a similarity index based on whether the two contestants are the same gender, the same race and similar in age. This uses the function `mutate()`.

```
> FriendFoe = FriendFoe |>
+   mutate(
+     similar = (sex == sex1) +
+       (max(abs(age - age1)) - abs(age - age1))/10 +
+          (white == white1)
+   )
```

For some of the analysis, it is helpful to reshape the data. The following code takes the columns associated with the first player and stacks them on top of the columns associated with the second player. The index `c(1:4, 12)` keeps the variables of interest for first player listed, while `c(8:11,13)` does the equivalent for the second player listed. The first data set is called `df1` and the second is `df2`. The code then labels the columns of `df2` using the column names of `df1`. It then joins the two data set together using the function `rbind()`, where the `r` is for row and calls the new data set `df`.

```
> df1 = FriendFoe[ , c(1:4, 12)]
> df2 = FriendFoe[ , c(8:11, 13)]
> colnames(df2) = colnames(df1)
> df = rbind(df1,
+             df2)
```

The data then create some new variables or recreate some previous variables lost from the original data. The function `rep()` repeats the vector or value in the first space by the number in the second. The variable `game` creates an id for each game in the data.

```
> df$game = rep(1:227, 2)
> df$round = rep(FriendFoe$round, 2)
> df$season = rep(FriendFoe$season, 2)
> df$cash = rep(FriendFoe$cash, 2)
> df$similar = rep(FriendFoe$similar, 2)
```

Now we can do some analysis on this data. What is the relationship between how much a player earns during the trivia part of the game, their demographics and how far they get in the game? Also, does their strategy choice in the prisoner's dilemma depend on their demographics, or the amount of cash on the table? For the first part we estimate the relationship with standard **linear regression** using the `lm()` function.`lm1` is an object created in **R** that can be used by other code to create regression result tables. For the second part, we use a **logit model** and a **probit model**. These are standard methods for analyzing problems where the outcome has two possibilities. Here `play` could be either `"friend"` or `"foe"`. In **R** we use the `glm()` function for generalized linear model. To specify a logit model, you need to state `family = binomial(link = "logit")`.

```
> lm1 = lm(cash ~ sex + white + age + round + season,
+           data = df)
> glm1 = glm(play == "friend" ~ sex + white + age +
+               similar + cash,
+             family = binomial(link = "logit"),
+             data = df)
> glm2 = glm(play == "friend" ~ sex + white + age +
+               similar + cash,
+             family = binomial(link = "probit"),
+             data = df)
```

We can display the results using the `stargazer` package.[5]

[5] Here the package is transforming the **R** output into LateX code.

TABLE 1.4
Linear regression estimates for the relationship between the amount earned in the trivia part of the game as well as logit and probit estimates of the probability of playing friend in the *Trust Box* part of the game.

```
> require(stargazer)
> stargazer(list(lm1, glm1, glm2),
+           keep.stat = c("n", "rsq"),
+           float = FALSE)
```

	Dependent variable:		
	cash	play == "friend"	
	OLS	*logistic*	*probit*
	(1)	(2)	(3)
sexmale	0.102	−0.105	−0.067
	(0.175)	(0.193)	(0.120)
whiteyes	0.309	0.441*	0.275*
	(0.234)	(0.265)	(0.163)
age	−0.005	0.041***	0.025***
	(0.011)	(0.013)	(0.008)
round2	2.242***		
	(0.215)		
round3	5.090***		
	(0.206)		
season2	−0.553***		
	(0.183)		
similar		0.077	0.049
		(0.121)	(0.075)
cash		0.007	0.005
		(0.034)	(0.021)
Constant	1.246***	−2.028***	−1.270***
	(0.449)	(0.770)	(0.476)
Observations	454	454	454
R^2	0.591		
Note:	*p<0.1; **p<0.05; ***p<0.01		

Table 1.4 presents the regression results. The first column of results is for the **linear regression**. It shows the relationship between the size of the prize, demographic characteristics, and characteristics of the team's performance. The last two columns present the logit and probit regressions respectively. These show the relationship between a player's demographic characteristics and playing the strategy friend in the *Trust Box* game.

Table 1.4 shows that the amount of cash earned increases significantly in the latter rounds. It also seems that the budget decreased in season 2. In addition, the table shows that the strategies of the players do seem to change with the player demographics. Older players seem to be much more willing to cooperate. The stakes, although in thousands of dollars, seem less important. Similarity of demographics does not seem important either. That said, according to List (2006), there is a selection round where players are matched into teams and that selection may account for the results.

1.7 Discussion and Further Reading

Games are all around us. Strategic behavior is very common. This book aims to give you the tools to understand how the real world works. It aims to make you comfortable understanding the strategic behavior generating the data you are analyzing.

The *Trust Box* analysis is based on List (2006) and Kalist (2004). The American economist, John List, has made his career pioneering the running of economic experiments in the real world, **field experiments**. One objective of these experiments is to test the predictions of the game theory models (Camerer, 2003). The book uses data from another field experiment in Chapter 8.

2

Nash Equilibrium

2.1 Introduction

The modern version of game theory began with the PhD dissertation of the American mathematician, John Nash. You may recognize the name from the hit movie, *A Beautiful Mind* with Australian Russell Crowe playing West Virginian Nash. Nash showed that for a large set of games, there is at least one outcome that is stable. Nash was interested in which outcome of the game is likely to occur. Stability seems like a prerequisite for any outcome that we are likely to observe. Stability says that if an outcome happens to occur then the outcome of the game is unlikely to change. In contrast, instability says that if an outcome happens to occur then the outcome of the game is likely to change. We are unlikely to observe an unstable outcome. The fact that Nash's solution concept is both a prerequisite for an outcome that is likely to occur and exists for a large set of games, makes it a super valuable idea. Today we call his proposed outcome a **Nash equilibrium**

Nash worked in the sub field of game theory called non-cooperative game theory. As he himself points out in his Nobel essay, this style of game theory was not in style.[1] It differed substantially from the ideas of Johnny von Neumann who was a senior mathematician at Princeton where Nash did his dissertation. Hungarian-American von Neumann helped develop game theory and was one of the first to suggest its use in economics. His collaboration with German-American economist, Oskar Morgenstern, produced the first game theory text, *Theory of Games and Economic Behavior* (1944). Luckily for us, Nash was stubborn and his efforts helped turn non-cooperative game theory into what today we just call game theory.

In this chapter, we will introduce this workhorse equilibrium concept as well as the concept of a **dominant strategy equilibrium**. The chapter illustrates these concepts by looking at the market structure in retail tire stores. How many stores are there going to be in the market? How does the number of stores relate to the price of tires in the market?

[1]https://www.nobelprize.org/prizes/economic-sciences/1994/ceremony-speech/. Accessed on 2/23/23.

DOI: 10.1201/9781003351603-2

2.2 What is a Nash Equilibrium?

What is an **equilibrium**? We are looking for an outcome that is a reasonable prediction of the game. We are not looking for the best outcome or the worst outcome, but the outcome with the most empirical content. The one that we are likely to observe in the wild. This outcome is not chosen in some ad hoc manner but based on a set of assumptions. A set of rules.

Definition 4. *An* equilibrium concept *is a set of rules for determining which outcome(s) of the game will occur.*

There are various assumptions we may want to hold in order to say that a particular outcome is a reasonable prediction of the game. The section presents the idea of **dominance** and how we may come up with a prediction of the game based only on assuming that the players of the game make choices that are rational.

2.2.1 Dominance

Before we get to **Nash equilibrium**, consider an alternative equilibrium concept, called **dominant strategy equilibrium**. Chapter 1 introduced the concept of a **dominant strategy** and **weakly dominant strategy**.

Definition 5. *A strategy is* dominant *if it gives the player the highest payoff irrespective of the strategies of the other players. A strategy is* weakly dominant *if for some strategies of the other players, the payoff is equal highest.*

A dominant strategy is always the best choice. It does not matter what any other player of the game does. Given that it is always the best choice, then presumably it is the one that a rational player is most likely to choose.

Definition 6. *A (weakly)* dominant strategy equilibrium *is an outcome of the game where all players play strategies that are (weakly) dominant strategies.*

If we assume that players are rational, and there exists an outcome which is a **dominant strategy equilibrium**, then we would expect that outcome to be the one we observe. The issue is that such outcomes may not exist in the game we are studying.

In the discussion above there is a subtle but important assumption. If a player has two choices, A and B, where A is part of a proposed equilibrium but B gives the player the exact same payoff, then A is still part of the equilibrium. For A not to be part of an equilibrium, there must be a B that makes the player strictly better off.

2.2.2 Prisoner's Dilemma

Consider a version of the prisoner's dilemma game presented in Chapter 1. In this version, the actions have generic names, but the payoffs have the same ordering as they did in the version presented in the previous chapter. Remember it is the ordering that matters, just like in Formula 1.

TABLE 2.1
Normal form representation of a prisoner's dilemma game with players P_1 and P_2 and actions BLACK and RED. P_1's actions are on the rows. The payoffs are in the cells with P_1's in the first position.

P_1, P_2	BLACK	RED
BLACK	3, 3	0, 5
RED	5, 0	2, 2

Table 2.1 presents the generic prisoner's dilemma game with two players P_1 and P_2 and two actions BLACK and RED. The payoffs are written out in **normal form**, which is a matrix-like representation of the game.

The outcome that has both players playing **dominant strategies** is {RED, RED}.

If you are player 1, you see that you are better off choosing RED irrespective of what player 2 does. In the table, player 1's payoff is the first element in the cell. Player 1's strategies are the rows. If player 2 chooses BLACK, we see that RED has a payoff of 5 which is greater than 3. If player 2 chooses RED, we that RED has a payoff of 2 which is greater than 0. RED is the **dominant strategy** for player 1. Similarly for player 2.

The reason the prisoner's dilemma is so famous is that it predicts an outcome which is bad for the players of the game. Even though two rational players will end up at the dominant strategy equilibrium, both players would be better off choosing {BLACK, BLACK} which has a payoff of {3, 3}. This outcome has a payoff that is strictly better for both players than the predicted outcome which has a payoff of {2, 2}. Our equilibrium concept predicts the outcome that we believe is most likely to occur not necessarily the outcome that is the best for the players.

The prediction of this game drives real world policy. The US Department of Justice (DOJ) has an explicit policy of giving leniency to firms that are the first to provide evidence of collusion in a market.[2] The policy explicitly aims to create a prisoner's dilemma among colluding firms.

Does dominant strategy equilibrium predict outcomes in real games? Chapter 1 analyzes data from a real TV program where players are playing a prisoner's dilemma game for real money. In *Friend or Foe*, if both players choose Foe then they each leave the show without any of their winnings. While

[2] https://www.justice.gov/atr/leniency-program accessed on November 12 2023.

if they both choose Friend then they would share their winnings. In the data, we see that {Foe, Foe} has the highest probability of occurring.[3]

2.2.3 Coordination Game

Now consider a slightly different game. We call the game presented in normal form in Table 2.2, a coordination game. We will soon see why.

TABLE 2.2
Normal form representation of a **coordination game** with players P_1 and P_2 and actions RED and BLACK. P_1's action choices are on the rows. Payoffs are in brackets with P_1 listed first.

P_1, P_2	BLACK	RED
BLACK	2, 5	0, 0
RED	0, 0	5, 2

Assume you are player 1. Can you work out your best strategy? We can go through the best choice for each choice made by player 2.

- Assume P_2 chooses BLACK. P_1's payoffs are:
  - BLACK: 2
  - RED: 0
- Assume P_2 chooses RED. P_1's payoffs are:
  - BLACK: 0
  - RED: 5

Is there a dominant strategy to this game? No. The best choice for you depends on the choice of player 2. If player 2 plays BLACK, then you should also play BLACK because $2 > 0$. If player 2 plays RED, then you should also play RED because $5 > 0$. See a **coordination game**! Both players should coordinate on the color.

What outcome do you predict will happen in this game? While both players want to coordinate they do not agree on which outcome to coordinate. Player 1 prefers that they coordinate on RED and Player 2 prefers that they coordinate on BLACK.

[3]The second part of the book considers changes to the prisoner's dilemma game that lead to predicted outcomes where the players have higher payoffs.

2.3 Nash Equilibrium

The section discusses the definition of a **Nash equilibrium** and illustrates how to find one or more Nash equilibria in games that have been presented earlier in the chapter, the prisoner's dilemma and the coordination game.

2.3.1 Definition

Definition 7. *A* Nash equilibrium *is a set of strategies such that each player's strategy has the highest payoff given the strategies of the other players.*

Like with the dominant strategy equilibrium, we assume that each player chooses the strategy that gives the highest payoff. Our base assumption is that the players of the game are rational. The difference here is that choice is dependent on the choice of the other players.

Stability is baked into Nash's concept. Each player is choosing the best option given the choices of all the other players. If all players are choosing the best option, then there is no need for any player to change. The outcome is stable.

2.3.2 Algorithm for Finding Nash Equilibrium

While Nash equilibrium has a number of nice properties, finding it is not one of them. We will use the following cumbersome algorithm. One issue new students have with game theory is that it seems like the teacher just picked some outcome at random and voila, it is an equilibrium! Of course, the teacher did not just pick at random. They already knew the answer. In reality, you do not know which one to pick and so unfortunately you just have to try them all. There are no short cuts.

- Step 1: Choose a candidate outcome (set of strategies).
- Step 2: Hold Player 1's strategy fixed. Is Player 2's strategy optimal?
  - Yes: Go to Step 3.
  - No: Not a Nash equilibrium.
- Step 3: Hold Player 2's strategy fixed. Is Player 1's strategy optimal?
  - Yes: Nash equilibrium.
  - No: Not a Nash equilibrium.

If the game has more than two players, then the algorithm can be expanded to go through each player in turn.

2.3.3 Prisoner's Dilemma Game

Let's test out our algorithm on a game we already know, the prisoner's dilemma. For this game, there exists a more efficient algorithm, but we are illustrating how this more general algorithm works. Remember the first step is just picking an outcome. No magic.

- Step 1: $\{BLACK, BLACK\}$
- Step 2: P_1 plays BLACK. Is BLACK optimal for P_2?
  - BLACK: 3, RED: 5
  - No: Not a Nash equilibrium.

Let's pick another one.

- Step 1: $\{RED, RED\}$
- Step 2: P_1 plays RED. Is RED P_2's optimal strategy?
  - BLACK: 0, RED: 2
  - Yes. Go to Step 3.
- Step 3: P_2 plays RED. Is RED P_1's optimal strategy?
  - BLACK: 0, RED: 2
  - Yes!
  - $\{RED, RED\}$ is a Nash equilibrium.

OK. There is some magic. You should try the algorithm on the other two outcomes.

2.3.4 Coordination Game

Let's try something a little more complicated, the coordination game presented in Table 2.2.

- Step 1: $\{BLACK, BLACK\}$
- Step 2: P_1 plays BLACK. Is BLACK optimal for P_2?
  - BLACK: 5, RED: 0
  - Yes. Go to Step 3.
- Step 3: P_2 plays BLACK. Is BLACK optimal for P_1?
  - BLACK: 2, RED: 0
  - Yes.
  - $\{BLACK, BLACK\}$ is a Nash equilibrium.

That was easy/lucky. Are there any other Nash equilibria of this game?

- Step 1: $\{RED, RED\}$
- Step 2: P_1 plays RED. Is RED P_2's optimal strategy?
  - BLACK: 0, RED: 2
  - Yes. Go to Step 3.
- Step 3: P_2 plays RED. Is RED P_1's optimal strategy?
  - BLACK: 0, RED: 5
  - Yes.
  - $\{RED, RED\}$ is a Nash equilibrium.

There are two Nash equilibria. Actually, there are even more, but we will come back to that in Chapter 5.

One of the valuable things about a Nash equilibrium is that it always exists.[4] Unfortunately, the cost is that there may be more than one Nash equilibrium in any game. Is one of the Nash equilibria of the coordination game more reasonable than the other? What is the most reasonable prediction of the game? This question of whether there exists more reasonable equilibria is called refinement. In this book, we will consider a number of refinements of Nash equilibrium. Chapter 4 analyzes what happens when the game we want to take to the data has multiple equilibria.

2.4 Entry Games

The section introduces empirical entry games. It presents the framework for how economists generally think about determinants of the number of firms in the market. You will often hear policy makers complain that prices are high because there is a lack of competition. There are too few firms in the market. Rarely, do policy makers step back and ask why there are too few firms in the market. The section presents a game in which a small number of firms choose to enter the market.

2.4.1 Bresnahan and Reiss

In a series of papers published in the early 1990s, the American economists, Tim Bresnahan and Peter Reiss, analyze the empirical implications of a simple entry game. Bresnahan and Reiss (1991b) analyze average prices for retail

[4] A Nash equilibrium exists for any finite game, that is any game where the number of players and the number of strategies is finite.

tire stores in various U.S. towns. The data shows that there is basically no relationship between observed average prices for tires and the number of tire retailers in the market.[5] Why do you think that is?

The problem is **endogeneity**.[6] The raw observations do not account for the fact that these tire retailers choose whether or not to be in the market based on various factors including the cost of selling tires. We are more likely to see more firms in larger markets, but some of these firms are going to have higher costs of selling tires and these higher costs will drive up prices. We have two countervailing forces, more competition driving down prices and less efficient firms driving up prices.

This is not just some academic issue. It was a central concern in the US Federal Trade Commission's (FTC) case against the merger of Staples and Office Depot (Ashenfelter et al., 2006). In the late 1990s, the two office supply super stores wanted to merge. As evidence against the merger, the FTC showed that prices where lower in cities with more stores. Is that a causal statement? Did the increase in the number of stores cause prices to fall? Moreover, would the opposite happen. If the merger took place and the number of independent firms in the market fell, would prices go up?

2.4.2 Two Firm Entry Game

Consider a relatively simple version of an entry game where there are just two firms. If only one firm enters then that firm earns monopoly profits and pays a fixed cost of entry, represented by the number 2. If that firm doesn't enter, nothing happens, which is represented by 0. If both firms enter there is competition but the firms also have to pay the fixed costs of entry. This payoff is represented by -1, which is bad. So a firm is willing to enter the market, but only if they are a monopolist.

More formally we have two players, the firm's strategies are entered or don't enter and the payoffs depend on whether the other firm enters as well.

- Players: Firm 1, Firm 2
- Strategies:
  - Firm 1: Enter, Don't Enter
  - Firm 2: Enter, Don't Enter
- Payoffs:
  - {Enter, Enter}: $\{-1, -1\}$
  - {Enter, Don't}: $\{2, 0\}$

[5] Surprisingly the authors state that their data shows that "entry lowers margins". It is unclear how they come to this conclusion.

[6] In econometrics we use the term **endogeneity** to mean that different cases observed in the data may not be determined at random. Importantly, case assignment may be directly related to the observed outcome.

- {Don't, Enter}: {0, 2}
- {Don't, Don't Enter}: {0, 0}

We can also represent this game in a normal form payoff matrix.

TABLE 2.3
A normal form representation of a two firm entry game. Firm 1's strategies are the rows and the Firm 2's strategies are the columns. The payoffs are in the cells, with Firm 1's payoff first.

Firm 1, Firm 2	Enter	Don't
Enter	−1, −1	2, 0
Don't	0, 2	0, 0

Table 2.3 presents the normal representation of the game. What is the Nash equilibrium of the game? It seems unlikely that both firms will enter and we see quickly that is not an equilibrium. If both firms enter, then Firm 1 would have been better off not entering as 0 is greater than −1. See the first column and the payoffs of the first element which are for Firm 1.

How about the case where both firms don't enter? Again we see that Firm 1 is better off entering. Look at the second column and see that 2 is greater than 0.

Let's check the other outcomes more systematically using our algorithm.

- Step 1: {Enter, Don't}
- Step 2: Firm 1 plays Enter. Is Don't optimal for Firm 2?
  - Enter: −1, Don't: 0
  - Yes. Go to Step 3.
- Step 3. Firm 2 plays Don't. Is Enter optimal for Firm 1?
  - Enter: 2, Don't: 0
  - Yes. It is a Nash equilibrium!

Is that the only Nash equilibrium of the game?

No. There is another Nash equilibrium where Firm 2 enters but Firm 1 does not.[7] This entry game is a coordination-type game.

Game theory makes an interesting prediction. It predicts that the market will be a monopoly, but it doesn't predict which firm will be the monopoly. This turns out to have some implications for the empirical analysis of bookstores analyzed in Chapter 4.

[7] Again there are other equilibria but we will get to them. Be patient!

2.4.3 Many Firm Entry Game

We can make the game more general by having up to $\bar{N}$ firms choosing whether or not to enter.

- Players: $\bar{N} > 0$ firms
- Strategies: Enter, Don't Enter
- Payoffs:
  - Enter: $\frac{(a-c)^2}{b(N+1)^2} - F$, where $N \leq \bar{N}$, where N is the number of firms that choose Enter.
  - Don't Enter: 0

In this case we have a small number of firms, $\bar{N}$. Each firm has a entry cost F. Entry costs may include finding a retail space, developing the space to sell tires, contracting with wholesalers and manufacturers, etc. Once firms enter (or not), the market mechanism determines the profits each firm will make.

The profit function is on the complicated side. Notice that the profits the firms make in the market are determined by the number of firms that enter (N). The more firms that enter, the lower the profits. Competition drives down prices and profits. Moreover, profits could be so low that they are below the fixed cost of entry. In that case, the firms are better off not entering.

To see where the complicated profit function comes from, assume price is determined by the following function.[8]

$$p = \frac{a}{N+1} + \frac{Nc}{N+1} \tag{2.1}$$

where c is the **marginal cost of production**, a is a demand parameter, and N is the number of firms that enter the market. As the number of firms enter the price falls and it converges to marginal cost as N gets large. The first part is close to zero when N becomes big. The second part is close to c because the fraction of N over $N+1$ is close to 1 when N is big. The **marginal cost of production** refers to the incremental costs of selling a tire, this may include hourly wages and the wholesale cost of the tire itself.

If a firm enters the market their profits are determined by quantity that they sell (q) multiplied by their profit margin ($p - c$). Demand in the market is assumed to be determined by the following linear function $Q(p) = a - bp$. The parameter a determines the level of demand and the parameter b determines the sensitivity of demand to price. Demand falls more dramatically for a particular price increase if b is larger. Each firm in the market is identical and so they just split demand evenly, $q = \frac{Q}{N}$.

$$q \times (p - c) = \frac{(a-c)^2}{b(N+1)^2} \tag{2.2}$$

Multiplying demand by margin gives the profit function.

[8]This is the solution to a Cournot game with N symmetric firms which is a game discussed in Chapter 3.

2.4.4 Nash Equilibrium

What is the outcome of this game? A Nash equilibrium requires that each firm is playing its optimal strategy given the actions of all the other firms.

Each firm will enter if and only if the following inequality holds.

$$\frac{(a-c)^2}{b(N+1)^2} > F \tag{2.3}$$

This means that if there are N firms in the market in equilibrium. The inequality above must hold. If it didn't some of the firms would not enter the market. Also, it can't be profitable for another firm to enter. What if one more firm enters? In that case, profits for each firm must fall below the fixed cost of entry.

$$\frac{(a-c)^2}{b(N+2)^2} < F \tag{2.4}$$

If there are $N < \bar{N}$ firms in the market, then it must be profitable for all of those firms to enter, but not profitable for any more firms to enter.

2.4.5 Fixed Cost of Entry

Let's add one more complication to the game. Let the fixed costs be a function of the number of firms that enter, $F = \theta(N+1)^K$, where K and θ are some parameters of costs. This assumption states that when more firms are in the market their costs of entry are higher. It may be that land or facility space becomes more expensive when there are more firms looking to use the land or space to sell tires. This is the idea that different firms have different fixed costs of entry. In markets with a small number of firms, those that enter will have low fixed cost of entering. In markets with a large number of firms in equilibrium, the firms will have higher fixed cost of entering.

2.4.6 Equilibrium Number of Firms

The number of firms in the market is an integer, a counting number, 1, 2, 85, etc. Unfortunately, while these numbers are easy they are annoying to use for solving equations. To make solving the equation easier we will make the unrealistic assumption that the number of firms in the market is a real number.[9] Real numbers are useful because they have the property that there exists a solution to our equilibrium equation.

In equilibrium the following equality holds.

$$\frac{(a-c)^2}{b(N+1)^2} = \theta(N+1)^K \tag{2.5}$$

[9] Real numbers include integers and rational numbers (fractions) but also more exotic numbers like π or $\sqrt{(2)}$.

If N firms enter then for each firm they get the same profits from entering or not entering. The Nash equilibrium does not predict which N firms enter just that there will be N firms that enter. Solving, the equilibrium number of firms is $N = \left(\frac{(a-c)^2}{b\theta}\right)^{\frac{1}{K+2}} - 1$. Below we will use logs and it will look nicer.

We see that the equilibrium number of firms is increasing in the profitability of the market and decreasing in the customer's sensitivity to price.

2.4.7 Simulation of Entry Game in R

To understand how this game works, we can simulate entry into thousands of markets. In the simulation, there are 1,000 markets with a maximum of 10 firms. Demand and cost parameters vary from market to market as do entry costs.

The code uses the function `runif()`. This generates a set of random numbers that are uniformly distributed between 0 and 1. The probability of drawing any particular number between 0 and 1 is the same. To be able to replicate the results exactly it uses `set.seed()`.[10] The different values used are just made up.

```
> set.seed(123456789)
> M = 1000
> N_bar = 10
> a = 4 + 1*runif(M)
> b = 0 + 0.6*runif(M)
> c = 1 + 2*runif(M)
> theta = 0 + 0.05*runif(M)
> K = 0.9
```

We can put the solution to Equation (2.5) in code form. Given the equilibrium entry, we can then plug the numbers back into the pricing equation to determine equilibrium prices in each of the markets. This is an example of where **R** shines. We can write out something that looks pretty similar to the math, but is hiding a lot more complexity. These two lines actually determine the equilibrium number of firms and equilibrium prices for all 1,000 markets.

```
> N = (((a - c)^2)/(b*theta))^(1/(K+2)) - 1
> p = a/(N+1) + (N*c)/(N+1)
```

We can compare equilibrium prices to the theoretical relationship between prices and the number of firms. To determine the equilibrium prices we calculate the average price at each equilibrium level of the number of firms. To do this, we use the package `data.table`.[11] This package is very useful for doing calculations

[10] Computers don't actually generate random numbers. The numbers come from a complicated non-linear function. They look random, but if you know the previous number in the sequence and the function used, then you can exactly determine the next number in the sequence.

[11] The syntax used by `data.table` is different from base **R** and also from `tidyverse()`.

on lots of different subsets of the data. Here we want to do the calculation for each market with the same number of equilibrium entrants. Round the number of firms using `round()` so that the number of firms is an integer. The resulting data set `dt`, for data table, has two variables "N" and "p". The next line takes the average price for each market with N firms and creates a new data set `dt1`.

```
> library(data.table)
> dt = data.table(N = round(N),
+                 p = p)
> dt1 = dt[, .(p = mean(p)), by = N]
```

To calculate the theoretical equivalent, we take the average for the parameter values and calculate the price if the number of firms in the market was determined exogenously rather than in equilibrium. The last line uses the price formula from Equation (2.1).

```
> a_m = mean(a)
> c_m = mean(c)
> N_m = 1:N_bar
> p_m = a_m/(N_m + 1) + (N_m*c_m)/(N_m + 1)
```

Figure 2.1 presents the relationship between the number of firms in the market and the price. This is only an example, but it does show that the theoretical relationship between price and the number of firms can differ from the empirical relationship. The theoretical relationship shows that prices fall with competition. In the simulated data, the relationship between the number of firms and the price in the market is less negative, at least for a small number of firms. The reason is that the empirical relationship is an equilibrium where the firms are taking account of what the price will be when deciding to enter the market.

If we just naively take the number of firms in a market and regress that on price, we are not finding the true relationship between price and competition. We are estimating a relationship mediated by equilibrium decisions of firms. Once we acknowledge this, we have two choices. Throw up our hands and give up or think seriously about these equilibrium decisions and how the data is generated.

2.5 Empirical Analysis: Tire Markets using R

Bresnahan and Reiss wrote a series of papers where they thought carefully about the empirical implications of entry games. In one of those papers, the authors get data on the number of firms in geographically distinct towns over a wide variety of industries including retail tire stores.

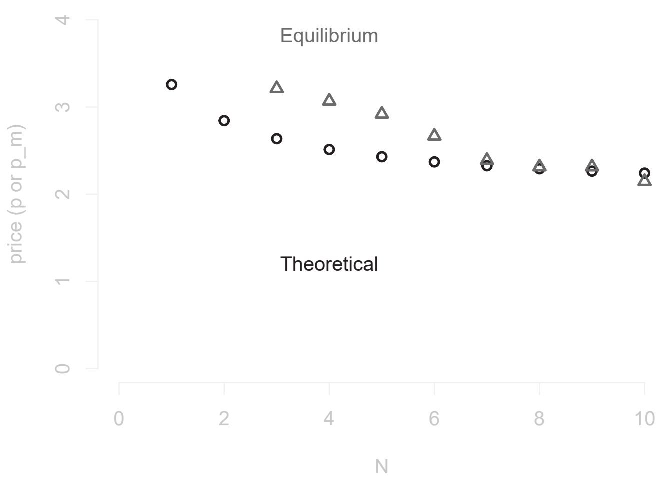


FIGURE 2.1
Scatter plot of prices from simulated data gives the number of firms (N) in the market. The equilibrium prices (`dt1$p`) (red triangles) fall less quickly than the theoretical prices (`p_m`) (black circles).

The section discusses the data on retail tire markets and how to take the model presented earlier in the chapter to the data. It estimates the parameters of the game and uses the game to simulate changes to policy such as reducing the costs associated with setting up a new firm.

2.5.1 Data

Those data provide information on the number of tire stores in each town, the population of each town, the number of commuters into the town, various economic indicators such as house prices and land prices and various demographic indicators such as age and family income.[12]

The code below also loads the data and plots the data. The data are in a csv file and is read in using `read.csv()`. The variable `dir` needs to be defined. This is set equal to the path where the data resides. The data is then plotted using `ggplot()` from the package `ggplot2` (or `tidyverse`).

[12]The data used here comes from Jeremy Bejara and his Github repository from a 2019 structural industrial organization course, `https://github.com/jmbejara/comp-econ-sp19/blob/master/lectures/5-14_Structural_IO_with_MLE/bresnahan-reiss-1991-discussion.ipynb`

```
> file = paste0(dir, "BresnahanAndReiss1991_DATA.csv")
> read.csv(file) |>
+   ggplot(mapping = aes(TPOP,
+                         TIRE)) +
+   geom_point() +
+   labs(x = "Total Population (000s log scale)",
+        y = "Number of Tire Stores") +
+   scale_x_log10()
```

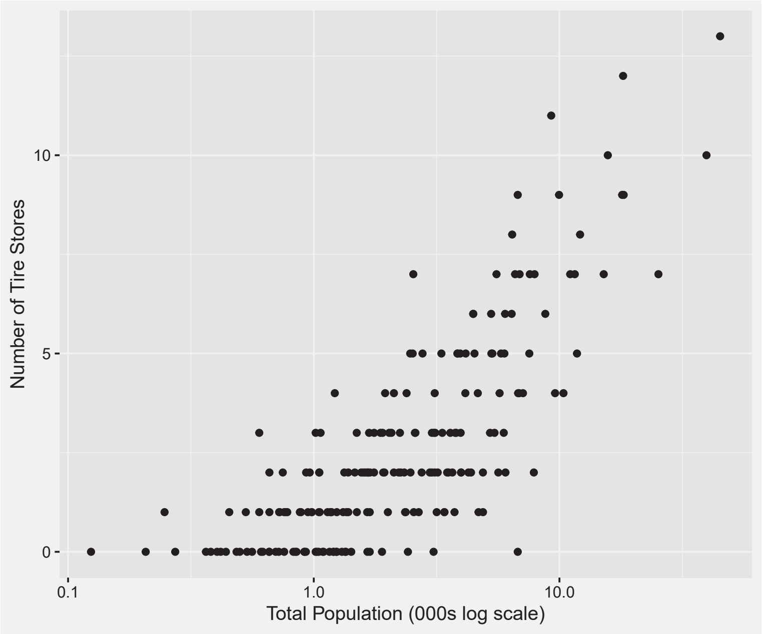


FIGURE 2.2
Scatter plot of the empirical relationship between the number of tire stores and the population of the town. It shows the general positive relationship with larger towns having more retail tire stores.

Our interest is figuring out what determines the number of stores in a market. Figure 2.2 shows that as the market gets larger, the number of firms increases.

2.5.2 Structural Model

We will follow Bresnahan and Reiss (1991b) and build a structural model. The basic idea is that the game theory provides the empirical relationship which we can then match to the data. From that matching, we can back out the parameters of the game theoretic model. Once we have the model parameters we can run policy simulations. At least that is the theory.

We can write Equation (2.5) in logs. We take logs of both sides of the equation and then rearrange to put the log of the number of firms on the left-hand side. This makes the equation look more like a linear regression equation.

$$
\begin{aligned}
&(a-c)^2 = b(N+1)^2\theta(N+1)^K \\
&\textit{or} \\
&(2+K)\log(N+1) = 2\log(a-c) - \log(b) - \log(\theta) \\
&\textit{or} \\
&\log(N+1) = \tfrac{2}{2+K}\log(a-c) - \tfrac{1}{2+K}\log(b) - \tfrac{1}{2+K}\log(\theta)
\end{aligned}
\tag{2.6}
$$

This transformation is useful for enabling us to use a standard **linear regression** estimator.

The relationship from the structural model states that the number of firms is increasing in average profits of the firm $(a-c)$ and decreasing in the substitution to other products (b) and the entry costs (θ). The rate of increase is determined by the parameter K.

2.5.3 Entry Estimator

Equation (2.6) states that in equilibrium the number of firms (log of the number of firms) is determined by factors determining demand size such as total population and per-capita income, factors determining costs such as wages, factors determining the slope of the demand function such as closeness of substitutes and finally factors determining the cost of entry such as property rental costs.

Unfortunately, it is not clear how our data maps into our theoretical parameters. We do have good measures for overall demand, like total population, but we don't have information on wages. We have some information about land prices which may affect rental prices and entry costs. We have information on the number of commuters which may be a measure of both market size and substitution out of the market.

2.5.4 Estimation in R

To do the estimation we can run a linear regression with the log of the number of retail tire stores against characteristics of the town such as the population, number of commuters, income, and land values. The code below reads in the data and then runs the regression. The `data.frame` called `data` removes any observations with missing values using `na.omit()`.[13] The code then runs two regressions. The first regression includes only the population of the town.

[13] **R** uses `NA` to represent missing values.

TABLE 2.4
OLS estimates of the relationship between market characteristics and the number of retail tire stores.

	Dependent variable:	
	log_stores	
	(1)	(2)
population	0.08***	0.07***
	(0.01)	(0.01)
cummuters		0.04
		(0.07)
income		0.13***
		(0.04)
land_value		0.07
		(0.19)
Constant	0.74***	−0.03
	(0.05)	(0.22)
Observations	202	202
R^2	0.36	0.40
Note:	*p<0.1; **p<0.05; ***p<0.01	

```
> data = read.csv(file) |>
+   mutate(
+     log_stores = log(TIRE + 1),
+     population = TPOP,
+     cummuters = OCTY,
+     income = PINC,
+     land_value = LANDV
+   ) |>
+   na.omit()
> lm1 = lm(log_stores ~ population, data)
> lm2 = lm(log_stores ~ population + cummuters +
+             income + land_value, data)
```

Table 2.4 shows the empirical relationship between number of stores and various economic factors. It shows that there is strong positive relationship between market size and the number of firms. The empirical relationship between other factors is not obvious, at least in this data.

2.5.5 Structural Estimation

The estimates presented in Table 2.4 provide the empirical relationship between the number of firms in the market and observed characteristics of the market. We are interested in mapping those estimates into our game theoretic model. Assume that the observed relationship captured by the second column of Table 2.4 is generated by the game described above (Equation (2.6)). Specifically, it represents a set of Nash equilibria of that game.

We will make the following assumptions.

- $a = a_0 \exp(population)^{a_1} \exp(commuters)^{a_2} \epsilon_a$
- $c = \epsilon_c$
- $b = b_0 \exp(income)^{b_1} \epsilon_b$
- $\theta = \theta_0 \exp(\text{land\_value})^{\theta_1} \epsilon_\theta$

That is the demand level is a function of population and commuters and some unobserved characteristic (ϵ_a). You will see in a sec why we wrote this down in such a weird way.[14] The marginal cost is unobserved. We will be forced to assume that $\epsilon_c = 0$ for what we do below.[15] The slope of demand is a function of income and an unobserved characteristic (ϵ_b). Lastly, the cost of entry is a function of land values and an unobserved characteristic (ϵ_θ).

Given these assumptions and the regression results in Table 2.4 we have the following relationships.[16] The coefficient estimates are on the changes in the values of the observed characteristics.

$$\begin{array}{l} 0.07 = \frac{2}{2+K} a_1 \\ 0.04 = \frac{2}{2+K} a_2 \\ 0.13 = -\frac{1}{2+K} b_1 \\ 0.07 = -\frac{1}{2+K} \theta_1 \end{array} \tag{2.7}$$

We can't uniquely determine the parameter values from these equations. In order to move forward we will assume that $a_1 = 1$. Given this assumption $K = 0.86$, $a_2 = 0.06$, $b_1 = -0.37$ and $\theta_1 = -0.20$. The parameter estimates are determined relative to the coefficient on population.

We have one more empirical relationship up our sleeve. The constant in the regression is -0.03. Given our assumptions, we have that $0.70 \log(a_0) - 0.35 \log(b_0) - 0.35 \log(\theta_0) = -0.03$. Again, we don't have enough information to pin down all the parameters. If we assume that $a_0 = b_0 = 1$, then $\theta_0 = 1.09$.

---

[14] In economics we generally refer to this specification as a Cobb-Douglas function.

[15] Given that we don't observe marginal costs our policy estimates will be OK as long as we don't try to model big changes.

[16] We are just taking the coefficient values from Model (2) and not worring about how well those coefficients are estimated.

2.5.6 Policy Simulation in R

What happens if the market size increases or if entry costs fall?

In this case to determine the equilibrium number of tire stores and equilibrium prices we create two functions using `function()`. These functions take in values for log of total population, log of commuters, log of income, and log of land values. They are based on Equations (2.6) and (2.1) respectively. The `p()` includes the equilibrium number of firms through the function `N()`.

```
> N = function() {
+   exp((2/(2 + K))*(log(a_0) + a_1*population + a_2*commuters) -
+     (1/(2 + K))*(log(b_0) + b_1*income) -
+       (1/(2 + K))*(log(theta_0) + theta_1*land_value) - 1)
+ }
> p = function() {
+   exp((log(a_0) + a_1*population + a_2*commuters))/
+     (N()+1)
+ }
```

We can use the estimates of the parameters of the game theoretic model to understand the likely impact of various policy changes.

```
> K = 0.86
> a_0 = 1
> a_1 = 1
> a_2 = 0.06
> b_0 = 1
> b_1 = -0.37
> theta_0 = 1.09
> theta_1 = -0.20
```

These are our baseline estimates given the observed data and our baseline estimates for the number of firms and the price in the market, `N0` and `p0` respectively.[17]

```
> population = mean(data$TPOP)
> commuters = mean(data$OCTY)
> income = mean(data$PINC)
> land_value = mean(data$LANDV)
> N0 = N()
> p0 = p()
```

Consider a policy that increases commuters by 10%. This may be a policy that makes it easier to drive to the town or increases railway capacity or

[17] Defining variables outside any function means that they are **global** variables that can be accessed by any function. In this case those variables can be accessed by `N()` and `p()`.

reduces parking costs downtown or requires government workers to return to the office. Note the weird way that the change comes in, this is because the measures are in logs.

```
> commuters = mean(data$OCTY) + log(1.1)
> N()/N0
  [1] 1.004007
> p()/p0
  [1] 1.002058
```

The policy has a very small effect, increasing the number of firms in the market by 0.4% and prices by about 0.2%.

Alternatively, consider a policy that increases the population by 10%. This might be something that increases the connectedness of the town to near by towns or allows more housing to be built.

```
> population = mean(data$TPOP) + log(1.1)
> commuters = mean(data$OCTY)
> N()/N0
  [1] 1.068922
> p()/p0
  [1] 1.034692
```

This policy increases the number of firms by 7% and prices by about 3.5%. Both of these policies increase the number of firms in the market but also increase the retail price of tires. Demand increases, which increases prices, but the impact is mediated by the equilibrium change in the number of firms and competition.

Lastly, consider a policy that decreases entry costs by 10%. This may be a policy that makes it easier for a retail tire store to enter the market, like reducing the permits required to set up the store or construct the store.

```
> population = mean(data$TPOP)
> theta_0 = 0.9*theta_0
> N()/N0
  [1] 1.037526
> p()/p0
  [1] 0.9667752
```

This policy increases the number of firms by 4% and reduces prices by 3%.

We may naively expect a policy that reduces entry costs to have a significant impact on both entry and prices. Both effects are mitigated in equilibrium. As it becomes cheaper to enter the market, firms understand that entering may not be that profitable because prices will fall. Similarly, because entry response is modest the price response is modest.

2.6 Discussion and Further Reading

In a series of papers, Tim Bresnahan and Peter Reiss showed how game theory could be used to improve empirical analysis of market structures. This chapter uses data from Bresnahan and Reiss (1991b) to illustrate a structural model of firm entry in the retail tire market. In the original paper, the authors examine a number of markets and include information on prices for the retail tire market.

This chapter introduces the idea of a Nash equilibrium. In the entry game, firms will enter as long as it is profitable given all the other firms that will also enter. Because firms can strategically respond to the decisions of other firms, policies that we may naively believe to significantly lower prices, may not.

The chapter introduces the idea of using a structural model to interpret the data. It assumes that the observed relationships in the data are a Nash equilibrium of a particular entry game. While a lot of questionable assumptions are required given the data available, the analysis helps us to understand why various policies may have pretty modest impact on the number of firms in the market and the prices of tires.

In the FTC's case against Staple's acquisition of Office Depot, the FTC presented evidence that there was a positive relationship between prices and the number of retail office supply stores (Ashenfelter et al., 2006). A concern with the analysis is that both the prices and the number of stores may be driven by other factors such as land values (Manuszak and Moul, 2008). The game theory can help us understand potential pitfalls in the empirical analysis and which econometric methods may provide solutions. Chapter 4 revisits firm entry games by analyzing the decisions of Barnes & Noble and Borders to enter various markets in the United States.

3

Oligopoly

3.1 Introduction

When most people, even most economists, discuss competition they have in mind the model presented in Econ 101. There are many firms all adding a small amount to the market. No firm has control over the price that they receive. Each firm can leave or enter the market at will, so any large profits get bid away as more firms enter. In this world, prices are equal to marginal cost and the amount of goods is such that consumers and firms cannot be made better off (without one or the other being made worse off). This is not what industrial organization economists mean when they discuss competition. Economists that work in antitrust and competition policy, distinguish between firms who work together to determine price and firms that work independently of each other.

This chapter presents the standard model of how firms set prices and output when they are doing so independently of each other. Chapter 8 will return to oligopoly and allow less independence in price setting. The chapter presents three standard models of competition, Cournot, Bertrand, and Hotelling.[1] The chapter uses **R** to simulate the Nash equilibrium of a Cournot game with three firms.

The Hotelling model is used to understand pricing of hamburgers in Santa Clara County California in the late 1990s. In particular, the chapter analyzes data on the price of the McDonald's Big Mac and the Burger King Whopper at the outlets throughout Santa Clara County. The model allows us to estimate how prices are affected by competition between firms that are differentiated by location. If one franchisee purchased all the McDonald's outlets in Santa Clara County, what would happen to the price of both the Big Mac and the Whopper in the county?

[1] Confusingly the Hotelling model is often called a Bertrand model. While this chapter presents two examples of a price setting game, many people assign the Bertrand name to any static price setting game where the products are not homogeneous.

DOI: 10.1201/9781003351603-3

3.2 Cournot's Model

In the 1850s a French applied mathematician made a far fetched claim. Augustine Cournot suggested having competing firms does not necessarily mean that prices will be equal to marginal cost. Tabarnak! How could this be? It is almost by definition that prices equal marginal cost in economics. Some fifty years later, another Frenchman, François Bertrand, argued that Cournot was full of merde. Even with just two firms, prices would equal marginal cost.

The section presents a two-firm version of the Cournot model and then generalizes that to a N-firm version. It uses **R** to numerically simulate the Nash equilibrium in a three-firm version of the model.

3.2.1 Two Firm Model

Formally we have the following game where we have two firms choosing quantity (q_i) and contemplating the impact of their collective choices on profits.

- Players: Firm 1 and Firm 2.
- Strategies:
  - Firm 1: $q_1 \geq 0$
  - Firm 2: $q_2 \geq 0$
- Payoffs:
  - Firm 1: $p(q_1, q_2) \times q_1 - c(q_1)$
  - Firm 2: $p(q_1, q_2) \times q_2 - c(q_2)$

The first part of Firm 1's payoff is revenue. The price $p(q_1, q_2)$ is a function of each firm's output in the market. The price is multiplied by the output produced by Firm 1 which is q_1. The second part is costs which is a function of the amount of output produced. To keep things simple let $p(q_1, q_2) = a - b \times (q_1 + q_2)$ and $c_j(q_j) = c \times q_j$. That is, we have linear demand and constant marginal cost. Assume also that $a - c > 0$. This will become important later.

Cournot's game assumes that both firms make identical goods. This may be a good model of a wheat market or an electricity market. In such markets each firm decides how much to produce and supplies to a centralized exchange that takes in demand and determines the price everyone gets. In the game, each firm does not know how much the other firm produces.

3.2.2 Best Response

One issue that people new to game theory find very confusing is the idea of a **best response function**. In the presentation of the game above, it is clearly stated that each firm does not know the other firm's choice when choosing its action. Then five minutes later we claim that firms have a function such that its action is a best response to the other firm's action. How can they both not know what the other firm is doing and have a best response to it? Makes no sense.

Both statements can be true because they refer to different concepts. One is a description of the game and the other is a description of the algorithm used to find the Nash equilibrium of the game. The games discussed in this part of the book assume the players choose actions once and simultaneously. In this sense, each firm does not know how much the other firm is producing. The best response function is an analytical tool used to find the Nash equilibrium. The Nash equilibrium is where each firm chooses the optimal output given the output chosen by the other firm. That is the Nash equilibrium is where each firm's output choice is a best response to the other firm's output choice.

The best response function is the solution to the following optimization problem.

$$\max_{q_1} \quad (a - bq_1 - bq_2)q_1 - cq_1 \tag{3.1}$$

The solution to this optimization problem is given by the first order condition.

$$\begin{aligned} & a - bq_1 - bq_2 - c - bq_1 = 0 \\ & or \\ & q_1 = \tfrac{a-c-bq_2}{2b} \end{aligned} \tag{3.2}$$

The best response function states that Firm 1's quantity is increasing in the difference $a - c$, decreasing in the quantity of Firm 2 (q_2) and decreasing in the willingness of customers to substitute out of the market (b). In this model, the actions of the two players are strategic substitutes, in words, when one firm increases its output, the other firm responds by decreasing its output.

3.2.3 Nash Equilibrium

The Nash equilibrium is where each firm is playing the best response to the other firm. That is, where the first order condition for Firm 1 (Equation (3.2)) and the equivalent condition for Firm 2 both hold. To solve for the equilibrium we need to solve for two unknowns from a system of two linear equations. Simply counting the number of unknowns and the number of independent linear equations we know that the solution exists and is unique.

In this case, we can use a nice trick to solve it. Because both firms are identical then it must be that in equilibrium they produce the exact same amount. That is $q_1 = q_2 = q$. Substituting this into Equation (3.2) allows us to find the solution.

$$\begin{aligned} & q = \tfrac{a-c-bq}{2b} \\ & 2bq = a - c - bq \\ & 3bq = a - c \\ & q = \tfrac{a-c}{3b} \end{aligned} \tag{3.3}$$

This is each firm's output in equilibrium. Equation (3.3) states that output will fall when marginal cost (c) increases and when substitutability (b) increases. In order for equilibrium output to be positive it must be the case that $a - c > 0$, which is an assumption made above.

If we substitute this back into demand we can determine price. We have two firms so we need to substitute back both equilibrium quantities.

$$\begin{aligned} p &= a - 2b\tfrac{a-c}{3b} \\ &= \tfrac{3a-2a+2c}{3} \\ &= \tfrac{a+2c}{3} \end{aligned} \tag{3.4}$$

Equation (3.4) states that price will increase with marginal cost, but are they higher than marginal cost?

$$\begin{aligned} & p > c \\ \Leftrightarrow\ & \tfrac{a+2c}{3} > c \\ \Leftrightarrow\ & a + 2c > 3c \\ \Leftrightarrow\ & a - c > 0 \end{aligned} \tag{3.5}$$

As long as there is positive output in the market, prices will be greater than marginal cost. Mayhem. Cats and dogs living together!

3.2.4 Cournot Model with N Firms

What happens as the number of firms increase? Do we get back to perfect competition?

In this case, Firm 1's best response becomes.

$$q_1 = \frac{a - c - b\sum_{j=2}^{N} q_j}{2b} \tag{3.6}$$

Again to solve for equilibrium with symmetric firms we can do the trick of setting all the output levels to be the same.

$$\begin{aligned} & q = \tfrac{a-c-b(N-1)q}{2b} \\ & 2bq = a - c - b(N-1)q \\ & (N+1)bq = a - c \\ & q = \tfrac{a-c}{(N+1)b} \end{aligned} \tag{3.7}$$

So in equilibrium, output is decreasing proportionately with the number of firms in the market.

Substituting equilibrium quantities back into demand we can determine prices. Remember we have to multiply by N.

$$\begin{aligned} p &= a - bN\tfrac{a-c}{(N+1)b} \\ &= \tfrac{(N+1)a-Na+Nc}{N+1} \\ &= \tfrac{a+Nc}{N+1} \\ &= \tfrac{a}{N+1} + \tfrac{N}{N+1}c \end{aligned} \tag{3.8}$$

Look familiar? It is the equation used to determine price in Chapter 2.

We can see that as N gets large, p converges to c. That is Cournot's model does give perfect competition but only for a large number of firms in the market.

3.2.5 Cournot Model in R

In order to better understand how the model works it helps to program it up in **R**. The function `price_cournot()` is used to determine the price given market output `q`. The function `br_cournot()` determine the firm's output given the output of the other two firms. This function is based on Equation (3.6). The notation `q[-i]` is used to refer to all the elements of `q` except the `i`th element.

```
> price_cournot = function(q) a - b*sum(q)
```

```
> br_cournot = function(q, i) {
+   if (price_cournot(q) > 0) {
+     return(max(c(0, (a - c[i] - b*sum(q[-i]))/(2*b))))
+   } else {
+     return(0)
+   }
+ }
```

The best response function checks to make sure prices and quantities are positive. It is based on Equation (3.6).

3.2.6 Solve for the Nash Equilibrium with R

The algorithm below looks for an equilibrium where the best response's of each firm lead to the same quantity. The algorithm chooses a starting level of output and then calculates the best response for each firm to get a new level of output. It then checks whether the new level of output is the same as the old level of output. If it is, then we have a Nash equilibrium. Remember a Nash equilibrium is where each firm is choosing a level of output that is a best response to all the other firms. If the new level of output is different from the old level of output, the algorithm sets the old level of output to the output just calculated and finds the best response to that amount. The algorithm stops when the new and old amounts become equal or close to equal.

In the code we have a default maximum number of iterations (`maxit = 100`) and a default level of convergence (`epsilon = 1e-5`).[2] The code uses `epsilon` to refer to a small number. It uses a `while()` loop to run until one of the conditions fails to hold.

The initial value for the output is determined by Equation (3.7) assuming all firms in the game have the same costs. The algorithm then calculates the

[2] `1e-5` is a way to write small numbers, in this case 0.00001.

best response for each firm and checks whether the new output is the same as the old output. The algorithm sets the old output to the new output and calculates the best response to that output. The algorithm continues until the difference between the new and old output is less than `epsilon` or the number of iterations exceeds `maxit`. The algorithm returns the output in equilibrium and whether the algorithm converged.

```
> ne_cournot = function(maxit=100, epsilon=1e-5,
+                       trace=FALSE, converged=TRUE) {
+   diff = 10000 # some big number
+   iter = 1
+   N = length(c)
+   q_old = rep((a - mean(c))/((N+1)*b), N)
+   # initial values for q_old
+   while(diff > epsilon & iter < maxit) {
+     q_new = rep(0, N)
+     for(i in 1:N) {
+       q_new[i] = br_cournot(q_old, i)
+     }
+     diff = sum(abs(q_new - q_old))
+     iter = iter + 1
+     if(trace) {
+       print(diff)
+       print(iter)
+     }
+     q_old = q_new
+   }
+   if(iter == maxit) {
+     converged = FALSE
+   }
+   return(list(q_star = q_new, converged = converged))
+ }
```

This algorithm is not super sophisticated and it takes advantage of uniqueness of the result. The algorithm does allow us to simulate more interesting models than the simple symmetric-firm model presented above.

3.2.7 Simulation of Cournot Model in R

The following simulation allows the costs to vary between firms and shows how variation in marginal costs leads to differences in market share for firms.

```
> set.seed(123456789)
> N = 3
> a = 0.5
> b = 0.2
```

```
> c = a*runif(N) # so costs vary between firms.
> q_star = ne_cournot()$q_star
> p_star = price_cournot(q_star)
```

The results are as follows. The marginal costs (`c`) are determined randomly. The output (`q_star`) and the price (`p_star`) are determined by the equilibrium algorithm.

```
> c
  [1] 0.3465879 0.3364405 0.3269508
> q_star
  [1] 0.1545377 0.2052715 0.2527168
> p_star
  [1] 0.3774948
```

In this example Firm 1 has the highest costs and the lowest quantity, while Firm 3 has the lowest cost and the highest quantity.

This simulation illustrates a standard result of the Cournot model, the less efficient (higher marginal cost) firms have lower market share and the more efficient (lower marginal cost) firms have higher market share.

It also makes it clear that less efficient firms could be in the market. There is nothing about the basic Cournot game that forces them to leave the market.

3.3 Bertrand's Model

Bertrand was having none of it. Perfect competition was not some edge case. In Cournot's game the firms choose quantity, but what happens if the firms choose price?

The section presents a two-firm version of Bertrand's game. Think about two ice-cream stands that are next to each other. Each stand displays the price of a cone of ice-cream cone. If one ice-cream stand charges more than the other, then everyone will buy from the cheaper stand. If both stands charge the same price, then they split the market.

3.3.1 Two-Firm Game

This time the two firms choose a price and payoffs are determined by the quantity which is a function of both prices.

- Players: Firm 1 and Firm 2.
- Strategies:

  - Firm 1: $p_1 \geq 0$
  - Firm 2: $p_2 \geq 0$

- Payoffs:
  - Firm 1: $p_1 \times q_1(p_1, p_2) - c_1 \times q_1(p_1, p_2)$
  - Firm 2: $p_2 \times q_2(p_2, p_1) - c_2 \times q_2(p_2, p_1)$

Again the payoffs for each firm are revenue less costs. This time, the firm chooses the price to charge and the market mechanism determines how much quantity the firm will sell.

$$q_1 = \begin{cases} 1 & \text{if } q_1 < q_2 \\ 0.5 & \text{if } q_1 = q_2 \\ 0 & \text{if } q_1 > q_2 \end{cases} \tag{3.9}$$

Let's assume that the market size is 1. Equation (3.9) says that if you have the lowest price, then everyone buys from you and if you have the highest price then no one buys from you. If both firms have the same price, they split the market.

This is a market where it is very easy for customers to substitute. One slight reduction in price can cause the whole market to shift.

A real-world example might be a supply contract request for quotes. Which ever firm offers the lowest price gets the full supply contract.

3.3.2 Nash Equilibrium

Consider a simple case where both firms have the same marginal cost, $c_1 = c_2 = c$. The unique Nash equilibrium is $p_1 = p_2 = c$.

To confirm that it is an equilibrium, assume that $p_1 = c$. What happens if $p_2 > c$? In this case, Firm 2's output is 0 and profits are 0. What if $p_2 = c$. In this case, Firm 2's output is 0.5, but Firm 2's profits are 0. Lastly, if $p_2 < c$, then Firm 2's output is 1, but Firm 2's profits are negative. So Firm 2 is indifferent between choosing $p_2 = c$ or $p_2 > c$. Given that Firm 2 cannot do better than the Nash equilibrium, this confirms $p_2 = c$ is optimal for Firm 2. We can make the exact same argument for Firm 1.

It is not only an equilibrium but a unique equilibrium. A candidate equilibrium is, $p_1 = p_2 > c$. Again, keep Firm 1's price at $p_1 > c$. If Firm 2 charges $p_2 > p_1$, then Firm 2's profits are 0. If Firm 2 charges $p_2 = p_1$ then Firm 2's profits are positive, $0.5(p_1 - c)$. If Firm 2 charges $p_2 = p_1 - \epsilon$ where ϵ is some small number, then Firm 2's profits are $(p_1 - c - \epsilon)$. These profits are a lot higher. Sure price is slightly lower, but Firm 2 went from selling to half the market to selling to the whole market. In this game there is a huge incentive to slightly undercut your rival in this market. Because of this incentive, there is no other Nash equilibrium.

Bertrand proved his point. With just two firms, price equals marginal cost, an important characteristic of perfect competition.

What happens in Bertrand's model if the two firms have different marginal costs (c_j)?

3.4 Hotelling's Model

To go from perfect competition being some edge case to it being a constant of the model all we needed to do was assume firms choose price instead of quantity! Nonsense. If you look more closely, the models proposed by the two Frenchmen are very different from each other. In particular, the demand in the Bertrand model is very particular.

Early in the Twentieth Century, the American statistician and economist, Harold Hotelling, suggested a compromise. He suggested a model where firms choose price but where demand was not nearly so particular as Bertrand assumes.

The section presents Hotelling's original model. It then presents a model of differentiated goods with linear demand and determines the Nash equilibrium in prices for that model.

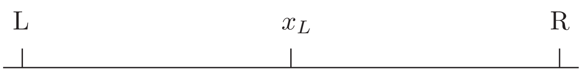


FIGURE 3.1
Hotelling's line with Firm L located at 0 and Firm R located at 1. Everyone "living" left of x_L purchases from Firm L.

3.4.1 Hotelling's Line

Figure 3.1 represents Hotelling's game. There are two firms L and R. Customers for the two firms "live" along the line. Consider two frozen custard (ice cream) places located at each end of a beach board walk. Your beach chairs may be located closer to one frozen custard place than the other. Customers prefer to go to the closer firm if the products and prices are otherwise the same. Hotelling's key insight is that while firms often compete by selling similar products, these products may not be identical. Moreover, some people may prefer one product to the other. Some people actually prefer Pepsi to Coke. The location on the line represents how much the customer prefers L to R.

The line and the distance between the customer and the firm represent how willing they are to purchase from a particular firm. Importantly, it represents how much the two firm's products are substitutes for each other. In the hamburger example presented below, we use the actual distance between stores to measure substitutability but more generally Hotelling's line is a metaphor for how similar or different two products are from each other. The closer the two firms are, the easier it is to substitute between them and the lower the price is likely to be.

3.4.2 Differentiated Goods Game

Now consider a game where there are N firms that all sell a product that is similar but not the same as each other. For example, a set of hamburger restaurants in Santa Clara. While the Big Mac may be the same, the location of each McDonald's outlet is different.

- Players: N firms where each firm has a location $\{x_i, y_i\} \in \Re^2$.
- Strategies: $p_i \geq 0$ for all $i \in \{1, ..., N\}$
- Payoffs: $(p_i - c_i) \times q_i(p_i, p_{-i})$

Each of the N firms chooses a price p_i. The profits are the price less marginal cost (c_i) multiplied by the quantity sold q_i. This quantity is determined by both the price the firm chargers p_i, by all the prices all the other firms charge p_{-i} and by the distances between the firms.[3]

To make things simpler and more concrete, assume that there are just two firms i and j, Firm i's sales are affected by p_i and p_j in the following way.

$$q_i(p_i, p_j) = \alpha + \beta p_i + \frac{\gamma p_j}{d_{ij}} + \epsilon_i \tag{3.10}$$

It is a linear demand model similar to the model presented earlier in the chapter. The quantity sold by Firm i is a function of the price Firm i charges, the price Firm j charges, and some unobserved term ϵ_i. Assuming, $\beta < 0$, then the higher the price the lower the demand for i's product. Demand is also a function of the competitor's price p_j. The extent of this is determined by γ, which is positive. The demand for i is higher when the competitor charges a higher price. The extent of the competitor's price matters depends on the distance between the outlets. The larger that distance (d_{ij}), the less the two firms compete with each other for customers. Again, d_{ij} could represent physical distance or just a measure of the difference between the two products.[4]

3.4.3 Best Response

Given all this, what will be the price in the market? We assume that the price is determined by the Nash equilibrium. The Nash equilibrium is the price such that Firm i is unwilling to change their price given the price charged by Firm j, and Firm j is unwilling to change their price given Firm i's price.

[3] $-i$ means not i.

[4] We will Euclidean distance, $d_{ij} = \sqrt((x_i - x_j)^2 + (y_i - y_j)^2)$. This is the as-the-crow-flies distance.

Firm i's problem is as follows.

$$max_{p_i} \quad (p_i - c_i)(\alpha + \beta p_i + \tfrac{\gamma p_j}{d_{ij}}) \tag{3.11}$$

The solution to the optimization problem is the solution to the first order condition.

$$(\alpha + \beta p_i + \frac{\gamma p_j}{d_{ij}}) + \beta(p_i - c_i) = 0 \tag{3.12}$$

Firm j's problem is similar. The two equations from the first order conditions are the best response functions for each firm.

3.4.4 Nash Equilibrium

Given the two equations derived from the first-order conditions, we can write down a system of equations. Each firm's best response function is as follows.

$$\begin{array}{l} p_i = \frac{c_i}{2} - \frac{\alpha}{2\beta} - \frac{\gamma p_j}{2\beta d_{ij}} \\ p_j = \frac{c_j}{2} - \frac{\alpha}{2\beta} - \frac{\gamma p_i}{2\beta d_{ij}} \end{array} \tag{3.13}$$

The prices are higher when costs are higher. Prices are strategic complements. While there is a negative sign in front of γ, we said above that β is negative. When p_j increases, then Firm i's best response is to *increase* p_i. Again, how much the two firm's interact depends on γ and the distance d_{ij}.

3.5 Empirical Analysis: Hamburger Competition with R

When he was doing his PhD at Stanford, Wash U Marketing professor, Raph Thomadsen, decided to study competition for hamburgers in Santa Clara County. Raph was interested in competition between the two big chains, McDonald's and Burger King as well as competition within the chain brands.

To study competition he got information on each outlet from the health department. Most importantly he got the outlet's location. He then physically visited each outlet determining the price of the hamburgers offered and whether the outlet had other features like a drive through or a playland.

This section models the prices for hamburgers in the late 90s in Santa Clara County using the model presented earlier.

3.5.1 Data

The data is provided by Raph Thomadsen and used in his RAND paper (Thomadsen, 2005).[5] For each outlet there is information about the brand,

[5] There is a slight descrepancy between the coordinates in the code and the location of the outlets using Google maps.

ownership structure, features of the outlet, age of the outlet, and price of the sandwich.

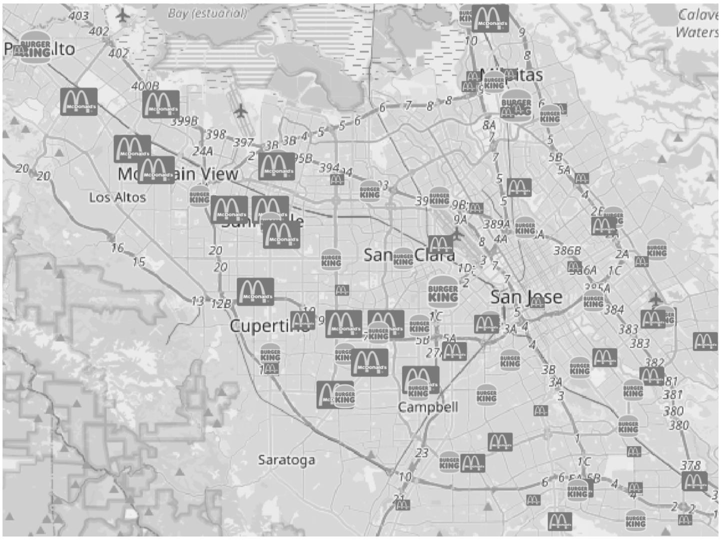


FIGURE 3.2
McDonald's and Burger King outlets in Santa Clara County. The size of the logo indicates whether the price of the sandwich is in the top third of prices, the middle third, or the bottom third.

Figure 3.2 is created using the `leaflet` package in **R** with the coordinates from the code and importing the icons. The figure gives some idea how competition works in Santa Clara county. There is a logo at each brand's location. The size of the logo represents the price charged for the signature sandwich. The higher the price, the bigger the icon. The outlets up near the Stanford campus have fairly high prices. Often a Burger King is paired with a McDonald's, but neither has really low prices. Prices on the west and south side of San Jose seem lower than on the east side and north east side.

You may be surprised by how much prices for a Big Mac vary across one city. You may also be surprised that the competition of the most interest is competition between McDonald's outlets. How can that be? Doesn't the McDonald's corporate headquarters set prices for the Big Mac? No. It depends. Some outlets in this data are owned by the McDonald's corporation and for those, yes, corporate headquarters would have a lot of say over price. But most of the outlets are owned by franchisees. Under California law, corporate headquarters is restricted in what it can require of them. Headquarters cannot determine the price charged by the outlet for its sandwiches.

3.5.2 Estimation (Part 1)

It is traditional when estimating the parameters of a game theory model, to first present estimates from a standard model like a linear regression model. The game theory helps us think about what to put into that regression model but also makes clear that we should be careful in interpreting the results. The raw prices are adjusted to make the regression results look nicer.

The data used in the rest of the analysis is restricted to outlets that are not corporate owned. The issue is that their pricing strategies may look very different from the franchise locations.

```
> file = paste0(dir,"outlets_gt_ch3.csv")
> data = fread(file) |>
+   mutate(
+     LPrice1 = log(100*(Price - 2.49)),
+     BK = BK == 1
+   )
> lm1 = lm(LPrice1 ~ BK, data)
> lm2 = lm(LPrice1 ~ BK + Playland +
+             DriveThru + Mall, data)
> lm3 = lm(LPrice1 ~ BK + Playland +
+             DriveThru + Mall + Race + Male, data)
```

Table 3.1 presents some regression results of log of sandwich price on characteristics of the outlet. This set up seems to match relatively closely to the table presented in Thomadsen (2005). Prices are lower for The Whopper (on average) and are lower at outlets with various amenities and in Malls as well as in various locations based on demographic characteristics. The coefficients on most of the characteristics are not statistically significantly different from zero. This is probably due to the small sample size. One coefficient that is statistically significantly different from zero is the dummy on being a Burger King outlet. The Whopper is cheaper than the Big Mac.

3.5.3 Empirical Equilibrium

As we did in the previous chapter, we will assume that the prices observed in the data are determined by the Nash equilibrium of the game described above. The implication is that for the whole set of prices we see from all the outlets, the whole set of best responses must hold. In the final data set we have prices and information for 79 outlets in Santa Clara County.

Our estimation problem is to find the α, β, and γ such that our 79 best response equations hold given the 79 prices we observe and the distance between each location. Actually, we are going to make our empirical model slightly more complicated by adding parameters for observed characteristics of the outlet.

$$2\beta p_i - \beta c_i + \mathbf{X}_i'\alpha + \frac{\gamma p_j}{d_{ij}} + \epsilon_i = 0 \tag{3.14}$$

TABLE 3.1
OLS estimates of the equilibrium relationship between price and observed characteristics of the outlet and their location.

	Dependent variable:		
	LPrice1		
	(1)	(2)	(3)
BK	−0.21***	−0.23***	−0.22***
	(0.05)	(0.05)	(0.05)
Playland		−0.08	−0.07
		(0.06)	(0.06)
DriveThru		−0.02	−0.04
		(0.06)	(0.06)
Mall		−0.13	−0.12
		(0.10)	(0.10)
Race			0.50
			(0.64)
Male			0.16
			(0.19)
Constant	4.54***	4.59***	4.50***
	(0.03)	(0.06)	(0.11)
Observations	79	79	79
R^2	0.19	0.22	0.23
Note:	*p<0.1; **p<0.05; ***p<0.01		

where $\mathbf{X}_i$ is a vector of observed characteristics for outlet i, e.g. $(1,\ brand_i,\ drivethru_i)$, where $brand_i$ indicates the brand of the outlet and $drivethru_i$ is 1 if the outlet has a drive through and 0 otherwise.[6] The parameter α is now a vector of parameters $(\alpha_0, \alpha_1, \alpha_2)$.[7]

Writing this down for all the outlets we have 79 equations that need to hold $(J = 79)$. How much they affect each other depends on the parameter γ and the distance between each outlet, d_{ij}. The other thing to notice is the assumption that the demand parameters $(\alpha,\ \beta,\ \gamma)$ are the same for every outlet. What differs between the outlets is the observed characteristics $\mathbf{X}_i$, their marginal cost c_i and the distance to the other outlets d_{ij} and unobserved characteristics

[6]The matrix notation $\mathbf{X}$ to represent a block of data. In this case, each row is an outlet and each column is a characteristic of the outlet.

[7]We are using matrix notation where two vectors of three elements $x'y$ means $(x_1 y_1 + x_2 y_2 + x_3 y_3)$. As you can see the matrix notation is a lot more compact.

of the outlet ϵ_i.

$$\begin{array}{l} 2\beta p_1 - \beta c_1 + \mathbf{X}_1'\alpha + \gamma \sum_{j=2}^{J} \frac{p_j}{d_{1j}} + \epsilon_1 = 0 \\ 2\beta p_2 - \beta c_2 + \mathbf{X}_2'\alpha + \gamma (\frac{p_1}{d_{12}} + \sum_{j=3}^{J} \frac{p_j}{d_{2j}}) + \epsilon_2 = 0 \\ \ldots \\ 2\beta p_J - \beta c_J + \mathbf{X}_J'\alpha + \gamma \sum_{j=1}^{J-1} \frac{p_j}{d_{Jj}}) + \epsilon_J = 0 \end{array} \tag{3.15}$$

This is quite a mess. It is a lot more compact to write this out using matrix notation.

To see what is going on look at the case where there are just two firms ($J = 2$). Also let's assume that marginal costs are zero and there are no observed characteristics or unobserved characteristics of the outlets.

$$\begin{array}{l} 2\beta p_1 + \gamma \frac{p_2}{d_{12}} = 0 \\ 2\beta p_2 + \gamma \frac{p_1}{d_{21}} = 0 \end{array} \tag{3.16}$$

We can write this out with matrices.

$$2\beta \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} + \gamma \begin{bmatrix} 0 & \frac{1}{d_{12}} \\ \frac{1}{d_{21}} & 0 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{3.17}$$

Remembering the matrix multiplication rules, we can write this out to get the equations above. In full matrix notation, we have the following.

$$2\beta p + \gamma \mathbf{D} p = 0 \tag{3.18}$$

where $\mathbf{D}$ is a matrix representing the distances between the stores with zeros on the diagonal and the inverse distance between the stores in each cell, and p is a vector of prices for the sandwiches.

The full empirical equilibrium can be written in matrix notation.

$$2\beta p - \beta c + \mathbf{X}\alpha + \gamma \mathbf{D} p + \epsilon = 0 \tag{3.19}$$

where p is the vector of prices for the sandwich, c is a vector of marginal costs for each outlet, $\mathbf{X}$ is the matrix of observed characteristics for each outlet, α is a vector of representing how customers value those characteristics and $\mathbf{D}$ is a full matrix with the distances between all of the outlets in Santa Clara County. You have to admit, it looks a lot nicer.

3.5.4 Estimator

Our estimation problem is to find the parameters of the model such that equilibrium prices in the model most closely matche the observed prices. To do this we will use a **method of moments** estimator. It sounds pretty fancy but it is just least squares. The idea is that because there are some unobserved characteristics that are determining the price, represented by ϵ in Equation (3.19) the equation will not precisely hold.

However we will assume that the Equation (3.19) holds on average for each outlet. We are assuming that the unobserved term is zero on average in equilibrium. Assumption 1 makes this idea formal.

Assumption 1. $\mathbb{E}(\epsilon_i | \mathbf{X_i}, p, c_i, \mathbf{D}_i) = 0$ *for all* $i \in \{1, ..., N\}$, *where* $\mathbf{D}_i$ *is the vector of distances from outlet* i *to every other outlet.*

Now we don't actually know the set of average prices for the outlets in the market. Rather we only observed a set prices for each outlet once. So we are going to assume that the analog of Assumption 1 holds in the data we observe. Rather than requiring that this equation is exactly zero, we will look for the parameters that make the average of ϵ_i squared as small as possible. We are going to minimize the sum of squares, or least squares.

$$\min_{\{\alpha,\beta,\gamma\}} \quad \frac{1}{N} \sum_{i=1}^{N} (2\beta p_i - \beta c_i + \mathbf{X}_i \alpha + \gamma \mathbf{D}_i p)^2 \tag{3.20}$$

In the actual estimation we will add one more complication. We don't observe c_i, the marginal cost at the outlets. We don't think that is a big deal because within a brand, the marginal costs will be very similar across the outlets in the data. One requirement from corporate headquarters is that each outlet use a supplier of similar quality to corporate's preferred supplier. They can't require a particular supplier be used but they can require certain standards be maintained. It is reasonable to think the cost of ingredients into the sandwich is pretty similar across all the outlets within a brand. In addition, all the outlets face similar labor markets and would pay similar wages. All this means that we assume $c_i = brand_i + \epsilon_c$.

3.5.5 Estimator in R

Now we need to translate all the math above into code so that we run the estimator on the data. The function `outlet_price_f()` maps Equation (3.20) into **R** code. It looks pretty similar but it is not quite the same. There is an extra matrix `Omega`. We will worry about this matrix in the next section. For the moment it is going to be the **identity matrix**, that is a matrix with all 1's on the diagonal and zeros every where else. Given this, the function is identical to Equation (3.20).[8] To do matrix multiplication in **R** the operation is `A%*%B` where `A` and `B` are two matrices where the number of columns of `A` is equal to the number of rows of `B`. While the operation `A*B` means that each cell of `A` is multiplied by each cell of `B`.[9]

```
> outlet_price_f = function(price, D, Omega, cost, X,
+                               alpha, beta, gamma) {
+     epsilon = beta*price +
```

[8] Do some matrix algebra with the identity matrix to convince yourself.

[9] `A` and `B` should have the same dimensions, but it is not strictly necessary but does lead to weird results if it does not hold.

```
+       gamma*D%*%price +
+       beta*(price - cost) +
+       gamma*(Omega*D)%*%(price - cost) +
+       X%*%alpha
+     sos = mean(epsilon^2)
+     return(sos)
+ }
```

The first function tries to look as close to the math as possible. The second function is a translation function. It translates from what works best as a function to optimize to what looks nicest. This translation function (`outlet_price_f_int()`) makes it clear to **R** that `X` is a matrix.[10]. It then translates a vector `par` which is used by the optimization function `optim()`, into our parameters, `alpha`, `beta`, and `gamma`. We force `beta` to be negative and `gamma` to be positive. We do this using the `exp()` function, which takes a number and makes it greater than 0.[11] As a rule optimization algorithms are not good at subtlety. Better to let it choose what ever parameter values it likes and then translate its choice into what ever restrictions you want to place on the parameter.

```
> outlet_price_f_int = function(par, price, D, Omega, BK, X) {
+   X = as.matrix(X)
+   J = dim(X)[2]
+   alpha = par[1:J]
+   beta = -exp(par[J+1])
+   gamma = exp(par[J+2])
+   cost = cbind(1,BK)%*%par[c(J+3,J+4)]
+   return(outlet_price_f(price, D, Omega, cost,
+                         X, alpha, beta, gamma))
+ }
```

3.5.6 Distances

In order to use the distance between outlets in our estimation we need to calculate distance between outlets. The code uses the `dist()` function to calculate the Euclidean distance between two points ("as the crow flies distance"). It is used in the loop to calculate the distance between all the outlets in the data set. The code runs a loop in **R** using `for()`. Whenever running a loop in **R** it is good practice to create an empty object, here it is a matrix `dist_mat`, that gets filled in during the loop.[12]

[10] **R** sometimes gets confused about what is a matrix and what isn't, so you need to repeat yourself a bit.

[11] It is the exponential function.

[12] This little trick speeds up the **R** substantially.

```
> dist = function(lat0, lon0, lat1, lon1) {
+   return(sqrt((lat0 - lat1)^2 + (lon0 - lon1)^2))
+ }
> J = dim(data)[1]
> dist_mat = matrix(NA, J, J)
> for(i in 1:J) {
+   for(j in 1:J) {
+     dist_mat[i,j] = dist(data$lat[i], data$lon[i],
+                          data$lat[j], data$lon[j])
+   }
+ }
```

Now we can create our matrix `D` by finding the inverse of the distance for each outlet combination and setting the values on the diagonal to 0.[13] Lastly we need to set the stores that have the same location to have a distance that is very close.[14]

```
> D = 1/dist_mat
> diag(D) = 0
> D[is.infinite(D)] = 1/0.006
```

3.5.7 Ownership

Almost there. There is one more thing to discuss before estimating prices and that is ownership. In the pricing model, we made the simplifying assumption that each outlet is priced independently. That is not true in the data. In the data, there are people that own multiple franchises. In this case, the owner is going to price their sandwiches accounting for the fact that they own other outlets.

Assume that one person owns two outlets and must choose the optimal price for each. Remember the demand function for each outlet depends on the prices of both outlets.

$$\max_{\{p_1,p_2\}} \quad q_1(p_1,p_2)(p_1-c_1)+q_2(p_2,p_1)(p_2-c_2) \tag{3.21}$$

For this case the first order conditions are as follows.

$$\begin{array}{ll} p_1: & q_1(p_1,p_2)+\frac{dq_1}{dp_1}(p_1-c_1)+\frac{dq_2}{dp_1}(p_2-c_2)=0 \\ p_2: & q_2(p_2,p_1)+\frac{dq_1}{dp_2}(p_1-c_1)+\frac{dq_2}{dp_1}(p_2-c_2)=0 \end{array} \tag{3.22}$$

So this time if they raise the price of the sandwich in outlet 1, they lose sales ($\frac{dq_1}{dp_1}$) but they gain back profits in proportion of the people that switch to

[13] In **R** the operation `1/D` calculates the inverse at each cell of the matrix, not the matrix inverse. To calculate the matrix inverse use `solve()`.

[14] 0.006 is the minimum distance of stores that don't have the same location.

the other outlet ($\frac{dq_2}{dp_1}$). Because the cost to the outlet of increasing price is mitigated by the recapturing of customers, prices will be higher.

In our problem, we have the following.

$$\begin{aligned} \frac{dq_1}{dp_1} &= \beta \\ \frac{dq_1}{dp_2} &= \frac{\gamma}{d_{12}} \end{aligned} \tag{3.23}$$

The omega notation allows us to keep track of the various ownership possibilities. Each outlet is assumed to own itself ($\omega_{11} = \omega_{22} = 1$), but there is a possibility that both outlets are owned by the same person ($\omega_{12} = \omega_{21} = 1$) or by different people ($\omega_{12} = \omega_{21} = 0$).

$$\begin{aligned} \beta p_1 + \gamma \frac{p_2}{d_{12}} + \omega_{11}\beta(p_1 - c_1) + \omega_{12}\gamma \frac{p_2 - c_2}{d_{12}} &= 0 \\ \beta p_2 + \gamma \frac{p_1}{d_{21}} + \omega_{22}\beta(p_2 - c_2) + \omega_{21}\gamma \frac{p_1 - c_1}{d_{21}} &= 0 \end{aligned} \tag{3.24}$$

We can write it in the following form.

$$2\beta p - \beta c + \gamma \mathbf{D} p + \gamma \mathbf{\Omega} \cdot \mathbf{D}(p - c) = 0 \tag{3.25}$$

where $\mathbf{\Omega} \cdot \mathbf{D}$ means cell by cell multiplication and $\mathbf{\Omega}$ is the ownership matrix.

$$\mathbf{\Omega} = \begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{bmatrix} \tag{3.26}$$

Remember that the diagonals of the matrix $\mathbf{D}$ are zeros.

3.5.8 Ownership of Outlets

In the data, there is an index for the different owners. If the ownership variable is 0 the outlet is independent and individually owned. If the variable is 1, then it is corporate owned (and has been dropped). If the variable is greater than 1 then the number is used to identify which outlets have the same owner. For example, outlets with ownership set to 6 have the same owner. The code below uses `which()` to find the column where the ownership code is the same. Each outlet is assumed to be owned by itself so the diagonal of the matrix is 1s.

```
> Omega = matrix(0, J, J)
> for(j in 1:J) {
+   if(data$Ownership[j] > 1) {
+     Omega[j,which(data$Ownership == data$Ownership[j])] = 1
+   }
+ }
> diag(Omega) = 1
```

3.5.9 Estimation (Part 2)

The final piece of the puzzle is to use the `optim()` function. This is a basic optimization algorithm used in **R**. It is used because it was made freely available by John C. Nash, no relation to John F. Nash, of Nash equilibrium fame. To use this function we give it a set of starting values (`init`). Here these are set to zero for `alpha` and very small numbers for `beta` and `gamma`. When the first function is called, it hits these values with `exp()` and $\exp(\log(a)) = a$. That is, log is the inverse of exp. The `diag()` function creates a matrix with values on the diagonal and zeros every where else. The function `dim()` finds the dimensions of a matrix, where the first element is the number of rows. We give `optim()` the initial values of the parameters, the function to optimize, then any extra information that the function will need. In this case the values for `D`, `price`, `Omega`, `BK`, and `X`. Lastly we can control the maximum number of iterations it will use before stopping. The default is small, so you may want to make this big using `maxit`. The parameter `trace` gives you a print out of what the function is doing when it is set to 1.

```
> init = c(rep(0, 5), log(0.001), log(0.001), c(0,0))
> a1 = optim(par = init,
+             fn = outlet_price_f_int,
+             D = D,
+             price = exp(data$LPrice1),
+             Omega = Omega,
+             BK = data$BK,
+             X = cbind(data$Playland,
+                       data$Mall,
+                       data$DriveThru,
+                       data$Race,
+                       data$BK),
+             control=list(trace = 0,maxit = 10000000))
```

```
> # alpha
> a1$par[1:5]
 [1] -0.005985475 -0.017518129  0.005481921 -0.006670482
     -0.005309910
> # beta
> -exp(a1$par[6])
 [1] -2.308111e-05
> # gamma
> exp(a1$par[7])
 [1] 2.583274e-08
> # cost
> a1$par[8:9]
 [1] 4.717501 8.238029
```

Standard errors have not been calculated for these estimates. How would you do that?

3.5.10 Goodness of Fit

It is standard in structural econometrics to present some type of goodness of fit analysis. Here we can use the parameter values to generate the predicted price. We can then compare the prediction to the observed prices.

To do the analysis we need to create a new function. This function finds the new set of prices given the parameter values we found above.

```
> outlet_price_f_int2 = function(par, D, Omega,
+                                cost, X, alpha, beta, gamma) {
+   X = as.matrix(X)
+   price = exp(par)
+   return(outlet_price_f(price, D, Omega,
+                         cost, X, alpha, beta, gamma))
+ }
```

```
> init = data$LPrice1
> b1 = optim(par = init,
+            fn = outlet_price_f_int2,
+            D = D,
+            Omega = Omega,
+            cost = cbind(1,data$BK)%*%a1$par[8:9],
+            X = cbind(data$Playland,
+                      data$Mall,
+                      data$DriveThru,
+                      data$Race,
+                      data$BK),
+            alpha = a1$par[1:5],
+            beta = -exp(a1$par[6]),
+            gamma = exp(a1$par[7]))
```

Figure 3.3 presents the goodness of fit of the simulation using the estimated parameters. In this exercise, we use the estimated parameters and then simulate the prices that satisfy the Nash equilibrium condition. The fitted curve is a little higher than the actual prices.

3.5.11 A Merger of McDonald's Outlets

It is not clear if the FTC has ever analyzed the impact of concentration among McDonald's franchisee owners. The economic theory is not different from analyzing the impact of a merger between hospitals or supermarkets.

Our experiment is for the independent McDonald's outlets to be purchased by the same person.

```
> data$Ownership2 = ifelse(data$BK == 2,
+                          10,
+                          data$Ownership)
> J = length(data$Ownership2)
> Omega2 = matrix(0, J, J)
> for(j in 1:J) {
+   if(data$Ownership2[j] > 1) {
+     Omega2[j, which(data$Ownership2 == data$Ownership2[j])] = 1
+   }
+ }
> diag(Omega2) = 1
```

We will assume the merger changes the ownership of the outlets but not the existence of the outlets. In our math and our code, this change is captured using `Omega` matrix. The code above finds all the McDonald's outlets that are independently owned and sets them to all have the same ownership.

The new firm's pricing decision for any outlet accounts for how that outlet's price affects demand at their other outlets. When one McDonald's increases the price for their Big Mac, some customers will switch to another outlet. After the merger many of these customers switch to outlets that are owned by the same firm. The merger reduces the loss in profits when prices are increased. The merger will lead to higher prices for Big Macs.

```
> init = b1$par
> c1 = optim(par = init,
+            fn = outlet_price_f_int2,
+            D = D,
+            Omega = Omega2,
+            cost = cbind(1,data$BK == 1)%*%a1$par[c(8,9)],
+            X = cbind(data$Playland,
+                      data$Mall,
+                      data$DriveThru,
+                      data$Race,
+                      data$BK),
+            alpha = a1$par[1:5],
+            beta = -exp(a1$par[6]),
+            gamma = exp(a1$par[7]))
```

The code below generates the `ggplot()` of the density of prices for the actual prices, the predicted prices from the estimated model and the predicted prices from the model of the merger of independent McDonald's outlets.

```
> ggplot_dens_outlets = data.frame(
+   Price = exp(data$LPrice1),
+   sim = exp(b1$par),
+   merger = exp(c1$par)
+ ) |>
+   ggplot(aes(x = Price)) +
+   geom_density(aes(y = ..scaled..), alpha = 0.5) +
+   geom_density(aes(x = sim, y = ..scaled..),
+                alpha = 0.5,
+                linetype = "dashed") +
+   geom_density(aes(x = merger, y = ..scaled..),
+                alpha = 0.5,
+                linetype = "dotted") +
+   geom_text(aes(x = 50, y = 1, label = "Actual")) +
+   geom_text(aes(x = 75, y = 0.3, label = "Sim")) +
+   geom_text(aes(x = 140, y = 0.4, label = "Merger Sim")) +
+   labs(x = "Price (normalized)",
+        y = "",
+        title = "") +
+   ## no numbers on y axis
+   scale_y_continuous(breaks = NULL) +
+   theme_minimal()


> ggplot_dens_outlets
```

Figure 3.3 shows both the goodness of fit of the model and the simulated impact on price of a merger between all the independent McDonald's outlets. The roll-up of independent McDonald's outlets would lead to a substantial increase in the prices of outlets in the market. Not only the McDonald's outlets but also the Burger King outlets that they compete with.

3.6 Discussion and Further Reading

Using game theory to analyze oligopoly models of competition actually predates game theory. In fact, all three models presented in this chapter predate Nash's analysis of game theory, even though all rely on the equilibrium concept.

Hotelling's 1929 paper, *Stability in Competition*, provides much of the intuition for the way many industrial organization economists think about competition for retail products (Hotelling, 1929). In the model, it is the "distance" between products that matters for competition. Not so much the exact number of competitors but how close the competitors are to each other in the minds of consumers.

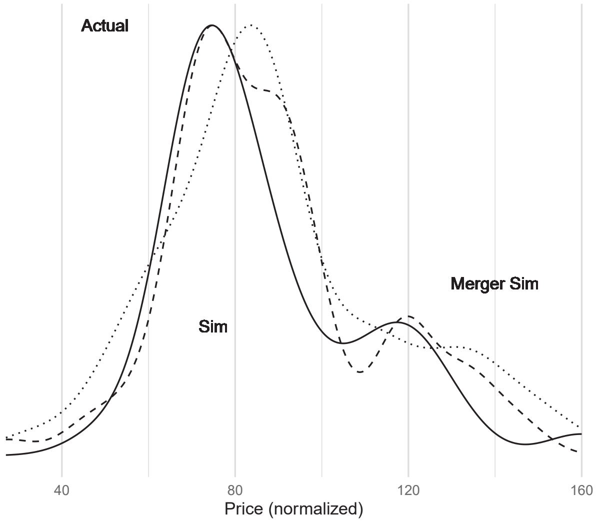


FIGURE 3.3
Plot of density of actual prices (solid line), predicted prices (dashed line) from the estimating model and simulated prices from the model of the merger of independent McDonald's outlets (dotted line). The dotted line is shifted up from the actual and simulated price distributions.

Somewhat confusingly we refer to a general model of differentiated price competition as a Bertrand model.[15] The classic paper taking this model to the data is Berry et al. (1995).

The subtlety of the Hotelling model didn't fit well into how US antitrust conducted merger review. The standard merger screen is a measure called the Herfindahl-Hirschman index (HHI) and the change the index caused by the merger. It is calculated by determining which firms are in the market, calculating each firm's share, squaring them and adding them up. While HHI is not a bad approximation of competition in homogeneous goods markets modeled by Cournot, it doesn't make a lot of sense for differentiated goods markets modeled by Hotelling. In the simple Hotelling line, the extent of competition can vary substantially without any change in the HHI.

The 2010 Merger Guidelines from the Department of Justice and the Federal Trade Commission, made an adjustment suggesting that a different screen may be better for differentiated goods mergers. The Upward Price Pressure screen measures how close two firms are by how many customers are diverted for a price increase.[16]

---

[15] Chapter 7 works through this model.

[16] `https://www.justice.gov/atr/horizontal-merger-guidelines-08192010` accessed on 11/21/23.

4

Empirical Entry Games

4.1 Introduction

This chapter revisits the entry game presented in Chapter 2. In this game, we have two firms considering whether or not to enter a market. The issue is that it is costly to enter and both firms would prefer to have the market to themselves. In equilibrium, we will tend to have one firm in the market, but the game theory may not predict which one.

The chapter applies an entry game to the question of which markets the mega bookstores, Borders and Barnes & Noble entered in the 1990s. The 90s saw massive changes in book retailing. Changes in distribution technology made it profitable for big box book stores to enter the market with a huge range of titles as well as other products such as music, games, and even coffee. The chains, Borders and Barnes & Noble led the charge by purchasing smaller chains and entering green fields sites. Which markets did these firms enter? What would have happened if they had been allowed to merge?

The chapter introduces an empirical model of entry based on the decision making model of Dan McFadden. The empirical model is extended to allow the firms to make decisions that are dependent on each other, a game. The chapter analyzes the effect of a merger, if that merger would have occurred in the early 90s.

4.2 Empirical Model

The section begins with a description of a standard choice model. This model was originally developed by Dan McFadden in the early 1970s to analyze consumer choice problems. McFadden was interested in who would use the new subway system that had been built in San Francisco, the Bay Area Rapid Transit (BART) system. Here we are adapting the model to analyze the choice of which market the firm will enter. The model is generalized to allow unobserved characteristics of the market to be correlated across firms. It is generalized again to allow firms to make choices that are dependent upon each other.

DOI: 10.1201/9781003351603-4

4.2.1 Single Firm Entry Model

By the year 2000, Barnes & Noble had bookstores in 283 counties in the United States, but this is out of over 3,000 counties. Placing these large stores in a location is not cheap, particularly if they involve new construction. So which counties dd Barnes & Noble choose to locate?

Assume that the firm's latent profits from locating in a particular county are as follows.

$$\pi_{1i} = \mathbf{X}_i'\beta_1 + \xi_{1i} \tag{4.1}$$

where $\mathbf{X}_i$ are market characteristics such as population size, β_1 determines how these characteristics are mapped into profits for the firm and ξ_{1i} is unobserved characteristics of the market and can be thought of as representing entry costs for the firm. These unobserved characteristics are unobserved by the econometrician (that's you) but they are observed by the firm itself.

If the data include information on the market such as population size (*pop*) and median income (*income*), then the decision to enter can be written as follows.

$$\pi_{1i} = \beta_{10} + \beta_{11} pop_i + \beta_{12} income_i + \xi_{1i} \tag{4.2}$$

For Firm 1, their profits in market i depend on pop_i and $income_i$ as determined by the parameters β_{11} and β_{12}, respectively.

In matrix notation we have that the vector of parameters is written as follows.

$$\beta_1 = \begin{bmatrix} \beta_{10} \\ \beta_{11} \\ \beta_{12} \end{bmatrix} \tag{4.3}$$

Similarly, the matrix of observed characteristics for N markets as follows.

$$\mathbf{X} = \begin{bmatrix} 1 & pop_1 & income_1 \\ 1 & pop_2 & income_2 \\ \dots & & \\ 1 & pop_N & income_N \end{bmatrix} \tag{4.4}$$

The first column is just 1's. In the matrix algebra, this column is multiplied by the β_{10} parameter to give a constant across all the markets. We use the notation $\mathbf{X}_i'$ to emphasize that we are using the ith row of the matrix and that is being multiplied by the vector of parameters β_1.

Barnes & Noble enter the market if and only if the following inequality holds.

$$\begin{aligned} \mathbf{X}_i'\beta_1 + \xi_{1i} & > 0 \text{ or} \\ \xi_{1i} > -\mathbf{X}_i'\beta_1 & \end{aligned} \tag{4.5}$$

So if the unobserved entry costs are low enough, then Barnes & Noble will enter the market.

We expect firm profits to determined by various factors affecting demand for books, costs associated with selling books, and fixed costs associated with the location of the store. If we assume that the unobserved costs of entry are distributed standard normal, $\xi_i \sim \mathcal{N}(0,1)$, then the probability of entry is $\Phi(-\mathbf{X}_i'\beta_1)$, where $\Phi()$ is the cumulative distribution of the standard normal. This is the **probit** model introduced in Chapter 1. We can estimate the parameters β_1 using the `glm()` procedure we used in Chapter 1.

$$\max_{\beta_1} \sum_{i=1}^{N} y_{1i} \log(\Phi(-\mathbf{X}_i'\beta_1)) + (1 - y_{1i}) \log((1 - \Phi(-\mathbf{X}_i'\beta_1)) \tag{4.6}$$

To estimate the model we can find the β_1 that maximizes the likelihood of the data, where y_{1i} is the observed entry decision for each county.[1]

4.2.2 Multiple Firm-Independent Entry Model

Taking baby steps to our full model, consider a model where we have two firms entering a market but the two firms are making decisions independently of each other. This could be a model of entry of Barnes & Noble and Best Buy. Both are big box stores whose decision to enter a market will be based on similar things, both observed and unobserved by the econometrician. However, with one focused on books and the other focused on electronics, it is unlikely that their decisions to enter a particular market will be dependent on each other.

The two firms will enter market i if and only if the following inequalities hold.

$$\begin{aligned} \mathbf{X}_i'\beta_1 + \xi_{1i} > 0 \\ \mathbf{X}_i'\beta_2 + \xi_{2i} > 0 \end{aligned} \tag{4.7}$$

Again $\mathbf{X}_i$ represents observed characteristics of the market such as population size. These characteristics are mapped into each firm's profit function by the parameter vectors β_1 and β_2.

So far this doesn't seem to be different from the model in the previous section. The difference is that we can allow the unobserved characteristics of the market to be correlated across the two firms. Assume that $\{\xi_{1i}, \xi_{2i}\} \sim \mathcal{N}(\mu, \mathbf{\Sigma})$ where $\mu = \{0, 0\}$ and

$$\mathbf{\Sigma} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \tag{4.8}$$

In words, we allow the unobserved characteristics of the markets to be distributed standard bivariate normal. The parameter $\rho \in [-1, 1]$ represents the correlation across the two firms. This parameter is likely to depend on how similar the two firms are in terms of their customer base and costs of setting up the store. A fast food restaurant and big box book store may have unobserved characteristics that are not vary correlated $\rho = 0$, while two big box stores may have a highly positive correlation (a high ρ).

[1] We log the probabilities so that we don't run into problems because the numbers are too small for the computer to represent.

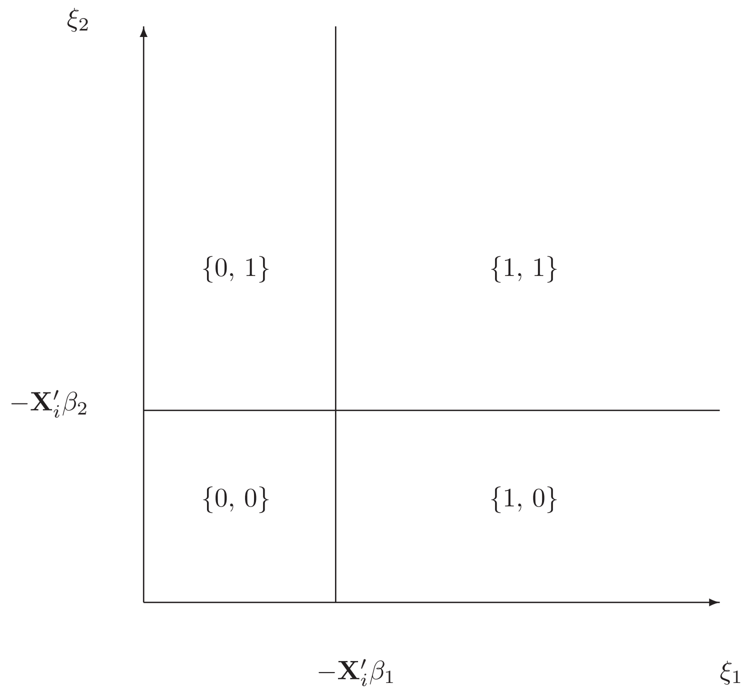


FIGURE 4.1
Empirical implications of the two firm independent entry model. Firm 1's entry is denoted in the first position of the brackets. If both firms have unobserved characteristics that are low then neither will enter $\{0, 0\}$, if Firm 1's unobserved characteristic is high then Firm 1 will enter $\{1, 0\}$, if Firm 2's unobserved characteristic is also high the Firm 2 will also enter $\{1, 1\}$.

Let 1 denote entry and 0 denote the choice not to enter. In this model we observe four cases, neither firm enters $\{0, 0\}$, Firm 1 enters but Firm 2 does not $\{1, 0\}$, Firm 1 doesn't enter but Firm 2 does $\{0, 1\}$, and both firms enter $\{1, 1\}$.

Figure 4.1 presents the four cases given the distribution of the unobserved characteristics of the markets for the two firms. The pattern of entry across markets will tell us if the unobserved characteristics are correlated across firms. If we see lots of cases where both firms do the same thing, both enter the market or neither enters the market, then that is consistent with positively correlated unobserved characteristics. If we see lots of cases where there is just one firm in the market but which is pretty evenly distributed, then that is consistent with the unobserved characteristics being uncorrelated or negatively correlated across the two firms.

Again we want to find the β_1, β_2 and ρ that maximize the likelihood of the observed data which is represented by y_1 and y_2.

4.2.3 Entry Game

After all of that set up, we can write down our empirical entry game. Like in the previous section we have two firms considering whether or not they should enter market i. They will enter if the following inequalities hold.

$$\begin{aligned} \mathbf{X}_i'\beta_1 - D_{2i}\alpha_1 + \xi_{1i} > 0 \\ \mathbf{X}_i'\beta_2 - D_{1i}\alpha_2 + \xi_{2i} > 0 \end{aligned} \tag{4.9}$$

where $D_{1i} \in \{0, 1\}$ and $D_{2i} \in \{0, 1\}$. There are a couple of differences between these inequalities and the inequalities in Equation (4.7). The first is that there are two extra parameters α_1 and α_2. These represent the reduced profits associated with competing with the other firm. This only occurs if the other firm enters the market. The profits for Firm 1 are lower if $D_{2i} = 1$ by the amount α_1, where $D_{2i} = 1$ if and only if the second inequality holds. The profits for Firm 2 are lower if $D_{1i} = 1$ which only occurs if the first inequality holds.

Now you may be able to start seeing the issue. The first inequality depends on the outcome of the second inequality which depends on the outcome of the first inequality! To see what is going on consider a version of the figure presented above.

Figure 4.2 represents the game. If the unobserved benefits for both firms are high (fixed cost of entry is low), then both firms will enter $\{1, 1\}$, if they are low, then neither will enter $\{0, 0\}$. If they are very high for Firm 1 and very low for Firm 2, then Firm 1 will enter $\{1, 0\}$. Similarly the other way $\{0, 1\}$. There is also the intermediate outcome where the model does not make a clear prediction on what will happen. It could be that Firm 1 enters while Firm 2 does not, or it could be the other way around. The model clearly predicts that one firm will enter, it just does not predict which.

4.3 Empirical Analysis: Bookstore Entry with R

The section presents the code for estimating the empirical models presented above. To show the progression from the original entry model to the entry game, the section presents code for all three models. The first model is just a probit, so you could use the `glm()` function baked into **R**. However, including

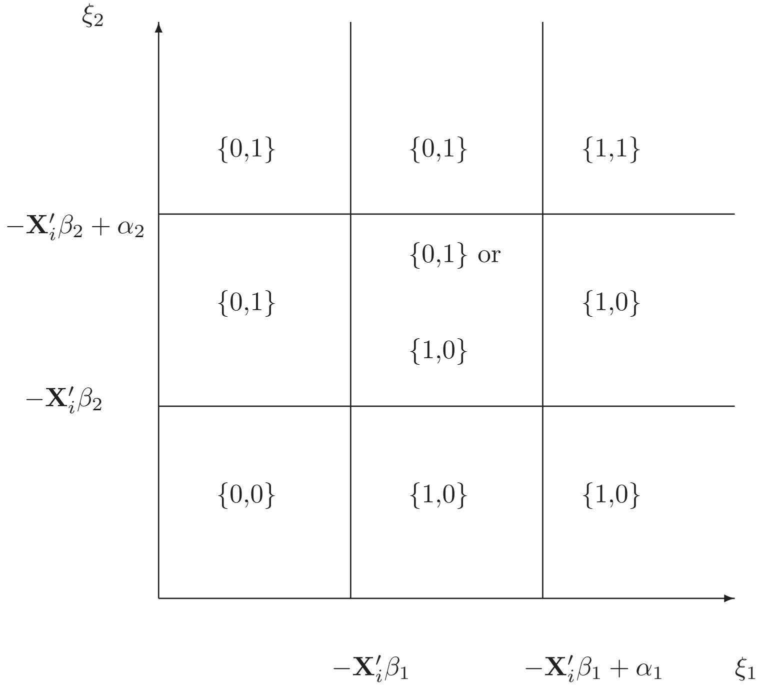


FIGURE 4.2
Empirical implications of the entry game. Firm 1's entry is denoted in the first position of the brackets. The region in the middle square has an in determinant outcome. One firm will enter, but it is not clear which.

the code for this case makes it easier to see how the more complicated models work.

4.3.1 Single Firm Entry Model in R

The estimation algorithm is **maximum likelihood**. In this method, we find the parameter values that maximize the likelihood that the data we observe was generated by a particular model. We will need a function that calculates the probability that the observed data occurred given a particular parameter value. There are a couple of different ways of doing that in **R**. In this section, we will calculate the probabilities **numerically**. That is the section presents a method for approximating the probability by having the computer do a lot of calculations. This method is not as efficient as using a built-in `C`-based function, but it is easier to see what is going on, particularly as the problem gets more complicated.

```
> set.seed(123456789)
> K = 10000
> Z_1 = rnorm(K)
```

This code generates `K = 10000` pseudo-random draws from a standard normal distribution. The function `rnorm()` is the **R** function for drawing from a standard normal. These are **global** variables, meaning that they are available to any function we write. The function `set.seed()` is used to make sure that the results can be exactly replicated.

```
> f_entry = function(X, beta1) {
+   X1 = as.matrix(cbind(1, X))
+   N = dim(X1)[1]
+   xi = Z_1
+   Xb = X1%*%beta1
+   p = rep(0, N)
+   for(k in 1:K) {
+     pi_k = Xb + xi[k]
+     p = p + (pi_k > 0)
+   }
+   return(p/K)
+ }
> f_loglik = function(X, y, beta) {
+   epsilon = 1e-5
+   p = f_entry(X, beta)
+   return(-mean((y == 1)*log(p + epsilon) +
+                  (y == 0)*log(1 - p + epsilon)))
+ }
```

The function `f_entry()` takes in the matrix of data `X` (without the column of 1's) and the vector of parameters `beta`. The matrix `X` consists of columns of data stating observed characteristics of each market, such as population size and median income.

The probability of entering market i is determined by the probability that profits will be positive given a large number of possible values for the unobserved characteristic `xi`.

The function `f_loglik` calculates the probability that the model is true given the observed data. This code is equivalent to Equation (4.6). This function takes a vector of outcomes, `y`. The vector `y` states whether or not the firm entered each of the markets in the data where 1 denotes entry and 0 denotes that the firm did not enter the market. The function transforms the probability into logs so that we don't run into problems where the numbers we are calculating are smaller than the smallest number the computer can handle. Later when we apply this function to an actual problem, we will benefit from the fact that the optimal value for `beta` is the same for the log-transformed function as it is for the original function. Notice that there is a negative sign in front of the `mean()` function. This is there because the optimize algorithm used defaults to find the minimum so the minimum of the negative log-likelihood is the maximum log-likelihood. The value `epsilon` is a small number designed to make sure that the computer doesn't crash if it tries to calculate the log of zero.

4.3.2 Multiple Firm-Independent Entry in R

```
> Z_2 = rnorm(K)
```

This time we need a distribution with two dimensions. So the first step is to create another large set of random numbers drawn from a standard normal function.

```
> f_2entry = function(X, beta_1, beta_2, rho) {
+   N = dim(X)[1]
+   xi_1 = Z_1
+   xi_2 = Z_2*sqrt(1 - rho^2) + rho*Z_1
+   Xb_1 = X%*%beta_1
+   Xb_2 = X%*%beta_2
+   p_00 = p_01 = p_11 = rep(0, N)
+   for(k in 1:K) {
+     pi_1k = Xb_1 + xi_1[k]
+     pi_2k = Xb_2 + xi_2[k]
+     p_00 = p_00 + (pi_1k < 0 & pi_2k < 0)
+     p_01 = p_01 + (pi_1k < 0 & pi_2k > 0)
+     p_11 = p_11 + (pi_1k > 0 & pi_2k > 0)
+   }
+   return(list(p_00 = p_00/K,
+               p_01 = p_01/K,
+               p_11 = p_11/K))
+ }
> f_loglik_2 = function(X, y, beta_1, beta_2, rho) {
+   epsilon = 1e-10
+   Lik = f_2entry(X, beta_1, beta_2, rho)
+   return((y[,1] == 0 & y[,2] == 0)*log(Lik$p_00 + epsilon) +
+          (y[,1] == 0 & y[,2] == 1)*log(Lik$p_01 + epsilon) +
+          (y[,1] == 1 & y[,2] == 1)*log(Lik$p_11 + epsilon) +
+          (y[,1] == 1 & y[,2] == 0)*log(1 -
+                                          Lik$p_00 -
+                                          Lik$p_01 -
+                                          Lik$p_11 +
+                                          epsilon))
+ }
> f_loglik_2_int = function(par, X, y) {
```

```
+   X = as.matrix(cbind(1, X))
+   J = dim(X)[2]
+   beta_1 = par[1:J]
+   beta_2 = par[(J+1):(2*J)]
+   rho = -1 + 2*exp(par[2*J+1])/(1 + exp(par[2*J+1]))
+   return(-mean(f_loglik_2(X, y, beta_1, beta_2, rho)))
+ }
```

Just by counting lines of code, we see that things are a lot more complicated when we have two firms whose decisions are independent but correlated. The function `f_2entry()` has parameters for Firm 1 (`beta_1`) and Firm 2 (`beta_2`) and the correlation between the unobserved term (`rho`). The unobserved term `xi_2` is a function of both `Z_1` and `Z_2`. The higher `rho` the more it weights `Z_1`. The function determines the probability of observing three cases neither enter, only Firm 1 enters, and both enter. The fourth case is calculated as the residual because probabilities must add to 1.

The code now includes an additional function `f_loglik_2_int()`. This is an intermediate function designed to be used by the **R**'s optimization algorithm `optim()`. In this case, the parameter `rho` is restricted to be between -1 and 1. It is good coding practice to allow the optimization algorithm to choose what ever values it likes, but then transform those values into the restricted set required by the model. To do this the code uses `exp(x)/(1 + exp(x))`, also known as the **softmax** function. This function takes any value of `x` and turns it into a number that lies between 0 and 1.

4.3.3 Entry Game in R

While the entry game is quite a bit more complicated than the previous model, the estimator is not that different. There are two additional parameters `alpha_1` and `alpha_2` but everything else is pretty much the same. The big difference is that the model cannot distinguish two cases in the data. The model predicts when one firm will enter the market, but it does not predict which firm that will be. Therefore, we need to combine those two cases in order to estimate our model.

```
> f_entry_game = function(X, beta_1, beta_2, alpha_1, alpha_2
+                         , rho) {
+   N = dim(X)[1]
+   xi_1 = Z_1
+   xi_2 = Z_2*sqrt(1 - rho^2) + rho*Z_1
+   Xb_1 = X%*%beta_1
+   Xb_2 = X%*%beta_2
+   p_00 = p_11 = rep(0, N)
+   for(k in 1:K) {
+     pi_1k = Xb_1 + xi_1[k]
+     pi_2k = Xb_2 + xi_2[k]
```

```
+     p_00 = p_00 + (pi_1k < 0 & pi_2k < 0)
+     p_11 = p_11 +
+       (pi_1k - alpha_1 > 0 & pi_2k - alpha_2 > 0)
+   }
+   return(list(p_00 = p_00/K,
+               p_11 = p_11/K))
+ }
```

The function `f_loglik_game_int()` is pretty similar to the equivalent function in the previous section. The difference is that there are two additional parameters. We will restrict these two parameters to be positive. We will impose the result that an increase in competition lowers profits (or does nothing).

```
> f_loglik_game = function(X, y, beta_1, beta_2, alpha_1,
+                          alpha_2, rho) {
+   epsilon = 1e-10
+   Lik = f_entry_game(X, beta_1, beta_2, alpha_1, alpha_2, rho)
+   return((y[,1] == 0 & y[,2] == 0)*log(Lik$p_00 + epsilon) +
+          (y[,1] == 1 & y[,2] == 1)*log(Lik$p_11 + epsilon) +
+          ((y[,1] == 1 & y[,2] == 0) +
+             (y[,1] == 0 & y[,2] == 1))*log(1 -
+                                          Lik$p_00 -
+                                          Lik$p_11 +
+                                          epsilon))
+ }
> f_loglik_game_int = function(par, X, y) {
+   X = as.matrix(cbind(1, X))
+   J = dim(X)[2]
+   beta_1 = par[1:J]
+   beta_2 = par[(J+1):(2*J)]
+   alpha_1 = exp(par[2*J+1])
+   alpha_2 = exp(par[2*J+2])
+   rho = -1 + 2*exp(par[2*J+3])/(1 + exp(par[2*J+3]))
+   return(-mean(f_loglik_game(X, y, beta_1, beta_2, alpha_1,
+                              alpha_2, rho)))
+ }
```

4.4 Empirical Analysis: Bookstore Entry using R

Barnes & Noble dates itself to the 1800s, but in the early 1990s, the firm developed the super bookstore. The big box of bookstores. The objective was to carry a huge range of books, music, games, and even food and coffee. It revolutionized book retail in the United States.

Where did Barnes & Noble choose to enter? What determined those locations?

4.4.1 Data

The data used here is from Adams and Basker (2025). The authors combine information from publicly available census data and published directories of retail bookstores.

TABLE 4.1
Mean County Characteristics by Presence of Barnes & Noble and Borders.

	Population	Income	College Share (%)	Bookstores (#)
None	37,857	33,513	15	2
Only Barnes & Noble	275,435	43,210	27	14
Only Borders	548,218	49,118	31	40
Both	272,191	46,476	26	12

Notes: Presence of Barnes & Noble and Borders referes to the year 2000. There are 2,792 counties with neither chain, 155 counties with only Barnes & Noble, 36 counties with only Borders, and 128 counties with both chains. Population, income, and college share use data from 2000. Income refers to median county-level household income in 2000. College share is the share of population aged 25 and older with a college degree in 2000. Bookstores is the total number of bookstores in the county in 1990.

Table 4.1 shows the difference between counties with and without Borders and Barnes & Noble stores. Unsurprisingly, the firms entered counties with larger populations, richer counties, counties with higher educated population, and counties with more bookstores in 1990.

4.4.2 Estimation of Single Firm Entry

The model presented in Equation (4.5) can be estimated by combining data on the location of Barnes & Noble stores with various economic and demographic information at the county level. In addition to these data, we have information on the number of book stores in the county in 1990, which is generally earlier than the Barnes & Noble super bookstores came into existence. We can similarly model the entry of Borders.

In this analysis, there is no game. Barnes & Noble and Borders are assumed to make their entry decisions optimally, but these decisions are completely independent of each other. To estimate the parameters of the model, we can either use a maximum-likelihood estimator and the functions `f_loglik()` and `f_enter()` or we can use the `glm()` procedure introduced in Chapter 1.

```
> file = paste0(dir, "book_2000.csv")
> dt = fread(file)

> glm1 = glm(enter ~ log_pop_2000 + med_income + college +
+              stores_1990,
+             data = dt,
+             family = binomial(link = "probit"))
```

This function is used to estimate cases where the outcome is binary, enter or did not enter, and where the unobserved characteristic of the market is distributed as a normal distribution with mean of 0 and standard deviation of 1. This is a **probit model**.

```
> init = glm1$coefficients
> data = data.frame(y = dt$enter,
+                   pop = dt$log_pop_2000,
+                   income = dt$med_income,
+                   college = dt$college,
+                   stores_1990 = dt$stores_1990)
> data = na.omit(data)
> res1 = bs(init, f_loglik,
+           y = data$y,
+           X = cbind(data$pop,
+                     data$income,
+                     data$college,
+                     data$stores_1990))
```

The `glm1` results are used as the initial starting value for the maximum likelihood estimator using `f_loglik()`. The code then creates the data to be used by the estimator, including information on which stores Barnes & Noble entered, population size of the county, percentage of college graduates in the county and the number of book stores in the county in 1990. The object `res1` stores the results. This uses a function called `bs()` which creates pseudo samples from the data set and uses the `optim()` function to determine the parameter values that maximize the likelihood using the `f_loglik()` function.[2]

```
> dt$enter2 = ifelse(dt$borders > 0, 1, 0)
> glm2 = glm(enter2 ~ log_pop_2000 +
+              med_income +
+              college + stores_1990, data = dt,
+             family=binomial(link="probit"))
```

The data set is adjusted to remove observations with missing values. The function `na.omit()` is used to do this.[3]

[2]The `bs()` function is available from the github page for the book.
[3]**R** uses `NA` to denote missing values.

```
> init = glm2$coefficients
> data = data.frame(y = dt$enter2,
+                   pop = dt$log_pop_2000,
+                   income = dt$med_income,
+                   college = dt$college,
+                   stores_1990 = dt$stores_1990)
> data = na.omit(data)
> res2 = bs(init, f_loglik,
+           y = data$y,
+           X = cbind(data$pop,
+                     data$income,
+                     data$college,
+                     data$stores_1990))
```

The code above similarly creates a data set for estimating the choice of market to enter for Borders stores. It estimates the probit model using the `bs()` function and stores the results in the object `res2`.

4.4.3 Estimation of Two Firm Entry Model

Table 4.2 presents the results from the different entry models discussed above. The first model assumes that the two firms make optimal decisions that are independent of each other. In addition, it assumes that the unobserved characteristics faced by the two firms in each market are independent across the two firms.

The two-firm model is similar to single firm entry model. The entry decisions of the two firms are independent of each other but the unobserved characteristics of the markets may be correlated across firms. We say that the firms are strategically independent but the markets are statistically dependent. This is a **biprobit model**. The code again creates the data set for the analysis removes missing values and uses the `bs()` function calling `f_loglik_2_int`. It stores the results in `res3`.

```
> init = c(glm1$coefficients,
+          glm2$coefficients,
+          0)
> data = data.frame(y_1 = dt$enter,
+                   y_2 = dt$enter2,
+                   pop = dt$log_pop_2000,
+                   income = dt$med_income,
+                   college = dt$college,
+                   stores_1990 = dt$stores_1990)
> data = na.omit(data)
> res3 = bs(init, f_loglik_2_int,
+           y = cbind(data$y_1,
+                     data$y_2),
```

```
+           X = cbind(data$pop,
+                     data$income,
+                     data$college,
+                     data$stores_1990))
```

4.4.4 Estimation of the Entry Game Model

For the third model, we have two firms entering the markets and again the unobserved characteristics of the two firms are correlated across markets. This time the decision to enter depends on the choice of the other firm. Decisions are now dependent. This model uses the same data as the previous model. It uses the `bs()` function with the `f_loglik_game_int()` likelihood function and saves the results in the object `res4`.

```
> init = c(glm1$coefficients,
+          glm2$coefficients,
+          0,
+          0,
+          0)
> res4 = bs(init, f_loglik_game_int,
+           y = cbind(data$y_1,
+                     data$y_2),
+           X = cbind(data$pop,
+                     data$income,
+                     data$college,
+                     data$stores_1990))
```

The table of results also presents information on the standard errors of the estimates of the parameters. These values are determined using a **bootstrap method**. This method approximates how estimates may vary when a different sample is used.[4]

Table 4.2 presents the estimated parameter values for the three models discussed above. The α parameters state that having two firms lowers profits. The store locations are positively associated with population size, education, and the number of existing bookstores. It is unclear if income has any impact. Finally, the unobserved entry costs are highly correlated across the two firms.

4.4.5 Model Fit

We can use the estimated parameters to simulate the model and compare the predicted entry decisions to the actual entry decisions. This provides one test of the model's fit. The code, not shown, simulates the model 1,000 times

[4]The bootstrap method approximates estimates from a new sample by resampling the current data and restimating the model using the new pseudo-sample.

TABLE 4.2
Results from estimates of the three models. The first set of columns labeled "Probit" are estimates assuming that the two firms are making entry decisions that are both strategically independent and statistically independent. The second set of columns labeled "BiProbit" assumes the firm entry decisions are strategically independent. It allows unobserved characteristics of the market to be correlated across firms. The two columns labeled "Game" refer to the case where the entry decisions of the two firms are both strategically and stochastically dependent. The columns labeled "SD" refer to the bootstrap standard errors.

	Probit	SD	BiProbit	SD	Game	SD
const_1	−15.06	0.22	−15.03	0.09	−15.11	0.25
Pop_1	1.08	0.02	1.09	0.01	1.07	0.01
Income_1	−1.03	0.44	−1.04	0.25	−0.76	0.48
College_1	5.68	0.66	5.51	0.34	5.65	0.55
Stores_1990_1	0.28	0.09	0.28	0.07	0.37	0.11
const_2	−11.57	0.26	−11.54	0.11	−11.37	0.20
Pop_2	0.66	0.03	0.66	0.01	0.65	0.02
Income_2	1.75	0.51	1.74	0.25	1.31	0.80
College_2	2.49	0.70	2.59	0.33	2.70	0.55
Stores_1990_2	0.52	0.08	0.49	0.06	0.79	0.11
alpha_1					0.73	0.18
alpha_2					0.70	0.12
rho			−0.08	0.10	0.47	0.10

and compares the predicted entry decisions to the actual entry decisions. The results are presented in Table 4.3.

Table 4.3 presents model fit results for the three models. There is not much difference between the three estimators, although the game theory model is slightly better at predicting the one-firm market.

4.5 Policy Analysis using R

What would happen if Barnes & Noble and Borders had merged? Here we are not going to worry about the effect on prices but on which counties the merged firm would enter. Assume that the α parameters remain the same post merger.[5] The difference the merger brings is that it allows the firms to coordinate entry. This analysis assumes that the merged firm will keep the two distinct brands.

To determine what would happen in this alternate universe, we can simulate

[5]Given that the α parameters are accounting for both diversion between stores and prices, we would expect them to be smaller but not zero post merger.

TABLE 4.3
Comparison of predictions of the three models to the actual outcomes. The higher percentage on the diagonal the better fit of the model. All models are good at predicting the high probability event, which is no entry by either firm. The non-strategic models are best at predicting when there will be two firms in two-firm markets, while the game theory model is best at predicting the one-firm market.

	None	One Firm	Two Firm
Probit: None	97.2	42.3	9.4
Probit: One	2.5	38.8	34.9
Probit: Two	0.3	18.9	55.6
BiProbit: None	97.1	41.6	9.0
BiProbit: One	2.6	40.0	35.9
BiProbit: Two	0.3	18.4	55.1
Game: None	97.1	41.0	9.0
Game: One	2.6	41.7	36.6
Game: Two	0.3	17.3	54.4

thousands of outcomes in the markets for which have data using the parameter values estimated above.

TABLE 4.4
Comparison of actual entry to simulated entry in 2000 and simulated entry under a merger.

	Actual	Sim	Merge
none	2919	2895	2895
BN or Borders	170	191	238
both	128	93	46

Table 4.4 presents summary of simulations of a merger between Borders and Barnes & Noble. The simulations tend to predict more counties with stores than we actually see, but fewer counties where both stores are present. The model predicts that the merger will lead to fewer markets with a Borders and Barnes & Noble. Not presented is the variation in the estimates, but it is pretty clear that there will be fewer markets with both stores. As some consumers prefer one or the other, people in those markets are worse off. We can think of that as a quality reduction. In addition, the reduced competition in those markets is likely to lead to higher prices.[6]

The value in the first row of columns 2 and 3 is the same. The merger changes whether or not the market is will have two firms or one, but not whether or not at least one firm will enter. The reason is that the game explicitly assumes that the two firms coordinate on entry. Chapters 5 and 8

[6]See the analysis presented in Chapter 3.

consider games where the firms cannot coordinate. In these cases, the merger may actually lead more firms to enter the market.

4.6 Discussion and Further Reading

The analysis is based on Adams and Basker (2025). The authors analyze the entry of mega bookstores in the US using publicly available census data and directories of retail bookstores. The entry game analysis is based on Bresnahan and Reiss (1991a) and Tamer (2003). Chapters 5, 6, and 8 revisit this problem under different modeling assumptions. Chapter 5 revisits the problem assuming that the outcome of the game in the data is a mixed strategy Nash equilibrium.

Game theory provides a better explanation of the entry decisions we observe, however it can also makes the analysis more challenging. Entry problems are a type of coordination game and like other coordination games they tend to have multiple equilibria. This means for a set of parameter values, a model makes multiple predictions about what we will see in the data. This makes the models a challenge to estimate. The solution presented here follows Bresnahan and Reiss (1991a) which simply combines all the equilibria into one observed outcome. Tamer (2003) presents a method for estimating a more efficient model. Many papers in empirical industrial organization make alternative assumptions that lead to unique predictions from the game. Chapter 6 discusses this option.

Borders and Barnes & Noble did not propose to merger in the 1990s or 2000s. Did they contemplate it? Were they dissuaded by the FTC's challenge of the merger between Blockbuster and Hollywood Video in 1999 and again in 2005?[7] Borders went into bankruptcy proceedings in 2011 and Barnes & Noble acquired some of its intellectual property.[8]

[7] https://www.nytimes.com/2005/03/26/business/media/blockbuster-ends-bid-for-rival.html

[8] https://www.barnesandnobleinc.com/about-bn/history/. Accessed on 2/20/23.

5

Mixed Strategies

5.1 Introduction

Chapter 2 stated that John Nash showed that the Nash equilibrium exists in a large set of games. This is a powerful result and is part of the reason why Nash equilibrium is so important. But Nash's proof relies on **mixed strategies**. Mixed strategies are a more complicated method for determining the outcome of a game. Mixed strategies are vital for the analysis of parlor games like chess or cards, and useful for the analysis of games with multiple equilibria like the **coordination game** presented in earlier chapters.

The chapter introduces mixed strategies using the classic childhood game Rock-Paper-Scissors. It analyzes penalty kicks in soccer's English Premier League. The chapter revisits the entry of big box bookstores analyzed in Chapter 4. This time it assumes players are playing a mixed strategy when there are multiple equilibria.

5.2 Zero-Sum Games

Zero-sum games are the oldest types of games studied in game theory. They are a natural way to represent parlor games like chess. In those games, there is a winner and a loser. What the winner wins, the loser loses. This section introduces some important zero-sum games and discusses how to find the equilibrium.

5.2.1 Rock-Paper-Scissors

Rock-Paper-Scissors is one of the more famous childhood games. It requires two players with at least one hand. The game provides solutions to many previously unsolvable problems - who gets the last slice of cake, who gets to ride shotgun, who gets to mow the lawn.

- Players: Player 1 and Player 2

DOI: 10.1201/9781003351603-5

- Strategies: Rock, Paper, Scissors
- Payoffs: See Table 5.1

Table 5.1 provides a normal form representation for the game. There are only three outcomes in the game, win, lose, and draw. We can represent these outcomes as payoffs, 1, −1 and 0 respectively. If both players choose the same shape, it is a draw. Then Rock beats Scissors, Paper beats Rock, and Scissors beats Paper. Notice that the numbers in each cell add to 0, this is what we mean by zero-sum.

TABLE 5.1
Normal form representation of Rock-Paper-Scissors with two players P_1 and P_2, the strategies for P_1 are on the rows and P_2's strategies are on the columns. Payoffs are each cell, with P_1's payoff listed first.

P_1, P_2	ROCK	PAPER	SCISSORS
ROCK	0, 0	−1, 1	1, −1
PAPER	1, −1	0, 0	−1, 1
SCISSORS	−1, 1	1, −1	0, 0

5.2.2 Nash Equilibrium

Does Rock-Paper-Scissors have a Nash equilibrium? To check whether a particular outcome is a Nash equilibrium we can follow the standard algorithm. In the first step, we posit the Nash equilibrium. Then in the second step, assume all but one player plays the posited strategy, determine the optimal strategy for the remaining player(s). If the posited strategy is not optimal, then it is not a Nash equilibrium.

- Step 1: ROCK, ROCK
- Step 2: P_1 plays ROCK. Is ROCK optimal for P_2?
  - ROCK: 0, PAPER: 1, SCISSORS: −1
  - No: Not a Nash equilibrium.

We see that if we posit that P_1 plays Rock, then it is optimal for P_2 to play Paper. Remember, Paper beats Rock. So $\{ROCK, ROCK\}$ is not a Nash equilibrium.

Similarly we can use the algorithm to determine if $\{PAPER, ROCK\}$ is a Nash equilibrium.

- Step 1: PAPER, ROCK
- Step 2: P_1 plays PAPER. Is ROCK optimal for P_2?

  - ROCK: -1, PAPER: 0, SCISSORS: 1
  - No: Not a Nash equilibrium.

Can you check other outcomes?

Chapter 2 states that there exists a Nash equilibrium for this game. Why can't we find it?

5.2.3 Mixed Strategies

Theorem 1. *For any finite game, there exists a Nash equilibrium, where that equilibrium may be in* mixed strategies*.*

Nash found that for any game with a finite number of players and strategies there exists a Nash equilibrium, but only if you allow players to play mixed strategies.

Definition 8. *A* mixed strategy *is a strategy that puts weights on each action, such that those weights are positive and sum to one.*

An example of a mixed strategy in Rock-Paper-Scissors is $\{0.2, 0.3, 0.5\}$, where that is 20% of the time the player chooses ROCK, 30% of the time they choose PAPER and 50% of the time they choose SCISSORS. This may not be a good mixed strategy but it is a mixed strategy. A strategy that puts 100 percent of the weight on a single action is a **pure strategy**. A **pure strategy Nash equilibrium** is just the Nash equilibrium discussed in Chapter 2.

5.2.4 Penalty Kicks

Let's consider another famous **zero-sum game**. In soccer there are situations where one player can kick the ball at the goal with only the goal keeper to try to stop it. This may occur when the defending team fouls the team with the ball. The result of the foul is that the team with the ball gets a free shot on goal. In soccer, goal keepers are small relative to the goal and the Kicker is close enough that it is probably the case that the Kicker will score if they kick to one side or the other, unless the Goalie moves that direction at the same time as the kick is made. Because the Kicker is so close, the Goalie can't wait to see where the ball is aimed before diving.

Here is a basic version of the game.

- Players: Kicker, Goalie
- Strategies:
  - Kicker: Kick Left, Kick Right
  - Goalie: Dive Left, Dive Right
- Payoffs: See Table 5.2

TABLE 5.2
Normal form representation of a penalty kicks game, where the Kicker's actions are in the rows, the payoffs are in the cells and the Kicker's payoff is listed first.

Kicker, Goalie	LEFT	RIGHT
LEFT	−1, 1	1, −1
RIGHT	1, −1	−1, 1

Table 5.2 presents the normal form representation of the game. Again, the values in each cell sum to zero. If the Kicker scores a goal, they get 1 and the Goalie gets −1. It is set up assuming that if the Kicker kicks left and the Goalie dives left then the goal is saved. The Goalie gets 1 and the Kicker gets −1. If the Goalie dives to the left and the kick goes right it is a goal! The Goalie gets −1 and the Kicker gets 1.

Below the chapter looks at real data from penalty kicks. In the real game, a score increases the probability that the scoring team wins (and decreases the probability that the Goalie's team wins).

5.2.5 Algorithim for Mixed Strategy Nash Equilibrium

To understand the algorithm for determining a mixed strategy, think about what must be true in equilibrium. If we are in a mixed strategy equilibrium then each player must be indifferent between their strategy choices when the other player's strategy is taken as given. The Goalie must be indifferent between diving to the left or diving to the right, given the Kicker's strategy. If that wasn't the case, the Goalie would prefer to dive left (or dive right) which is a **pure strategy**. So for there to exist a mixed strategy equilibrium both players must be exactly indifferent to each pure strategy. They must be sitting on this knife's edge. Sounds painful.

Let p be the probability that Kicker chooses LEFT. Let q be the probability Goalie chooses LEFT.

- Find q such that for Kicker, the payoff for LEFT equals the payoff for RIGHT.
- Find p such that for Goalie, the payoff for LEFT equals the payoff for RIGHT.

If the Goalie's strategy is q, the weight on diving LEFT, then the Kicker's payoffs are:

- LEFT: $(q)(-1) + (1-q)(1) = 1 - 2q$
- RIGHT: $(q)(1) + (1-q)(-1) = -1 + 2q$

Reading Table 5.2 and the first row, we see that if the first column is chosen, the Kicker gets -1 and 1 if second column is chosen (the first element).

The algorithm states that we need to find the q such that the Kicker is indifferent between LEFT and RIGHT. What strategy by the Goalie makes the Kicker indifferent between kicking LEFT or RIGHT? What strategy of the Goalie gives the Kicker the same expected payoff? What q is such that $1 - 2q = -1 + 2q$? To repeat. We need to find the strategy of the Goalie that makes the Kicker indifferent between her two choices.

$$\begin{aligned} &-4q = -2 \\ &4q = 2 \\ &q = \tfrac{2}{4} \\ &q = 0.5 \end{aligned} \tag{5.1}$$

Now let Kicker choose p, the weight on kicking LEFT. Goalie's payoffs are:

- LEFT: $(p)(1) + (1-p)(-1) = -1 + 2p$
- RIGHT: $(p)(-1) + (1-p)(1) = 1 - 2p$

To see this look at the first column of Table 5.2 and the second element, the Goalie gets 1 if LEFT is chosen and -1 if RIGHT is chosen.

To find the equilibrium level for p, determine where the Goalie is indifferent between LEFT and RIGHT. This time, we need to find the strategy of the Kicker that makes the Goalie indifferent between his choices.

$$\begin{aligned} &-1 + 2p = 1 - 2p \\ &4p = 2 \\ &p = \tfrac{2}{4} \\ &p = 0.5 \end{aligned} \tag{5.2}$$

Nash equilibrium is $\{p = 0.5, q = 0.5\}$

One of the greatest penalty scorers was the Argentinian, Diego Maradona. According to Maradona, his secret to penalty kicks was to wait and see what the goalie did and then do the opposite. He cheated! Just like you used to do with your younger sibling when playing Rock-Paper-Scissors. Maradona was playing a different game. A dynamic game.

5.3 Rock-Paper-Scissors

We saw above that there is no pure strategy Nash equilibrium for the Rock-Paper-Scissors game. This section finds the mixed strategy Nash equilibrium (MSNE). It then explores how the equilibrium changes when the game becomes more complicated.

5.3.1 Mixed Strategy Nash Equilibrium

Now we know how to find a mixed strategy Nash equilibrium, consider a standard version of Rock-Paper-Scissors. Can you guess what the equilibrium is?

Let Player 2 choose $\{q_1, q_2\}$. So q_1 is the probability Player 2 chooses ROCK, q_2 is the probability that Player 2 chooses PAPER and $1 - q_1 - q_2$ is the probability that Player 2 chooses SCISSORS. Find $\{q_1, q_2, 1 - q_1 - q_2\}$ such that Player 1 is indifferent.

Quick algorithm: Guess and confirm.

Guess: $\{q_1 = \frac{1}{3}, q_2 = \frac{1}{3}\}$ Confirm:

- ROCK: $(q_1)(0) + (q_2)(-1) + (1 - q_1 - q_2)(1) = \frac{0}{3} + \frac{-1}{3} + \frac{1}{3} = 0$
- PAPER: $(q_1)(1) + (q_2)(0) + (1 - q_1 - q_2)(-1) = \frac{1}{3} + \frac{0}{3} + \frac{-1}{3} = 0$
- SCISSORS: $(q_1)(-1) + (q_2)(1) + (1 - q_1 - q_2)(0) = \frac{-1}{3} + \frac{1}{3} + \frac{0}{3} = 0$

Confirmed!

Given that the two players are identical we can guess that the Nash equilibrium choices are the same: $\{\{p_1 = \frac{1}{3}, p_2 = \frac{1}{3}\}, \{q_1 = \frac{1}{3}, q_2 = \frac{1}{3}\}\}$. Can you check that this is in fact the Nash equilibrium?

Are there any others? How would you check?

5.3.2 More Complicated Version

Now let's make things a bit more interesting. Assume Player 1 gets a high value of playing ROCK when Player 2 plays SCISSORS. And because this is a zero sum game, Player 2 gets a low value from playing SCISSORS when Player 1 plays ROCK. Probably best not to ask too many more questions. Given this preference for ROCK, what do you think will be the Nash equilibrium?

That's a good guess, but it is wrong.

TABLE 5.3
Normal form representation of a more complicated Rock-Paper-Scissors. The payoff changed in the top right cell.

P_1, P_2	ROCK	PAPER	SCISSORS
ROCK	0, 0	−1, 1	2, −2
PAPER	1, −1	0,0	−1, 1
SCISSORS	−1, 1	1, −1	0, 0

What are the payoffs for Player 1? Let Player 2's strategy be $\{q_1, q_2, 1 - q_1 - q_2\}$.

- ROCK: $(q_1)(0) + (q_2)(-1) + (1 - q_1 - q_2)(2)$
- PAPER: $(q_1)(1) + (q_2)(0) + (1 - q_1 - q_2)(-1)$

- SCISSORS: $(q_1)(-1) + (q_2)(1) + (1 - q_1 - q_2)(0)$

This lets us work out which mixed strategy of Player 2 that makes Player 1 indifferent between their choices. OK. That is, for what values of q_1 and q_2 are all three values equal to each other? This looks complicated. Let's use the computer.

5.3.3 Using R to Determine when Player 1 is Indifferent

We can write the payoffs for Player 1 has functions of Player 2's mixed strategy choice. We can then plot the functions to see if there is a point where Player 1 is indifferent. We can keep adjusting q_1 and q_2 until all three lines intersect.

```
> rock = function(q1, q2) {
+   q1*0 + q2*(-1) + (1 - q1 - q2)*(2)
+ }
> paper = function(q1, q2) {
+   q1*1 + q2*0 + (1 - q1 - q2)*(-1)
+ }
> scissors = function(q1, q2) {
+   q1*(-1) + q2*(1) + (1 - q1 - q2)*(0)
+ }
```

The code creates the object `ggplot_rps` which is a plot of the expected value of playing ROCK, PAPER, and SCISSORS for different 10 different values of q_1 and $q_2 = 0.42$. We are cheating a bit by assuming we already know the value of q_2.

```
> ggplot_rps = data.frame(
+   q_1 = seq(0, 1, by = 0.1),
+   rock = rock(seq(0, 1, by = 0.1), rep(0.42, 10)),
+   paper = paper(seq(0, 1, by = 0.1), rep(0.42, 10)),
+   scissors = scissors(seq(0, 1, by = 0.1), rep(0.42, 10))
+ ) |>
+   ggplot(aes(x = q_1)) +
+     geom_line(aes(y = rock)) +
+     geom_line(aes(y = paper), linetype = "dotted") +
+     geom_line(aes(y = scissors), linetype = "dashed") +
+     geom_vline(xintercept = 0.33,
+               linetype = "dashed",
+               color = "gray") +
+     scale_x_continuous(breaks = seq(0, 1,
+                                     by = 0.1)) +
+     scale_y_continuous(breaks = seq(-2, 2,
+                                     by = 1)) +
+     labs(title = "Payoff to Player 1",
```

```
+          x = "Probability P2 chooses ROCK",
+          y = "") +
+      geom_text(aes(x = 0.5, y = -1, label = "rock")) +
+      geom_text(aes(x = 0.5, y = 1, label = "paper")) +
+      geom_text(aes(x = 0.8, y = 0, label = "scissors"))
```

```
> ggplot_rps
```

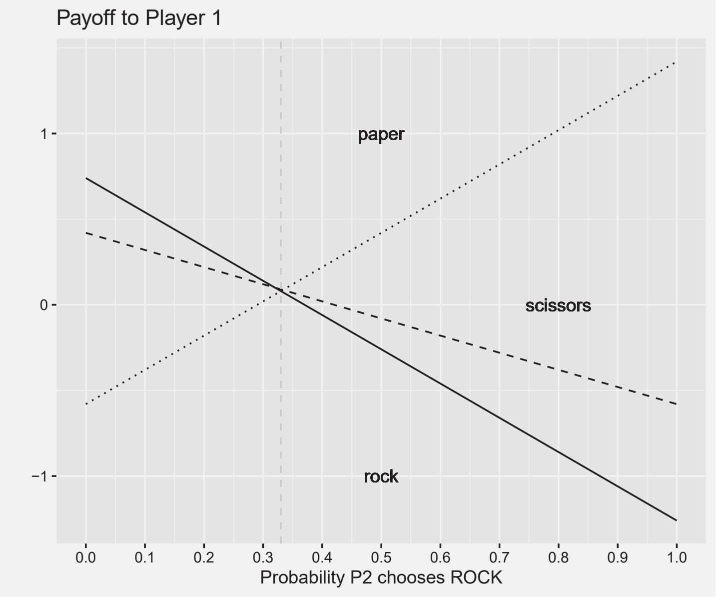


FIGURE 5.1
The figure plots the expected value of playing ROCK, PAPER and SCISSORS when $q_2 = 0.42$ for different values of q_1. Player 1 is indifferent between all three strategies when $q_1 = 0.33$.

Figure 5.1 shows where Player 1 is indifferent for different values of q. Player 2 tends to play PAPER! If Player 2 chooses ROCK about one-third of the time and PAPER 42% of the time, then Player 1 is indifferent.

5.3.4 Solving for MSNE using R

What about Player 2? Let Player 1 choose $\{p_1, p_2\}$.

- ROCK: $(p_1)(0) + (p_2)(1) + (1 - p_1 - p_2)(-1)$
- PAPER: $(p_1)(1) + (p_2)(0) + (1 - p_1 - p_2)(-1)$

- SCISSORS: $(p_1)(-2) + (p_2)(1) + (1 - p_1 - p_2)(0)$

For what strategy choice, is Player 2 indifferent between their three choices? For what values of p_1 and p_2 are these three values equal to each other?

```
> rock1 = function(p1, p2) {
+   p1*0 + p2*(1) + (1 - p1 - p2)*(-1)
+ }
> paper1 = function(p1, p2) {
+   p1*1 + p2*0 + (1 - p1 - p2)*(-1)
+ }
> scissors1 = function(p1, p2) {
+   p1*(-2) + p2*(1) + (1 - p1 - p2)*(0)
+ }
```

For this case, let's use computer muscle to solve the problem. The function `f_rps()` determines the values of p_1 and p_2 such that Player 2 is indifferent between the three choices. It does this by calculating the Euclidean distance between the choices, or the sum of squared differences. The function `f_rps_int()` is the intermediate function for `optim()`. This function translates the values chosen by the optimization algorithm into probabilities using the soft-max function.

```
> f_rps = function(p1, p2) {
+   (rock1(p1, p2) - paper1(p1, p2))^2 +
+            (rock1(p1, p2) - scissors1(p1, p2))^2
+ }
> f_rps_int = function(par) {
+   p = exp(par)/(1 + sum(exp(par)))
+   return(f_rps(p[1], p[2]))
+ }
> init = c(0,0)
> a_rps = optim(init, f_rps_int)
> exp(a_rps$par)/(1 + sum(exp(a_rps$par)))
 [1] 0.2499962 0.2500044
```

Player 1 plays ROCK with probability of about one quarter and plays PAPER with probability one quarter and SCISSORS half the time.

What the heck is going on? We said that Player 1 prefers ROCK but ends up playing SCISSORS most of the time. The reason is that Player 2 doesn't like it if Player 1 chooses ROCK so they tend to play PAPER and because Player 2 plays PAPER, Player 1 tends to play SCISSORS. Clear? No?

5.4 Empirical Analysis: Penalty Kicks using R

In soccer penalty kicks occur for two reasons. The first may occur if a player with the ball is fouled near goal, actually in what is known as the penalty box. This area extends 16 meters either side of the goal and 16 meters forward of the goal. Once fouled in this area, the team with the ball gets a penalty kick. The second is when there is a drawn game and penalty kicks are used to determine the winner. In both cases, the team with the ball gets to kick the ball from 11 meters in front of goal with only the goal keeper between the player and the goal. The Goalie has so little time to react to the kick that we can think of the Goalie and the Kicker choosing their strategies simultaneously.

The section analyzes data from the English Premier League, which is the highest league in English and Welsh soccer.

5.4.1 Penalty Kick Game

TABLE 5.4
Normal form representation of a penalty kicks game, where the Kicker's actions are in the rows, the payoffs are in the cells and the payoffs are the parameters to be estimated. It is a zero-sum game.

Kicker, Goalie	LEFT	CENTER	RIGHT
LEFT	p_{ll}	p_{lc}	p_{lr}
CENTER	p_{cl}	p_{cc}	p_{cr}
RIGHT	p_{rl}	p_{rc}	p_{rr}

Table 5.4 presents the normal form representation of the game we will take to the data. Each player has three action choices, LEFT, CENTER, and RIGHT. The payoffs are given by the 9 parameters to be estimated. The parameter p_{ij} is the probability that the Kicker scores a goal given that the Kicker chooses $i \in \{l, c, r\}$ and the Goalie chooses $j \in \{l, c, r\}$. The value to the Goalie is just the negative. It is a zero-sum game.

We will allow the two players to use mixed strategies. The vector representing the Kicker's strategy is as follows.

$$q_k = \left[\begin{array}{ccc} q_{kl} & q_{kc} & q_{kr} \end{array} \right] \tag{5.3}$$

The weights must some to one, $\sum_{i \in \{l,c,r\}} q_{ki} = 1$. Similarly, we can represent the Goalie's strategy.

$$q_g = \left[\begin{array}{ccc} q_{gl} & q_{gc} & q_{gr} \end{array} \right] \tag{5.4}$$

5.4.2 Data

The data used here are from Kaggle and covers penalty kicks in the English Premier League during the 2016/2017 season.[1]

```
> file = paste0(dir, "penalty_data.csv")
> data = fread(file) |>
+   filter(
+     Kick_Direction != ""
+   )
```

The code brings in the data and removes the cases where the kick direction is missing. The code uses `fread()` which is part of `data.table`. The first step is calculating the strategies for both players. The assumption is that every Kicker and Goalie is using the same strategy. Alternatively, you can think of this as the average strategy choice from a distribution of strategies. The code uses `table()` to count the number of each case and then divides by the total number of observations to get the probability of each strategy choice. The vector q_k is then reordered to match the order in Table 5.4. The same is done for the Goalie.

```
> q_k = table(data$Kick_Direction)
> q_k = q_k/sum(q_k)
> q_k = q_k[c(2, 1, 3)]


> q_g = table(data$Kick_Direction)
> q_g = q_g/sum(q_g)
> q_g = q_g[c(2, 1, 3)]
```

We can also calculate the parameter value from Table 5.4.

```
> action = c("L", "C", "R")
> res_mat = matrix(NA, 3, 3)
> for(i in 1:3) {
+   for(j in 1:3) {
+     dt_ij = data[data$Kick_Direction == action[i] &
+                  data$Keeper_Direction == action[j]]
+     res_mat[i,j] = mean(dt_ij$Scored == "Scored")
+   }
+ }
```

Table 5.5 presents the empirical results of the penalty kick game. Does it look like what you expected? We would expect a lower probability of scoring

[1] https://www.kaggle.com/datasets/mauryashubham/english-premier-league-penalty-dataset-201617 license: CCO: Public Domain, ShubhamMaurya, English Premier League Penalty Dataset, 201617.

TABLE 5.5
Strategies and conditional probabilities for penalty kicks in the English Premier League in the 2016/2017 season. The first column is Kicker's strategy and the first row is the Goalie's strategy. The cells the probability that the Kicker scores given the action chosen by the Kicker and the Goalie.

	q_k	G:LEFT	G:CENTER	G:RIGHT
q_g		0.46	0.17	0.38
K:LEFT	0.46	0.65	1.00	0.88
K:CENTER	0.17	1.00	0.00	0.89
K:RIGHT	0.38	0.83	1.00	0.56

when the Goalie chooses the same direction as the Kicker. These are on the diagonal and we see that in fact the lower probabilities are on the diagonal. Kicking to the center leads to a really low probability of scoring when the Goalie also chooses not to move. You may be surprised how often the Kicker still scores even when the Goalie chooses the same direction. It is 65 percent when kicking LEFT and 56 percent when kicking RIGHT. It is also interesting that it is not symmetric. The probability of scoring when kicking LEFT is higher than the probability of scoring when kicking RIGHT. Kickers are a little more likely to kick LEFT than RIGHT.

Right footed kickers may have a better chance of scoring when kicking to the LEFT. From a right footed kicker, the ball will naturally curve across and away from the Goalie. We can check to see if there are differences in payoffs and strategies between left and right footed kickers.

5.4.3 GMM Estimator

One issue with looking at differences by the footedness of the kicker is that the data is limited. A solution is to bring in more information from the game theory.

We know a few things that must be true in the game. In particular, in equilibrium, the strategy for the Kicker must make the Goalie indifferent between the choices. And the strategy for the Goalie must make the Kicker indifferent between the choices.

The Kicker and the Goalie choose q_k and q_g, respectively, such that the following equalities hold. The periods in the subscripts mean that it represents any choice.

$$\begin{array}{l} q_k' p_{.l} = q_k' p_{.c} = q_k' p_{.r} \\ q_g' p_{l.} = q_g' p_{c.} = q_g' p_{r.} \end{array} \tag{5.5}$$

The vector q_k is laid down on its side and we use matrix multiplication rules to multiply it with the conditional probabilities when the Goalie chooses LEFT. These conditional probabilities are the first column of Table 5.4. The elements in q_k are multiplied with the corresponding elements in $p_{.l}$ and the three numbers are summed together to give the probability of a goal when the Kicker plays strategy q_k and the Goalie plays LEFT.

We will guess the matrix of strategies, p, then use equilibrium relationship to solve for q_k and q_g. In addition, we directly observe q_k, q_g and p in the data. Don't we have too many conditions? Yes we do. Won't requiring that there is a mixed strategy Nash equilibrium lead to different estimates? Yes. Yes it will. The data combined with the game theory provide too many conditions and thus too many estimates. The solution is to average over the estimates. We are over identified. The generalized method of moments (GMM) algorithm provides a way to find that average. A method of moments estimator was used in Chapter 3. Here we have multiple **moments**.[2]

5.4.4 GMM Estimator in R

The first step is to create functions `f_mixed()` and `f_mixed_int()`. These functions are used to determine the equilibrium strategy of the Kicker and the Goalie given a set of conditional probabilities of scoring. Given a set of conditional probabilities `p`, `optim()` is used to determine the equilibrium strategies `q_k` and `q_g`. This is done by finding the strategies that minimize the sum of squared differences between payoffs of the three options for the other player.

```
> f_mixed = function(q_k, q_g, p) {
+   p = matrix(unlist(p), nrow=3)
+   pi_1 = t(q_k)%*%p[,1]
+   pi_2 = t(q_k)%*%p[,2]
+   pi_3 = t(q_k)%*%p[,3]
+
+   pi_4 = t(q_g)%*%p[1,]
+   pi_5 = t(q_g)%*%p[2,]
+   pi_6 = t(q_g)%*%p[3,]
+   return((pi_1 - pi_2)^2 + (pi_1 - pi_3)^2 +
+             (pi_4 - pi_5)^2 + (pi_4 - pi_6)^2)
+ }
> f_mixed_int = function(par, p) {
+   q_k = exp(par[1:3])/sum(exp(par[1:3]))
+   q_g = exp(par[4:6])/sum(exp(par[4:6]))
+   return(f_mixed(q_k, q_g, p))
+ }
```

The function `f_penalty()` takes in the proposed estimate of the conditional probabilities `p` and the observed actions and outcomes `X`. It then calculates

[2]The term refers to a characteristic of the distribution such as the mean (the first moment) or the variance (related to the second moment). But here we use the term more generally and include characteristics of the equilibrium.

7 **moments**. For instance, it determines the difference between the estimate of the Kicker's probability of kicking LEFT against the observed probability that the Kicker kicked LEFT. It also compares the proposed estimate of the conditional probabilities to the actual conditional probabilities in the data. The function `f_penalty_int()` is an intermediate function called by `optim()`. It turns the parameter values into numbers between 0 and 1. It then calls `f_gmm()`, which is a generic function for calculating the GMM optimization problem.[3] The standard errors are calculated using the bootstrap.

```
> f_penalty = function(p, X) {
+   N = dim(X)[1]
+   init = rep(0, 6)
+   p = matrix(p, nrow = 3)
+   q = optim(par = init, fn = f_mixed_int, p = p)
+   q_k = exp(q$par[1:3])/sum(exp(q$par[1:3]))
+   q_g = exp(q$par[4:6])/sum(exp(q$par[4:6]))
+   g_k = g_g = matrix(0, 3, N)
+   for(i in 1:3) {
+     g_k[i,] = q_k[i] - (X$kicker == action[i])
+     g_g[i,] = q_g[i] - (X$goalie == action[i])
+   }
+   g_s = rep(0, N)
+   for(i in 1:3) {
+     for(j in 1:3) {
+       index_ij = which(
+         X$kicker == action[i] & X$goalie == action[j]
+       )
+       g_s[index_ij] = p[i,j] -
+         (X$score[index_ij] == "TRUE")
+     }
+   }
+   G = rbind(g_k,
+             g_g,
+             g_s)
+   return(G)
+ }
> f_penalty_int = function(par, X) {
+   p = exp(par)/(1 + exp(par))
+   p = matrix(p, nrow = 3)
+   G = f_penalty(p, X)
+   return(f_gmm(G, K = 7))
+ }
```

[3]The `f_gmm()` function is available from the github site for the book.

5.4.5 Difference By Footedness

Does the behavior of the Kicker and the Goalie change depending on the which foot the Kicker generally kicks with? In general a right-footed Kicker will have an easier time scoring to the LEFT. The Kicker will generally strike the ball with their foot slightly to the right of center, that will put a counter-clockwise rotation on the ball which will then curve in the air from right to left.

So kicking LEFT for a right-footed Kicker will generally be better, holding the Goalie's strategy constant. Of course the Goalie gets a say in this. The Goalie may want to choose LEFT more often and thus be more likely to stop the ball when the right-footed Kicker chooses LEFT. For left-footed Kickers, the opposite is true.

```
> dt_R = dt[dt$Foot == "R"]
> p = matrix(NA, 3, 3)
> for(i in 1:3) {
+   for(j in 1:3) {
+     index_ij = which(dt_R$Kick_Direction == action[i] &
+                          dt_R$Keeper_Direction == action[j])
+     p[i,j] = mean(dt_R$Scored[index_ij] == "Scored")
+   }
+ }
> p[is.nan(p)] = NA
```

In the code the function `is.nan()` is used replace infinite values with missing values (`NA`). The matrix `p` gives the conditional probability that the Kicker scores given the Kicker is right footed.

```
> X = data.frame(
+   "kicker" = dt_R$Kick_Direction,
+   "goalie" = dt_R$Keeper_Direction,
+   "score" = dt_R$Scored == "Scored"
+ )
> epsilon = 1e-10
> init = log(p + epsilon)
> a = optim(par = init,
+           fn = f_penalty_int,
+           X = X,
+           control = list(trace = 0,
+                          maxit = 100000))
```

Similarly for Left footers.

```
> dt_L = dt[dt$Foot == "L"]
> p = matrix(NA, 3, 3)
> for(i in 1:3) {
+   for(j in 1:3) {
```

```
+     index_ij = which(dt_L$Kick_Direction==action[i] &
+                        dt_L$Keeper_Direction == action[j])
+     p[i,j] = mean(dt_L$Scored[index_ij] == "Scored")
+   }
+ }
> p[is.nan(p)] = 0
```

TABLE 5.6
Strategy estimates for Right and Left footed players. The first and fourth columns are the averages from the bootstrap for the two cases respectively. The top three rows are the strategies for the Kicker, while the bottom three rows are the strategies for the Goalie.

	Right	0.05	0.95	Left	0.05	0.95
Kicker - L	0.53	0.44	0.61	0.47	0.28	0.69
Kicker - C	0.19	0.13	0.26	0.01	0.00	0.06
Kicker - R	0.28	0.20	0.35	0.52	0.31	0.67
Goalie - L	0.54	0.44	0.62	0.19	0.03	0.36
Goalie - C	0.07	0.02	0.12	0.01	0.00	0.11
Goalie - R	0.39	0.31	0.48	0.80	0.64	0.95

Table 5.6 presents the strategies given the footedness of the Kicker. Right-footed kickers will generally more accurate kicking LEFT and we see that they choose LEFT 53% of the time, while left-footed kickers choose RIGHT 52% of the time. Goalie's respond by playing LEFT to a right-footed kicker 54% of the time and RIGHT to a left-footed kickers 80% of the time!

The analysis suggests that the Goalie's strategy is completely different depending on the footedness of the kicker.

5.5 Multiple Equilibria

Another reason why we may see mixed strategies in a game is because there are multiple equilibria. We saw this in the coordination game. The two players would like to coordinate but it is unclear which choice they should coordinate on. In that situation a mixed strategy Nash equilibrium may be a more reasonable prediction of the outcome. Similarly, there was a coordination problem in our analysis of entry of mega bookstores. This section reconsiders the problem assuming that Borders and Barnes & Noble are playing a mixed strategy Nash equilibrium.

5.5.1 Coordination Game

Here is the normal form representation presented in Chapter 2.

TABLE 5.7
Normal form representation of a coordination game from Chapter 2.

P_1, P_2	BLACK	RED
BLACK	2, 5	0, 0
RED	0, 0	5, 2

We use the same algorithm as for zero-sum games to find the MSNE. What is your guess on the equilibrium? Remember that we don't have the zero-sum game weirdness here.

Let q_1 and q_2, respectively, be the probability that Player 1 and Player 2 choose BLACK. What is the q_2 that makes Player 1 indifferent between BLACK and RED?

$$\begin{aligned} 2q_2 &= 5(1 - q_2) \\ 7q_2 &= 5 \\ q_2 &= \tfrac{5}{7} \end{aligned} \tag{5.6}$$

What is the q_1 that makes Player 2 indifferent between BLACK and RED?

$$\begin{aligned} 5q_1 &= 2(1 - q_1) \\ 7q_1 &= 2 \\ q_1 &= \tfrac{2}{7} \end{aligned} \tag{5.7}$$

Yes. Player 1 prefers RED and so weights their play to RED, while Player 2 prefers BLACK and weights their play to BLACK.

5.5.2 Bookstore Entry

Chapter 4 presents a game describing entry of two mega bookstores, Barnes & Noble and Borders. In the game Barnes & Noble will enter market i if the following inequality holds.

$$\mathbf{X}_i'\beta_1 - D_{2i}\alpha_1 + \xi_{1i} \geq 0 \tag{5.8}$$

where $D_{2i} \in \{0, 1\}$ represents the choice to enter market i of Borders. Under the assumptions of the equilibrium the value of D_{2i} is known by Barnes & Noble when they make their entry decision. Both bookstores also know ξ_{1i}, which is the value the econometrician doesn't observe.

In Chapter 4 we showed that there is a case where Barnes & Noble will enter but only if Borders does not, and similarly Borders will enter but only if Barnes & Noble does not. At these values of the ξs, we have a coordination game with multiple equilibria.

We can use the same algorithm to determine the mixed strategy Nash equilibrium. Let q_1 be the probability that Barnes & Noble enters and q_2 be the probability that Borders enters. What q_2 makes Barnes & Noble indifferent between entering and not entering?

$$\begin{array}{l} \mathbf{X}_i'\beta_1 - q_{2i}\alpha_1 + \xi_{1i} = 0 \\ or \\ q_{2i} = \frac{\mathbf{X}_i'\beta_1 + \xi_{1i}}{\alpha_1} \end{array} \tag{5.9}$$

Similarly we can solve for q_1

$$q_{1i} = \frac{\mathbf{X}_i'\beta_2 + \xi_{2i}}{\alpha_1} \tag{5.10}$$

If we assume that Barnes & Noble and Borders are at a mixed strategy Nash equilibrium, then we no longer have a indeterminacy problem with our estimator. For every value of ξ_1 and ξ_2 we have a known probability over which outcome will occur.

5.5.3 MSNE Estimator in R

The estimator is pretty similar to the one used in Chapter 4. There are a couple of differences with `f_entry_mix()`. First, the area where there is indeterminacy about which firm will enter, the mixed strategy Nash equilibrium determines what outcome will occur. Second, because there is no indeterminacy we can ask the probability that Firm 2 enters while Firm 1 does not.

To determine the mixed strategy Nash equilibrium we use Equations (5.10) and (5.9). The mixed strategy only occurs when the profits are such that it is only profitable for one of the two firms to enter. Because it is a mixed strategy, there is a possibility that any of the four outcomes occur. If there is an indeterminant outcome the mixed strategy Nash equilibrium determines the probabilities. In the code these are given by `q_1k` and `q_2k`. The probability of an indeterminant outcome is given by `p_ind`.

```
> f_entry_mix = function(X, beta_1, beta_2, alpha_1, alpha_2,
+                        rho) {
+   N = dim(X)[1]
+   xi_1 = Z_1
+   xi_2 = Z_2*sqrt(1 - rho^2) + rho*Z_1
+   Xb_1 = X%*%beta_1
+   Xb_2 = X%*%beta_2
+   p_00 = p_01 = p_11 = rep(0, N)
+   for(k in 1:K) {
+     pi_1k = Xb_1 + xi_1[k]
+     pi_2k = Xb_2 + xi_2[k]
+     q_1k = max(c(min(c((pi_2k)/alpha_2,1)),0))
+     q_2k = max(c(min(c((pi_1k)/alpha_1,1)),0))
+     p_ind = (pi_1k > 0 &
```

```
+                pi_1k - alpha_1 < 0 &
+                pi_2k > 0 &
+                pi_2k - alpha_2 < 0)
+     p_00 = p_00 +
+       (pi_1k < 0 & pi_2k < 0) +
+       p_ind*(1 - q_1k)*(1 - q_2k)
+     p_01 = p_01 +
+       (pi_1k < 0 & pi_2k > 0) +
+       p_ind*(1 - q_1k)*q_2k
+     p_11 = p_11 +
+       (pi_1k - alpha_1 > 0 & pi_2k - alpha_2 > 0) +
+       p_ind*q_1k*q_2k
+   }
+   return(list(p_00 = p_00/K,
+               p_01 = p_01/K,
+               p_11 = p_11/K))
+ }
```

5.6 Empirical Analysis: Bookstore Entry with MSNE using R

The analysis presented in this section uses exactly the same data used in Chapter 4. It also uses much of same empirical machinery.

Table 5.8 shows that the two different assumptions lead to similar results. Assuming that the outcome is a mixed strategy Nash equilibrium leads to an estimate that Borders is less impacted by competition with Barnes & Noble than the other way around. It also estimates less statistical dependence between the two stores.

Table 5.9 simulates the effect of a merger between Borders and Barnes & Noble on the willingness to have both brands in a market. The welfare impact of the merger is ambiguous. While it is the case that we see a reduction in competition. The merged firm is less like to have both brands in the market. In this way the merger leads to fewer brands in a market which reduces quality and the reduced head to head competition leads to higher prices. However, you see that there is a small reduction in the number of markets without a bookstore. The merger has the benefit of allowing the brands to coordinate their entry decision. While we have a game of complete information, the players cannot coordinate their choice in the mixed strategy Nash equilibrium. Because there is a possibility of having too many firms enter a market, firms reduce their willingness to enter some markets. The merger solves the coordination problem, reduces the possibility of having too many firms enter a market, and increases the willingness for the combined firm to enter some markets.

TABLE 5.8
Results from estimates of the game theory model from Chapter 4 and the model assuming a mixed strategy Nash equilibrium. The two columns labeled "Pure" refer to the case where the entry decisions of the two firms are both strategically and statistically dependent, but we assume a pure strategy Nash equilibrium. The two columns labeled "Mix" refer to the same model but assuming the outcome is a mixed strategy Nash equilibrium.

	Pure	SD	Mix	SD
const_1	−15.11	0.25	−14.82	0.20
Pop_1	1.07	0.01	1.05	0.02
Income_1	−0.76	0.48	−0.97	0.23
College_1	5.65	0.55	5.71	0.28
Stores_1990_1	0.37	0.11	0.39	0.07
const_2	−11.37	0.20	−11.48	0.20
Pop_2	0.65	0.02	0.63	0.02
Income_2	1.31	0.80	1.91	0.38
College_2	2.70	0.55	2.82	0.39
Stores_1990_2	0.79	0.11	0.77	0.10
alpha_1	0.73	0.18	0.70	0.21
alpha_2	0.70	0.12	0.56	0.15
rho	0.47	0.10	0.30	0.10

TABLE 5.9
Comparison of actual entry to simulated entry in the year 2000 and simulated entry under a merger. These results are based on the assumption that Borders and Barnes & Noble are playing a mixed strategy Nash equilibrium

	Actual	Sim	Merge
none	2919	2895	2881
BN	155	135	176
Borders	15	54	72
both	128	95	51

5.7 Discussion and Further Reading

Mixed strategy Nash equilibria are weird. Many people find them unintuitive and the algorithm for finding them is less than obvious. But for some games, they seem like the correct prediction.

While soccer penalty kicks may not represent the types of games you are interested in, it does provide an example of real people making real decisions with real consequences. Sporting contests provide researchers with access to large amounts of data on relatively simple strategic situations. This makes sports a good laboratory for testing game theory's predictions. Adams (2020) presents an analysis of play choice in NFL games.

The chapter introduces one of the most important empirical methodologies in structural estimation, the **generalized method of moments** (Hansen, 1982). Games often provides more moments than parameters to estimate, GMM provides a way to average over the moments.

The chapter also revisits the entry game analyzed in Chapter 4. The entry game has multiple equilibria. This chapter assumes that the outcome is the result of a mixed strategy Nash equilibrium. It revisits the merger simulation under the mixed strategy Nash equilibrium assumption. As before the merger reduces the number of markets with both brands, but unlike Chapter 4, the merger increases the number of markets with at least one firm.

Part II

Dynamic Games of Complete Information

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

6

Dynamic Games

6.1 Introduction

The first part of the book considers static games of complete information. These games have very simple strategies. Perhaps with the exception of games with mixed strategies, these games are straightforward to analyze. This part of the book considers dynamic games. These games have strategies that can be substantially more complex than those we have seen. This complexity will force us to make important simplifying assumptions.

A **dynamic game** refers to the information available to the player at the time of their choice of action. If the player has no information about what the other players are doing, then it is a static game. If at least one player is able to observe the actions of the other player prior to making their choice, then we have a dynamic game.

To restate whether or not a game of complete information is static or dynamic has to do with information, not time. We can have static games where players move at different times, but they don't get to observe the other player's action prior to their move. We can have dynamic games where players move more or less at the same time, but where at least one player observes the action of the other player prior to their move.

Definition 9. *In a* dynamic game *at least one player observes information about the other player's actions prior to choosing their action.*

This chapter introduces an alternative representation of a game, the **extensive form** representation. Chapter 1 introduced the **normal form** representation of a game. This is the matrix like object with one player on the columns and the other on the rows. The normal form representation is associated with static games while the **extensive form** representation is associated with dynamic games. There is nothing to stop you from using normal form representations for dynamic games or extensive form for static games, but we will see the value of the extensive form representation for dynamic games.

The chapter introduces a new equilibrium concept, **subgame perfection**. **Perfection** refers to equilibrium refinements. We are interested in reducing the number of predictions for the game that we believe are reasonable. The chapter revisits our analysis of entry game and the choices of Borders and

DOI: 10.1201/9781003351603-6

Barnes & Noble about which counties to enter. We can reduce the multiplicity of equilibria in the entry game by assuming that Barnes & Noble moves first and using subgame perfection.

6.2 Extensive Form

The section makes an adjustment to the **coordination game** introduced in Chapter 2 to illustrate the **extensive form** representation.

6.2.1 Coordination Game

Consider a variation on the coordination game. Assume that Player 1 chooses their action first and Player 2 observes that action before choosing their action.

- Players: Player 1 and Player 2
- Strategies:
  - Player 1: {BLACK, RED}
  - Player 2:
    * If Player 1 plays BLACK, then BLACK. If Player 1 plays RED, then BLACK.
    * If Player 1 plays BLACK, then RED. If Player 1 plays RED, then BLACK.
    * If Player 1 plays BLACK, then BLACK. If Player 1 plays RED, then RED.
    * If Player 1 plays BLACK, then RED. If Player 1 plays RED, then RED.
- Payoffs
  - {BLACK, {{BLACK: BLACK}, {RED: BLACK}}}: $\{2,5\}$
  - {BLACK, {{BLACK: RED}, {RED: BLACK}}}: $\{0,0\}$
  - {BLACK, {{BLACK: BLACK}, {RED: RED}}}: $\{2,5\}$
  - {BLACK, {{BLACK: RED}, {RED: RED}}}: $\{0,0\}$
  - {RED, {{BLACK: BLACK}, {RED, BLACK}}}: $\{0,0\}$
  - {RED, {{BLACK: RED}, {RED: BLACK}}}: $\{0,0\}$
  - {RED, {{BLACK: RED}, {RED: RED}}}: $\{5,2\}$
  - {RED, {{BLACK: BLACK}, {RED: RED}}}: $\{5,2\}$

Wow. Things got complicated right quick! Player 1's strategy is simple. This player does not have any information before they choose their action, so their strategies are just their actions. Things are a lot more complicated for Player 2. Player 2 gets to observe Player 1's action prior to making their choice. Therefore Player 2's strategy must account for this information. Player 2 has four possible strategies. This is because there are two possible states, Player 1's action choice, and two choices for each state. Two times two is four.

We will use the notation {BLACK: BLACK} to mean if Player 1 plays BLACK, then Player 2 plays BLACK.

6.2.2 Strategies Revisited

Strategies are one of the three main components of a game. In Chapter 1, we stated the definition of a strategy.

Definition 10. *A* strategy *is a function that maps from the player's information set to the player's actions.*

While in the first part of the book, a strategy was simply an action. Here it is a function. It maps from the information observed to an action. Moreover, it is a function that maps from every possibility to an action. Sure, this is just another way of saying it is a function, but it is really important to remember that it is a complete plan. It states what the player will do in every possible and every conceivable case.

6.2.3 Nash Equilibrium of the Coordination Game

Remember the Nash equilibrium algorithm asks us to posit a candidate set of strategies for all the players and then check that each player's strategy is optimal given the posited strategies.

Player 1's strategy is just RED or BLACK, while Player 2's strategy describes what they will do for the two possible cases. The strategy states what Player 2 will do for a situation that never actually occurs. This is what is meant by a complete plan. Player 2's strategy states what they will do under *every* circumstance, not just what happens in the proposed equilibrium.

Candidate: $\{RED, \{\{BLACK : BLACK\}, \{RED : RED\}\}\}$

- Assume Player 1 plays RED
- Player 2's payoffs
  - $\{\{BLACK : BLACK\}, \{RED : RED\}\}$: 2
  - $\{\{BLACK : RED\}, \{RED : RED\}\}$: 2
  - $\{\{BLACK : RED\}, \{RED : BLACK\}\}$: 0
  - $\{\{BLACK : BLACK\}, \{RED : BLACK\}\}$: 0
  - Yes. Player 2's strategy is optimal (not dominated).

- Assume Player 2 plays $\{\{BLACK : BLACK\}, \{RED : RED\}\}$
- Player 1's payoffs
  - RED: 5
  - BLACK: 2
  - Yes. It is a Nash equilibrium.

It is an equilibrium for Player 1 to choose RED and for Player 2 to also choose RED.

Are there any others?
Candidate: $\{BLACK, \{\{BLACK : BLACK\}, \{RED : BLACK\}\}\}$

- Assume Player 1 plays BLACK
- Player 2's payoffs
  - $\{\{BLACK : BLACK\}, \{RED : RED\}\}$: 5
  - $\{\{BLACK : RED\}, \{RED : RED\}\}$: 0
  - $\{\{BLACK : RED\}, \{RED : BLACK\}\}$: 0
  - $\{\{BLACK : BLACK\}, \{RED : BLACK\}\}$: 5
  - Yes. Player 2's strategy is optimal (not dominated).
- Assume Player 2 plays $\{\{BLACK : BLACK\}, \{RED : BLACK\}$
- Player 1's payoffs
  - RED: 0
  - BLACK: 1
  - Yes. It is a Nash equilibrium.

There are at least two Nash equilibrium.

Is $\{BLACK, \{\{BLACK : BLACK\}, \{RED : BLACK\}\}\}$ a likely outcome from the game? The two players coordinate on black, which is fine. What about the idea that Player 2 would choose BLACK when they know Player 1 has chosen RED. Is that reasonable?

6.2.4 Game Tree

One way to see the issue with the Nash equilibria is to look at the **game tree**, the extensive form representation presented in Figure 6.1.

A **game tree** is a directed graph, it consists of **nodes** and **edges**. The nodes represent places where a player makes a decision. An edge is the line and arrow from node to a node further along the game tree. The edge represents the flow of information.

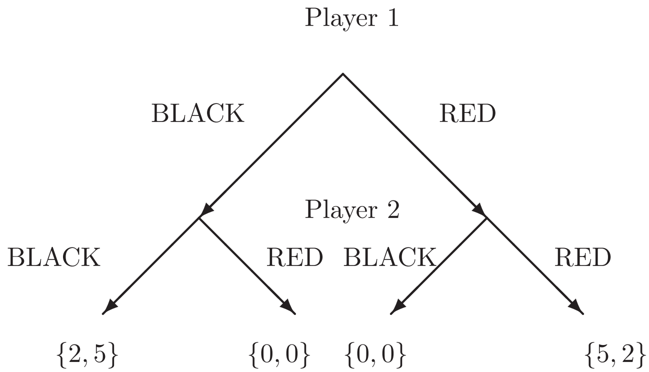


FIGURE 6.1
Extensive form representation of the dynamic coordination game tree.

Definition 11. *A* node *of a game tree is a place where the player makes a choice.*

Consider the Nash equilibrium $\{BLACK, \{BLACK : BLACK, RED : BLACK\}\}$. We checked that it was in fact a Nash equilibrium above. Look at what happens on the tree. If Player 1 plays RED, then Player 2 has a choice of BLACK or RED. In the equilibrium, they choose BLACK, but their payoff would have been higher if they had chosen RED.

Working down the tree, the equilibrium states that Player 1 chooses RED, so we go down the right branch. Now it is Player 2's turn. They can choose RED or BLACK. If they choose RED they get 2, while if they choose BLACK they get 0. In equilibrium they choose BLACK. Does that make sense? In the next section, we will consider an equilibrium refinement that rules out this case.

6.3 Subgame Perfection

The first equilibrium refinement we will consider is **subgame perfection**. An equilibrium refinement is an additional property of the predicted outcome that must be true. **Subgame perfection** requires a set of strategies be a Nash equilibrium but also that the subset of strategies associated with each **subgame** be a Nash equilibrium of the subgame.

The section defines subgame perfection and then works through the implications for the coordination game presented above.

6.3.1 Definition

Definition 12. *A* subgame *is a game that can be played from any node of the game tree.*

Definition 12 introduces the idea of a **subgame**. When you look at the game tree in Figure 6.1, there are three distinct nodes. For Player 2, there is the node after which Player 1 plays black and the node after which Player 1 plays red. There is also the initial node where Player 1 makes their choice of black or red. At each of these nodes we describe a separate game. This game is a subgame.

Definition 13. *A* subgame perfect Nash equilibrium *is a Nash equilibrium where the strategies in each subgame are a Nash equilibrium from that subgame.*

If for a particular Nash equilibrium strategy set we can look at each subgame and associate the strategies with a Nash equilibrium for that subgame, then the strategy set is **subgame perfect**.

6.3.2 Coordination Game

In the **coordination game**, there are three subgames. The whole game is a subgame. The other two begin at Player 2's decision node. In these two subgames, there is just one player (Player 2) and their strategies are just the actions {BLACK, RED}.

Is {BLACK, {{BLACK: BLACK}, {RED: BLACK}}} subgame perfect?

Consider the subgame at the node where Player 1 choose RED.

- Players: Player 2
- Strategies: {BLACK, RED}
- Payoffs: BLACK: 0, RED: 2

It is not a Nash equilibrium of the subgame for Player 2 to choose BLACK because they would be better off choosing RED.

Subgame perfection removes outcomes that allow non-credible strategies. Nash equilibrium requires the players to choose strategies that are optimal given the strategies of the other players. Subgame perfection requires players to choose actions that are optimal even if in equilibrium these actions will never be played.

Is it better to be Player 1 or Player 2 in this game? You may think Player 2, as they get to see what Player 1 does and react to it. If we rule out non-credible threats, then Player 1 always gets their way. Player 1 has a **first-mover advantage**. If the first move in the game can commit to a strategy then they have a distinct advantage. They can force the other player to choose actions that the first player prefers.

6.3.3 Empirical Entry Game

Consider a different version of entry game analyzed in Chapter 4. Instead of having Borders and Barnes & Noble enter at the same time, assume that Barnes & Noble moves first. Borders observed Barnes & Noble's decision to enter or not and then decides to enter.

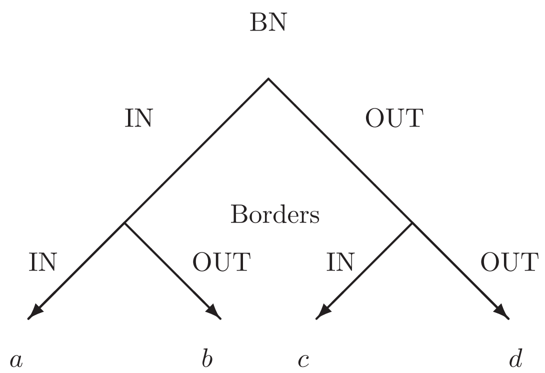


FIGURE 6.2
Dynamic entry game tree. $a = \{\mathbf{X}_i'\beta_1 - \alpha_1 + \xi_{1i}, \mathbf{X}_i'\beta_2 - \alpha_2 + \xi_{2i}\}$, $b = \{\mathbf{X}_i'\beta_1 + \xi_{1i}, 0\}$, $c = \{0, \mathbf{X}_i'\beta_2 + \xi_{2i}\}$ and $d = \{0, 0\}$.

Figure 6.2 represents the entry game. Barnes & Noble moves first and chooses whether or not to enter the market and then Borders chooses. The payoffs state that upon entry the firm earns $\mathbf{X}_i'\beta_1$, but if they face competition then those profits fall by α_1. The entry costs are captured by ξ_{1i}.

6.3.4 Equilibrium in Entry Game

To determine the **subgame perfect Nash equilibrium**, we solve the last subgame first. If Barnes & Noble entered the market, Borders will enter if and only if the following inequality holds.

$$\mathbf{X}_i'\beta_2 - \alpha_2 + \xi_{2i} \geq 0 \tag{6.1}$$

Borders will only enter if the duopoly profits more than outweigh the entry costs. While if Barnes & Noble has not entered, Borders will enter if the following inequality holds.

$$\mathbf{X}_i'\beta_2 + \xi_{2i} \geq 0 \tag{6.2}$$

This time, Borders enters if monopoly profits are high enough.

Now we go back and consider Barnes & Noble's choices. If they enter, there are two cases. Case 1 is that Borders enters ($\mathbf{X}_i'\beta_2 - \alpha_2 + \xi_{2i} \geq 0$), they will also enter if the following inequality holds.

$$\mathbf{X}_i'\beta_1 - \alpha_1 + \xi_{1i} \geq 0 \tag{6.3}$$

Case 2 is that Borders does not enter, and so Barnes & Noble will enter if and only if $\mathbf{X}_i'\beta_1 + \xi_{1i} \geq 0$.

If Barnes & Noble chooses not to enter, then there are also two cases. In Case 1, where Borders enters, it is an equilibrium if and only if $\mathbf{X}_i'\beta_1 - \alpha_1 + \xi_{1i} \leq 0$. In Case 2, it is an equilibrium if and only if $\mathbf{X}_i'\beta_1 + \xi_{1i} \leq 0$.

6.3.5 Empirical Implications

To summarize our four observed outcomes are a subgame perfect Nash equilibrium if the following inequalities hold.

- Both enter: $\mathbf{X}_i'\beta_1 - \alpha_1 + \xi_{1i} \geq 0$ and $\mathbf{X}_i'\beta_2 - \alpha_2 + \xi_{2i} \geq 0$.
- BN enters only: $\mathbf{X}_i'\beta_1 + \xi_{1i} \geq 0$ and $\mathbf{X}_i'\beta_2 - \alpha_2 + \xi_{2i} \leq 0$
- Borders enters only: $\mathbf{X}_i'\beta_1 - \alpha_1 + \xi_{1i} \leq 0$ and $\mathbf{X}_i'\beta_2 + \xi_{2i} \geq 0$.
- Neither firm enters: $\mathbf{X}_i'\beta_1 + \xi_{1i} \leq 0$ and $\mathbf{X}_i'\beta_2 + \xi_{2i} \leq 0$

How do these compare to the outcomes in Chapter 4? There is no indeterminacy. If you go back and look at Figure 4.2 and the middle square it has $\{0, 1\}$ or $\{1, 0\}$. Under this set up with subgame perfection it becomes $\{1, 0\}$. Only Barnes & Noble will enter. Where before it there were two equilibria (more if you include mixed strategies), here there is always a unique subgame perfect Nash equilibrium. Of course, the price is that we need to make very strong assumptions about how the game is played.

6.4 Empirical Analysis: Bookstore Entry with Subgame Perfection in R

The estimator for the subgame perfect Nash equilibrium is basically the same as in Chapter 4. The difference is that `f_entry_spne()` can separately estimate the cases where there is just one firm. Where previously there were multiple equilibria, now Barnes & Noble enters, while Borders does not.

6.4.1 SPNE Estimator

The code for the estimator is somewhat longer than the estimator used in Chapter 4. The reason is that we now assume a unique equilibrium and so we can separately estimate all four possible cases.

```
> f_entry_spne = function(X, beta_1, beta_2, alpha_1,
+                         alpha_2, rho) {
+   N = dim(X)[1]
+   xi_1 = Z_1
+   xi_2 = Z_2*sqrt(1 - rho^2) + rho*Z_1
+   Xb_1 = X%*%beta_1
+   Xb_2 = X%*%beta_2
+   p_00 = p_01 = p_11 = rep(0, N)
+   for(k in 1:K) {
+     pi_1k = Xb_1 + xi_1[k]
+     pi_2k = Xb_2 + xi_2[k]
+     p_00 = p_00 + (pi_1k < 0 & pi_2k < 0)
+     p_01 = p_01 + (pi_1k - alpha_1 < 0 & pi_2k > 0)
+     p_11 = p_11 + (pi_1k - alpha_1 > 0 &  pi_2k - alpha_2 > 0)
+   }
+   return(list(p_00 = p_00/K,
+               p_01 = p_01/K,
+               p_11 = p_11/K))
+ }
```

6.4.2 SPNE Estimates

Table 6.1 shows that the assumption about simultaneous move gives very similar estimates to the assumption that Barnes & Noble moves first. The big difference is on the estimates of the impact of competition on the two firms. In order to reconcile the observed entry decisions with modeling assumptions, the estimator states Barnes & Noble is not affected much by competition, while Borders is.

Table 6.2 compares the model predictions to the actual data for the model presented in Chapter 4 and for a model presented here where Barnes & Noble moves first and there is a subgame perfect Nash equilibrium. The subgame perfect Nash equilibrium model does a better job of predicting the case where there is only one firm, but a substantially worse job of predicting the two-firm case.

6.5 Discussion and Further Reading

Notice that so far in this part of the book, there has been no discussion of time in these so-called dynamic games. This goes back to the point made earlier, dynamics has to do with information not time. Chapters 7 and 8 introduce time.

TABLE 6.1
Results from estimates of the game theory model from Chapter 4 and the model assuming Barnes & Noble moves first with a subgame perfect Nash equilibrium. The two columns labeled "Multi" refer to the case where we assume there could between two pure strategy Nash equilibrium. The two columns labeled "SPNE" refer to the model where Barnes & Noble moves first and there is a subgame perfect Nash equilibrium.

	Multi	SD	SPNE	SD
const_1	−15.11	0.25	−15.12	0.06
Pop_1	1.07	0.01	1.04	0.02
Income_1	−0.76	0.48	−1.01	0.21
College_1	5.65	0.55	5.61	0.19
Stores_1990_1	0.37	0.11	0.15	0.06
const_2	−11.37	0.20	−11.15	0.15
Pop_2	0.65	0.02	0.65	0.02
Income_2	1.31	0.80	1.93	0.17
College_2	2.70	0.55	2.85	0.17
Stores_1990_2	0.79	0.11	0.63	0.08
alpha_1	0.73	0.18	0.50	0.16
alpha_2	0.70	0.12	1.08	0.14
rho	0.47	0.10	0.39	0.06

TABLE 6.2
Comparison of predictions of the two models. The subgame perfect Nash equilibrium is somewhat better at fitting the case where there is only one firm, but not when there are two firms.

	None	One Firm	Two Firm
Multi: None	97.1	41.0	9.0
Multi: One	2.6	41.7	36.6
Multi: Two	0.3	17.3	54.4
SPNE: None	97.5	48.8	14.0
SPNE: One	2.4	43.2	53.3
SPNE: Two	0.1	8.0	32.7

The empirical analysis in this chapter is based on work presented in Adams and Basker (2025). The authors use information on the location of Barnes & Nobel and Borders collected from bookstore directories and firm websites to analyze the dynamics of the retail bookstore industry.

7

Repeated Games

7.1 Introduction

Repeated games are the lens through which we are adapting our thinking on competition and competition policy. In the late 1990s, the economics of competition policy changed pretty dramatically. Game theory and structural econometrics provided antitrust authorities with new tools for modeling mergers. Chapters 3 and 4 introduce methods used to analyze the impact of retail mergers. While these models substantially improved our ability to understand competition and predict the outcomes of mergers, something was not quite right. Our approach to collusion and markets with collusive pricing remained rudimentary. Using the new models to analyze collusion was like pushing a square peg into a round hole.

The analysis presented here on collusive pricing is influenced by two papers. First, economists Nathan Miller, Gloria Sheu, and Matthew Weinberg presented compelling evidence that our new models worked poorly when used to analyze the US beer industry. Their paper, "Oligopolistic Price Leadership and Mergers: The United States Beer Industry," was published in the *American Economic Review* in 2021, and suggested a major rethink in the models we need for analyzing competition. Second, Canadian and Australian economists, David Byrne and Nic de Roos, analyzed pricing in the retail gasoline market in Perth Western Australia. Their paper, "Learning to Coordinate: A study in retail gasoline" was published in the *American Economic Review* in 2019. The authors use daily pricing data from a large number of retailers to show evidence of price leadership and coordination.

This chapter shows how the lens of repeated games can be used to explain firm behavior and pricing. Repeated interactions change the strategic relationships quite dramatically. In a single shot game, players don't have to account for the consequences of their actions. In repeated games they do.

Using data from Perth retail gasoline stations, the chapter presents two models of competition and pricing based on repeated interactions. The first model is a standard static pricing model that was introduced in Chapter 3. The second is a collusive oligopoly pricing model. This model shows how pricing is constrained by the incentives of firms to cheat and choose a lower price. The chapter estimates the parameters using data from gas stations in Perth

DOI: 10.1201/9781003351603-7

in a period where margins where low in 2008. It then compares the profit margins predicted by the collusive model to the actual profit margins of the same firms in 2012. While well-intentioned, price transparency regulation by the state government seems to cause Perth motorists to pay more for their petrol (gasoline).

The chapter runs a merger simulation using each of the two models of pricing behavior. It suggests that merger analysis should account for these changes when analyzing the likely effect of the merger. The predicted price increase using a repeated game model may be substantially higher than the predicted price increase using the standard static game presented earlier in the book.

7.2 Repeated Prisoner's Dilemma

The prisoner's dilemma is the most famous games in game theory. The game shows that the predicted outcome may not be the outcome that is best for both players. It may not be Pareto efficient. How much of that result is due to the set up of the game? What if players had to account for each other's actions? What if the game repeated?

This section presents the prisoner's dilemma game and shows how the Nash equilibrium change when dynamics are added. In particular, it shows when playing Cooperate is supported as a subgame perfect Nash equilibrium.

7.2.1 Prisoner's Dilemma

The game set up is as follows. We will relabel the strategies of the game presented in Chapters 1 and 2.

- Players: Player 1 and Player 2
- Strategies: Cooperate or Defect
- Payoffs:
  - {Cooperate, Cooperate}: {3, 3}
  - {Cooperate, Defect}: {0, 5}
  - {Defect, Cooperate}: {5, 0}
  - {Defect, Defect}: {2, 2}

In this game, the Nash equilibrium is {Defect, Defect}, even though both would be better off with the {Cooperate, Cooperate} outcome. The question is whether making a slight change to the set up of the game changes the predicted outcome.

7.2.2 Normal Form Representation

Chapter 1 introduced the normal form representation of a game.

TABLE 7.1
Normal form representation prisoner's dilemma game, with two players, Player 1 and Player 2. For Player 1, their choices are the rows and their payoffs are listed first in each cell.

P_1, P_2	Cooperate	Defect
Cooperate	3, 3	0, 5
Defect	5, 0	2, 2

Table 7.1 presents the **normal form** of the prisoner's dilemma game. We can determine the Nash equilibrium by looking at the second column and seeing whether Player 1 is better off choosing the top row or the bottom row. Player 1 gets nothing from choosing the top row and 2 from the bottom. Similarly for Player 2, we can check the Nash equilibrium by looking at the bottom row and seeing if Player 2 wants the first column or the second column. Player 2 gets 0 from the first column and 2 from the second column. {Defect, Defect} is a Nash equilibrium.

7.2.3 Finitely Repeated Game

Consider a version of the game where the static prisoner's dilemma (above) is repeated a finite number of times (T times). This could be 20 periods, for example.

The game is different. The players are the same but the strategies are completely different. A **strategy** is a function that maps from the complete history of the game to an action. In period 1, the complete history is null, so the strategy is just the action choices {Cooperate, Defect}. In period 2, the complete history is what ever the outcome was in period 1. There are 4 possibilities as listed above. From each possibility, the strategy states which action the player will choose. In period 3, things are even more complicated. Now the history includes 16 possibilities. For each of the 4 possible outcomes in period 1, there are 4 possible outcomes in period 2. Again the strategy states what the player will do in period 3 given each of the 16 possible histories. As you can imagine, for a game with twenty periods, there are an awful lot of possibilities.

What is the Nash equilibrium of this game? It is probably better to ask, what isn't. Consider any set of strategies where one player plays Cooperate in the last period. The game in the last period is essentially the same as a one period game. We know from the analysis of the one period game that the best response must be a strategy where they play Defect in the last period. Now consider sets of strategies where the player always plays Defect in the last period but Cooperate in the second to last period. Again, the best response must be a strategy that states the player plays Defect in the second to last period and Defect in the last period. Using this logic, we can show that the Nash equilibria are associated with playing Defect in every period. That is to say, despite adding a lot of complexity to the game, the prediction doesn't change.

7.2.4 Infinitely Repeated Prisoner's Dilemma

What if we make one more change to the game? This time, we repeat the game above for an infinite number of periods. Does this seem like a reasonable change? Infinity is quite a long time. For the strategic behavior to change, we don't literally need an infinite number of periods, we just need the players to be unsure when the game is going to end. It is better to think of a finite period game as one where every player knows exactly when the game is going to end. While an infinitely repeated game is one where players don't know exactly when the game will end.

For this case, we need to add another parameter to the payoffs, $r \in [0, 1)$. This represents a discount rate and is a number between 0 and 1. The practical reason for adding a discount rate is that with an infinite number of periods we cannot analyze the game. The payoffs are infinite under all possible strategies. As you may remember from calculus, an infinite sum of a non-decreasing sequence is infinite. But with a discount rate, we can create a decreasing sequence that decreases fast enough that our infinite sum sums to a finite number. A player's utility given a particular strategy s is an infinite sum of a discounted sequence of per period payoffs.

$$U(s) = \sum_{t=1}^{\infty} r^t \pi_t(s) \tag{7.1}$$

where s is the strategy. If we assume that the payoff to the players in each period ($\pi_t(s)$) is bounded and there is a discount rate so that the payoff is lower in the future than today ($r < 1$), then the infinite sum, $U(s) < \infty$. Now we can analyze the game because different strategies may have different payoffs.

We can also use a very useful trick. If $\pi_t(s) = \pi(s)$, that is, if the per-period payoff is constant for a given strategy s and r is strictly between 0 and 1, which it is, we have the following simplification.

$$U(s) = \frac{\pi(s)}{1 - r} \tag{7.2}$$

So if the discount rate is 0.9, then the total payoff is 10 times the per-period payoff for the strategy s.

Where does this discount rate come from? What does it mean? The most obvious way to think about it is as an interest rate. Actually, one minus the interest rate. When we are thinking about dynamic decisions it makes sense for decision makers to refer to the interest rate when determining the value of future decisions. Here, it may be reasonable to think of the discount rate as representing the probability that the game continues into the next period.

7.2.5 Cooperation as a Nash Equilibrium

Can having both players cooperate every period be supported as a Nash equilibrium of the infinitely repeated prisoner's dilemma? Yes.

To construct the supporting strategies, we need to allow players to punish defection. There are many strategies that have this feature but the simplest is called the **grim-trigger strategy**. In this strategy, the player plays Cooperate unless the history includes one of the players playing Defect. In that case, the player plays Defect forever. So "grim" is for the fact that the threat is the worst possible kind and "trigger" is for the fact that the strategy involves a simple event that changes behavior.

A strategy maps from the complete history of the game into the choice of Defect or Cooperate in each period. It is a set of functions for each period t, $\{s_t\}_{t=1}^{\infty}$. Each function is $s_t : H_t \to \{Defect, Cooperate\}$, where H_t is the complete history of possible actions that could have happened in the previous $t-1$ periods. In this case, if H_t includes Defect, then $s_t(H_t) = Defect$ while if it doesn't, then $s_t(H_t) = Cooperate$.

Is the grim-trigger a subgame perfect Nash equilibrium? Consider the case where both players are playing the grim-trigger. Payoffs are as listed in the previous section.

1. Grim-trigger: $\frac{3}{1-r}$
2. In period t play Defect: $5 + r\left(\frac{2}{1-r}\right)$

So against the grim-trigger in which the other player is playing Cooperate, if you play Cooperate each period you get 3 each period. If you play Defect in one period, you get 5 for that period but the grim strategy kicks in and you get 2 for every period. For the grim-trigger to be an equilibrium choice (1) must give a higher payoff than choice (2).

The player would prefer to play grim-trigger if the following inequality holds.

$$\begin{aligned} &\tfrac{3}{1-r} > 5 + r\left(\tfrac{2}{1-r}\right) \\ &3 > 5(1-r) + r2 \\ &3 > 5 - 5r + 2r \\ &0 > 2 - 3r \\ &3r > 2 \\ &r > \tfrac{2}{3} \end{aligned} \tag{7.3}$$

As long as the players care about the future enough, r is high enough, they will play Cooperate against the grim-trigger. If they are very myopic, r is low, then they will cheat and take the high payoff today.

Just to tie things up we should also check that both players will keep playing Defect after the trigger has been pulled. Looking back at the discussion above we see that this must be the case.

7.3 Bertrand Competition

This section revisits the Bertrand pricing game presented in Chapter 3. It presents this model assuming logit demand, which is the more common assumption in the literature. It uses this model to estimate parameters of demand using data for retail gasoline from Perth Australia.

7.3.1 Two Firm Model

Consider a model where we have two firms that choose prices. The two firms sell similar products but are not exactly the same. For example, the products may be stores that are located in different places. The differentiation between the products is enough to induce some market power for each firm. That is for a small price increase, the firm is not going to lose all its customers. Below we will use this model to analyze retail gasolone. The gasoline itself is identical as it literally comes from the same pipe.[1] What is different between stations is there location and their brand.

In this model, Firm 1 chooses their price (p_1) to maximize the margin which is price less cost (c_1) multiplied by Firm 1's share ($s_1(p_1, p_2)$)

$$\max_{p_1} \quad (p_1 - c_1) s_1(p_1, p_2) \tag{7.4}$$

The logit model assumes demand has the following form.

$$s_1(p_1, p_2) = \frac{\exp(\delta_1)}{\exp(\delta_1) + \exp(\delta_2)} \tag{7.5}$$

where $\delta_1 = -\alpha_1 p_1 + \xi_1$ and $\delta_2 = -\alpha_2 p_2 + \xi_2$. In this model, demand for Firm 1 is determined by firm specific value ξ_1 and by Firm 1's price p_1. Sensitivity to price is determined by the parameter α_1.

7.3.2 J Firm Model

Assume that we have J firms and logit demand.

$$\delta_j = -\alpha_j p_j + \xi_j \tag{7.6}$$

$$s_j(p_j, p_{-j}) = \frac{\exp(\delta_j)}{1 + \sum_{j'=1}^{J} \exp(\delta_{j'})} \tag{7.7}$$

[1] Some stations may add in additives at the end.

where p_j is the price charged by firm j, α_j represents the price sensitivity of firm j's customers, and ξ_j is a set of characteristics other than price that determine demand. The notation $-j$ means all the other firms that are not j.

7.3.3 Nash Equilibrium in R

The following are the functions for determining the Nash equilibrium using the logit demand system. The function `f_share()` determines the shares from the logit model given the vectors of ξs, αs and prices. The function `pi_f()` determines the vector of firm profits, while `pi_i()` determines the profits for firm i given a vector of prices for the other firms. The function `pi_i_opt()` determines firm i's best response to a set of prices, while `pi_opt()` determines the vector of best response. The function `ne_f()` determines the Nash equilibrium vector of prices.

```
> f_share = function(xi, price, alpha) {
+   delta = xi - alpha*price
+   exp_delta = c(1,exp(delta))
+   return((exp_delta/sum(exp_delta))[-1])
+ }
> pi_f = function(xi, price, alpha, cost) {
+   return((price - cost)*f_share(xi, price, alpha))
+ }
> pi_i = function(i, p_i, xi, price, alpha, cost) {
+   price[i] = p_i
+   return(pi_f(xi, price, alpha, cost)[i])
+ }
> pi_i_opt = function(i, xi, price, alpha, cost) {
+   a_i = optimize(f=pi_i, interval=c(0, 400),
+                  i = i,
+                  xi = xi,
+                  price = price,
+                  cost = cost,
+                  alpha = alpha,
+                  maximum = TRUE)
+   return(a_i$maximum)
+ }
> pi_opt = function(xi, price, alpha, cost) {
+   N = length(xi)
+   price_new = rep(NA, N)
+   for(i in 1:N) {
+     price_new[i] = pi_i_opt(i,xi,price,alpha,cost)
+   }
+   return(price_new)
+ }
```

Nash equilibrium is determined using a similar iterative approach that we used to determine equilibrium of the Cournot model in Chapter 3. It uses a `while()` to determine when the sequence of price choices converges.

```
> ne_f = function(xi, price, alpha, cost,
+                 tol=1e-10, maxiter=10000,
+                 trace=0) {
+   price_0 = price
+   diff = sum(abs(price_0))
+   iter = 1
+   converge = FALSE
+   while(diff > tol & iter < maxiter) {
+     price_1 = pi_opt(xi, price_0, alpha, cost)
+     diff = sum(abs(price_1 - price_0))
+     price_0 = price_1
+     if(trace > 0) {
+       print(diff)
+       print(iter)
+     }
+     iter = iter + 1
+   }
+   if(diff < tol) {converge=TRUE}
+   return(list(price=price_1,converge=converge))
+ }
```

7.4 Empirical Analysis: Retail Gasoline Pricing using R

This section estimates the parameters of the model assuming the Perth retail gas oline market is priced competitively, i.e., consistent with static Bertrand Nash equilibrium.

7.4.1 Perth Gas Price Data

Figure 7.1 presents the average margins by week for the gas stations. While there is a lot of variation in margins in 2008 and 2009, the average stays pretty steady. Things seems to change in 2010, with margins steadily increasing. Why would that be?

7.4.2 Estimating Parameters

In the analysis below, we aggregate up to the brand level and estimate the pricing game between the brands. We assume that margins and shares are the

```
> file = paste0(dir, "perth_gas_data.csv")
> dt = fread(file)
> dt[,.(
+   margin = mean(margin),
+   date = mean(date)
+   ),
+          by = c("week", "year")] |>
+   ggplot(aes(x=date, y=margin)) +
+   geom_point(color = "gray") +
+   geom_smooth(se = FALSE) +
+   labs(title = "Margin (cents/liter)",
+        x = "",
+        y = "")
```

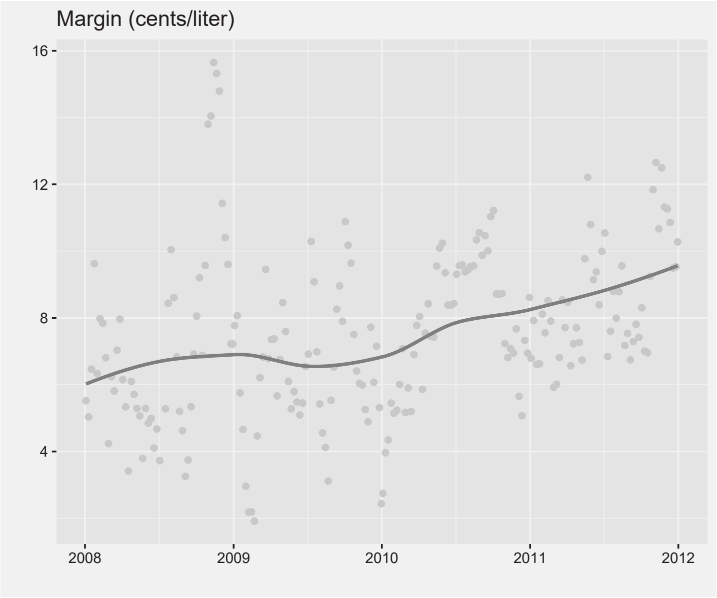


FIGURE 7.1
Plot of weekly margins from 2008, 2009, 2010, and 2011. There is a lot of variation from 2006 to 2010 but prices are moving around the same average. After 2010, the average price starts to increase.

result of Nash equilibrium of the static pricing game. The data provide daily prices, daily wholesale price, the location of the stations and the number of stations for each brand.

To estimate the firm's marginal costs, we regress prices on the distance between the station and the Kwinana terminal, which is located south of Perth.

The assumption is that variation in prices due to distance is determined by the trucking costs of the fuel. We use a quadratic on distance, brand dummies and week dummies to estimate marginal costs for each station. These values are added to the terminal price to get the estimated cost for each station. As we don't have access to quantity information, the brand share is assumed to be equal to the proportion of stations that the brand has.

The model requires estimates of two parameters for each firm. The price sensitivity parameter, α_j, and the unobserved quality parameter, ξ_j. The first is found from the firm's first order condition when demand is determined by a logit model.

$$\alpha_j = \frac{1}{(p_j - c_j)(1 - s_j)} \tag{7.8}$$

The second value is from inverting the logit demand to get the unobserved characteristic as a function of the observed prices, shares, and the parameter α_j.

$$\xi_j = \alpha_j p_j + log(s_j) - log(s_0) \tag{7.9}$$

where s_0 is the outside share.

7.4.3 Parameter Estimates

From Equation (7.8) and (7.9), we determine α and ξ for each firm from the observed prices and shares. We are assuming that the observed prices and shares are determined as the outcome of a static Bertrand Nash equilibrium. We are also assuming that the outside good are the independent stations and the smaller brands, Mobil, Wesco and Better Choice. Finally, we are assuming that Caltex and Caltex Woolworths make pricing decisions as if they are the same firm.

To do this analysis we will aggregate up to annual average prices, costs, margins, and shares.

The first step is to select the variables and the year (2008) we will use. Also redefine some of the brand names to make the analysis easier.

```
> dt2 = dt |>
+   filter(year == 2008) |>
+   select(
+     date,
+     store = TRADING_NAME,
+     brand = BRAND_DESCRIPTION,
+     price = PRODUCT_PRICE,
+     margin
+     )
> dt2$brand[grep("Caltex", dt2$brand)] = "Caltex"
> dt2$brand[which(dt2$brand %in% c("Independent",
+                                  "Mobil",
```

```
+                               "Wesco",
+                               "Better Choice"))] = "Independent"
```

The next step is to calculate the shares by determining the number of stations for each brand and then calculating the share for each brand.

```
> stores = dt2[, .N, by = brand]
> stores$shares = stores$N/sum(stores$N)
> dt2 = merge(dt2, stores, by = c("brand"))
```

The next step creates a data set with prices, margins, shares, and costs averaged up to the brand level for 2008. This also calculates α for each brand.

```
> dt3 = dt2[, .(
+   price = mean(price, na.rm = TRUE),
+   margin = mean(margin, na.rm = TRUE),
+   share = mean(shares, na.rm = TRUE),
+   cost = mean(-margin + price, na.rm = TRUE),
+   alpha = 1/(mean(margin, na.rm = TRUE)*(1 - mean(shares,
+                                               na.rm = TRUE)))
+   ),
+           by = brand]
```

Next, is calculating ξ for each brand. This calculation assumes that the independent stores are the outside option.

```
> index_ind = grep("Independent", dt3$brand)
> dt3$xi = NA
> dt3$xi[-index_ind] =
+   dt3$alpha[-index_ind]*dt3$price[-index_ind] +
+   log(dt3$share[-index_ind]) -
+   log(dt3$share[index_ind])
```

The final step is to use the function `ne_f()` to determine the Nash equilibrium.

```
> a_f = ne_f(dt3$xi[-index_ind],
+            dt3$price[-index_ind],
+            dt3$alpha[-index_ind],
+            dt3$cost[-index_ind])
```

Table 7.2 presents the estimates for α and ξ for each brand. These values reconcile the observed equilibrium margins and the observed equilibrium shares. BP is able to have higher margins and only slightly lower share because its customers are less price sensitive than Caltex.

Can you compare the parameter estimates using the algebraic approach to the numeric approach to determining the Nash equilibrium? Are they exactly the same? Would you expect them to be?

TABLE 7.2
The table presents the prices, margins, market share, and estimates for α and ξ. BP is able to have higher margins and only slightly lower share because its customers are less price sensitive than for Caltex.

	Brand	Price	Cost	Share	α	ξ
1	Ampol	153.70	142.53	0.02	0.09	12.93
2	BP	148.53	140.77	0.23	0.17	26.16
3	Caltex	146.14	139.46	0.30	0.21	32.56
4	Coles Express	144.65	139.92	0.10	0.23	34.28
5	Eagle	152.32	147.32	0.00	0.20	26.47
6	Gull	144.00	138.06	0.11	0.19	27.70
7	Liberty	145.23	138.06	0.03	0.14	19.70
8	Peak	139.98	136.33	0.04	0.29	39.42
9	Shell	152.11	142.57	0.07	0.11	17.20
10	United	138.25	135.42	0.02	0.36	48.59

7.5 Repeated Oligopoly

The standard differentiated goods Bertrand model of price competition suggests that the outcome we would expect, while higher than perfect competition, it is not collusion. In the static game, it is always better for the firms to "cheat" and lower their prices in order to increase profits. The question then is whether we would expect to see collusion when firms interact repeatedly.

The section adapts the static Bertrand game to an infinitely repeated setting and determines the optimal pricing under collusion.

7.5.1 Collusive Equilibrium

Choosing the static equilibrium Bertrand prices in each period can be supported as a subgame perfect Nash equilibrium of the infinitely repeated game. Like with prisoner's dilemma, it is always fine to play Defect every period.

Given this, we ask whether or under what circumstances can a trigger strategy support collusion. Let π_N denote the per-period profits in a Nash equilibrium of the static game, π_C denote the collusive profits and π_D the profits from defecting and choosing an optimal price when the other firm is offering the collusive price. We have $\pi_D > \pi_C > \pi_N$. This is exactly the same as for a prisoner's dilemma.

Like the prisoner's dilemma, collusion can be supported by a trigger strategy if the following inequality holds.

$$\begin{aligned} &\tfrac{\pi_C}{1-r} > \pi_D + r\left(\tfrac{\pi_N}{1-r}\right) \\ &\pi_C > (1-r)\pi_D + r\pi_N \\ &r(\pi_D - \pi_N) > \pi_D - \pi_C \\ &r > \tfrac{\pi_D - \pi_C}{\pi_D - \pi_N} \end{aligned} \tag{7.10}$$

The number on the bottom is larger than the number on the top, and so for a large enough r collusion is a Nash equilibrium of the infinitely repeated oligopoly game.

7.5.2 Identifying Collusion

Identifying collusion for policy or academic purposes is quite different from identifying collusion for criminal prosecution. A criminal case requires hard evidence, not some cool econometric specification. The best evidence includes credible witnesses, audio recordings, video recordings, and written documents. You may be surprised to learn that there isn't that much economics involved prosecuting a criminal collusion case. There is undercover work, wire tapping, etc, but no economics. Once the criminal case is proven, economists are called in to estimate damages.[2] Here again, identifying the collusion is not that difficult. At least it is not that difficult given that the FBI has already completed the task. The case record includes the dates when the collusion occurred and (hopefully) dates when the collusion did not occur. The hard part for the econometrician is working out which part of the difference in prices is due to the collusion and which is due to other changes.

Identifying collusion without the assistance of the FBI is difficult. To observe prices and quantities from a market, there is enough exogenous variation such that we can estimate the elasticity of demand. If the products are differentiated, we can use Bertrand Nash equilibrium to identify marginal costs and mark-ups. While assuming a static equilibrium or a dynamic collusive equilibrium gives different estimates of the mark ups and marginal cost, without some other information we can't tell the difference from the data. If we had data on marginal costs, then suddenly things get a lot easier. We can match the implied marginal costs from our proposed behavioral assumptions and see which ones fits better. Alternatively, if the courts tell us that there was a period where the static equilibrium determined prices, then we can use that period to identify marginal costs. We can then compare the implied markups from the static equilibrium to the observed markups to estimate the **super markups** associated with collusion.[3]

[2]This is usually the amount equal to the consumer surplus lost from the collusion.

[3]These are markups that are larger than we would expect from firms choosing price consistent with equilibrium of a static Bertrand game.

7.5.3 Choosing Super Markups

Choosing the collusive price seems like a simple enough problem for the firms. The collusive price should be equal to the price that a monopoly would charge. Not so fast Sonny Jim! Sure, the monopoly price would be optimal if all the firms behaved as one firm. They are not one firm and they do not behave as such. In each period, each firm may prefer to renege on the deal and choose a lower price, gaining share while everyone else is charging the collusive price.

The collusive price is the solution to the following maximization problem. The super markup (smu) is the difference between the collusive price and the static Nash equilibrium price where the marginal cost is normalized to zero.

$$\begin{array}{ll} \max_{p_c} & \pi_c(p_c) \\ s.t. & \pi_d(p_c) + \frac{r\pi_{NE}}{1-r} \le \frac{\pi_c(p_c)}{1-r} \end{array} \tag{7.11}$$

where $\pi_c(p_c)$ are the per-period profits that one firm makes when everyone chargers p_c, $\pi_d(p_c)$ is the per-period profit when one firm is able to deviate and charge a lower price knowing that everyone else is charging p_c, and π_{NE} is the per-period profit in a static Bertrand Nash equilibrium. The firm's preferences over profits in future periods is determined by the parameter r.

The problem states that firms can charge any collusive price they like as long as all the firms are unwilling to cheat on the agreement. In general, we find that the higher the price, the greater the value in cheating on the agreement.

In general, the colluding firms cannot charge anything they want, in fact they can't even charge the monopoly price. They are constrained both by the demand system and by the incentive of firms to cheat on the agreement.

7.5.4 Estimating Collusive Prices in R

The following functions are used to determine the collusive price when demand is determined by the logit model. The collusive price is assumed to be equal to the Nash equilibrium price plus a **super markup**. The function `f_smu_share()` determines the vector shares at a particular smu. The function `pi_d()` determines the optimal deviation profits given a particular smu. This places a constraint on what smu can be chosen.

```
> f_smu_share = function(smu, xi, price, alpha) {
+   f_share(xi, price + smu, alpha)
+ }
> pi_d = function(p_d, i, smu, xi, price, alpha, cost) {
+   price = price + smu
+   price[i] = p_d
+   share = f_share(xi, price, alpha)
+   return(share[i]*(p_d - cost[i]))
+ }
```

The function `pi_smu()` determines the vector of profits for a particular smu and `pi_smu_int()` is an intermediate function for `optim()`.

```
> pi_smu = function(smu, xi, price, alpha, cost,
+                   r, lambda, pi_ne) {
+   N = length(xi)
+   pi_c = (price + smu - cost)*f_smu_share(smu, xi, price,
+                                           alpha)
+   PI_d = matrix(NA, N, 2)
+   for(i in 1:N) {
+     ai = optimize(f = pi_d,
+                   interval = c(0,2000),
+                   i = i,
+                   smu = smu,
+                   xi = xi,
+                   price = price,
+                   alpha = alpha,
+                   cost = cost,
+                   maximum = TRUE)
+     PI_d[i,1] = i
+     PI_d[i,2] = ai$objective
+     #print(i)
+   }
+   return(pi_c -
+          lambda*(((1 - r)*PI_d[,2] +
+          r*pi_ne - pi_c)^2))
+ }
> pi_smu_int = function(par, xi, price, alpha,
+                       cost, r, pi_ne) {
+   smu = par[1]
+   lambda = 1
+   return(sum(pi_smu(smu, xi, price,
+                     alpha, cost, r,
+                     lambda, pi_ne)))
+ }
```

7.6 Empirical Analysis: Collusion in Perth Gas Stations using R

This section estimates the extent of collusion in the the Perth retail gasoline market.

7.6.1 Super Markups

With all the parameters of the model estimated from the data under the static Bertrand Nash assumption, we use the parameters to determine the super markup.

As above, to determine the super markup, we need to determine the static Nash profits, the profits from cheating, and we need to make an assumption about discounting. We assume `r` is equal to 0.9. What happens at different values?

```
> margin_ne = (a_f$price - dt3$cost[-index_ind])
> share_ne = f_smu_share(0, dt3$xi[-index_ind],
+                        a_f$price, dt3$alpha[-index_ind])
> pi_ne = margin_ne*share_ne
> a1 = optimize(pi_smu_int,
+               c(0,20),
+               xi = dt3$xi[-index_ind],
+               price = a_f$price,
+               alpha = dt3$alpha[-index_ind],
+               cost = dt3$cost[-index_ind],
+               r = 0.9,
+               pi_ne = pi_ne,
+               maximum = TRUE)
> a1$maximum
  [1] 4.589335
> a1$maximum/mean(margin_ne)
  [1] 0.7117553
```

In this set up, the super markup is 4.59 cents a liter or 71 percent of the average static Nash margins. Not bad! If we look at Figure 7.1 we see that margins increased about 4 cents a liter between 2008 and 2012. This suggests that the collusive model is a more accurate representation of pricing behavior than the static pricing model.

7.6.2 Analyzing Mergers with Collusion

Previously we have discussed the effect of mergers when firm competition and pricing can be modeled as a static Bertrand Nash equilibrium. What if the firms prices are really being determined as a collusive agreement? Will the merger affect prices? You may think that the collusive price is as high as prices could get. What impact could the merger have?

Mergers can cause prices to increase when firms are colluding. With collusion, the merger leads to two changes in the market. First, it has an effect on the static Bertrand Nash equilibrium price and thus the punishment that can be imposed. Second, it changes the set of constraints that are placed on the problem. Post merger, there is one less firm that is trying to cheat on the deal.

In theory, these two effects work in different directions. The merger increases profits from the punishment phase, which makes incentivizing firms to collude harder. On the other hand, the number of firms that must be kept in line has been reduced.

Consider a merger between BP and Peak. Peak is very small and so is unlikely to be of concern using standard static Bertrand Nash pricing. To model the effect of the merger, we need to calculate the observed characteristics of the merged firm.

```
> index_m = 9
> dt3$brand[index_m]
  [1] "Peak"
> bp_s = dt3$share[2]/(dt3$share[2] + dt3$share[index_m])
```

The variable `bp_s` is the share of the merged firm that is BP. The following code calculates the new characteristics of the merged firm. The new firm's costs are the weighted average of each firms average costs. Similarly, the new firm's price sensitivity parameter and unobserved characteristics are weighted averages of the two firms. Do these assumptions make sense? What happens under alternative assumptions?

```
> brand_m = dt3$brand[-c(index_ind, index_m)]
> brand_m[2] = paste0(dt3$brand[index_m], " and ", dt3$brand[2])
> xi_m = dt3$xi[-c(index_ind, index_m)]
> alpha_m = dt3$alpha[-c(index_ind, index_m)]
> alpha_m[2] = bp_s*dt3$alpha[2] + (1  - bp_s)*dt3$alpha[index_m]
> price_m = dt3$price[-c(index_ind, index_m)]
> price_m[2] = bp_s*dt3$price[2] + (1 - bp_s)*dt3$price[index_m]
> cost_m = dt3$cost[-c(index_ind, index_m)]
> cost_m[2] = bp_s*dt3$cost[2] + (1 - bp_s)*dt3$cost[index_m]
```

Given the change caused by the merger, we can calculate the new Bertrand Nash equilibrium as well as the new equilibrium collusive price.

```
> a_f_m = ne_f(xi_m, price_m, alpha_m,
+            cost_m, maxiter=1000000)
> margin_m = (a_f_m$price - cost_m)
> share_m = f_smu_share(0, xi_m, a_f_m$price, alpha_m)
> pi_ne_m = margin_m*share_m
> a2 = optimize(pi_smu_int,
+               c(0,20),
+               xi = xi_m,
+               price = a_f_m$price,
+               alpha = alpha_m,
+               cost = cost_m,
+               r = 0.9,
```

```
+                pi_ne = pi_ne_m,
+                maximum = TRUE)
```

If the firms are playing Bertrand static game, then merger has the following impact on prices. Prices will increase 0.4 percent.

```
> (mean(a_f_m$price) - mean(a_f$price))/mean(a_f$price)
  [1] 0.004122352
```

If the firms are playing a collusive game, then the merger has the following impact on prices. Prices will increase 1.2 percent.

```
> (mean(a_f_m$price+a2$maximum) -
+     mean(a_f$price+a1$maximum))/mean(a_f$price+a1$maximum)
  [1] 0.01238215


> sum_tab_m = cbind(
+   as.numeric(a_f$price),
+   as.numeric(c(a_f_m$price[1:7], NA, a_f_m$price[8:9])),
+   as.numeric(a_f$price) + a1$maximum,
+   as.numeric(c(a_f_m$price[1:7], NA, a_f_m$price[8:9]))
+ a2$maximum
+ )
> rownames(sum_tab_m) = dt3$brand[-index_ind]
> colnames(sum_tab_m) = c("Price", "Price Merge",
+                         "Collusive", "Collusive Merge")
```

Table 7.3 presents the impact of the merger on prices under the two different models. Note that while BP's price goes down with the merger, the average price of BP and Peak increases. Remember the new firm is a weighted average of BP and Peak's observed and unobserved characteristics.[4] The impact of the merger is quite different under the two models of pricing behavior. The average price effect of the merger is small. It is 0.4% increase in prices in the static Bertrand case and 1.2% increase in the collusive case. The difference between the price increase assuming static Bertrand and collusion is over 100%. The assumed method by which firm's determine prices may have a large effect on our prediction of the merger.

7.7 Discussion and Further Reading

In the 1990s, industrial organization economists made great gains in understanding markets and competition. But there was something missing. The

[4] What may be other assumptions that could be made about the new post-merger firm?

TABLE 7.3
The table presents impact of the merger under the two pricing models. The first and third column are the pre-merger prices for the two models. The second and fourth columns are the new prices after the simulated merger between BP and Peak. Note that the new prices are given for BP. The impact of the merger is quite different under the two pricing models. For example, Shell's prices don't really change with the merger in the Bertrand static pricing game, while they increase substantially in the repeated game.

	Price	Price Merge	Collusive	Collusive Merge
Ampol	153.70	153.81	158.29	159.67
BP	148.53	145.77	153.12	151.62
Caltex	146.14	146.84	150.72	152.70
Coles Express	144.65	144.85	149.24	150.70
Eagle	152.32	152.33	156.91	158.18
Gull	144.00	144.27	148.59	150.13
Liberty	145.23	145.31	149.82	151.17
Peak	139.98		144.57	
Shell	152.11	152.41	156.70	158.26
United	138.25	138.28	142.84	144.14

standard models didn't always capture the pricing behavior. A number of recent papers have shown that we need better models in order to capture the possibility of collusion in markets. Repeated games allow us to use much richer models of pricing behavior and show how collusion can be a more natural feature of how firms determine prices.

The paper by David Byrne and Nic de Roos does not provide us with any model techniques. Rather, the paper takes a very close look at an actual market and shows how firms in that market actually behave (Byrne and de Roos, 2019). Nathan Miller and Matt Weinberg review the joint venture between Miller and Coors. The authors use information on the premerger market to estimate parameters of the model. They show that the observed price increase post merger cannot be explained by our standard static Nash model (Miller and Weinberg, 2017). Nathan Miller, Gloria Sheu, and Matt Weinberg argue that a model that explicitly accounts for collusion is necessary to predict the effect of mergers in some industries (Miller et al., 2021).

A number of papers formally compare the predictions of static pricing models to observed pricing behavior. Miller and Weinberg (2017) use pre-merger pricing to estimate the parameters, then compare simulated pricing to actual pricing after the merger. Nevo (2001) and Backus et al (2021) compare the predictions of collusive pricing models to actual pricing in the ready-to-eat cereal market.

The pricing patterns seen in retail gasoline is very strange (Lewis, 2012). In the Perth data, we see the firms move to a very ordered pattern of pricing on a weekly basis. If you zoom in even closer you see that the brands are using one or two stations to signal which price they will move to for that week. Our analysis in this chapter is based on the work of DOJ Economist, Zhongmin Wang. Wang (2009) uses a mixed strategy dynamic model of to analyze pricing patterns in the Perth data. Nobel prizing winning economists, Eric Maskin and Jean Tirole, show that short term commitment to a price is necessary to get the type of pricing dynamics we see in the data (Maskin and Tirole, 1988). The West Australian state government introduced price regulation that allowed for this type of equilibrium. We generally call it “post and hold” regulation. The post means that the price is publicly displayed and the hold means that the price cannot be changed for some period of time, say 24 hours. By making the price publicly available, it allowed for strategies that are a function of each other’s prices. Hold means that the firms can commit to the price which means that the equilibrium cannot devolve into a competitive pricing process. That is, each firm will choose a price to under cut its competitor, like in the original Bertrand game presented in Chapter 3.

8

Bargaining

8.1 Introduction

When you take your first economics class you learn about two price setting mechanisms, perfect competition and monopoly. You learn about the case where no seller in the market has any power to determine the price and the case where the only seller has the power to determine the price. You may have also been exposed to the idea that the buyer has the power to determine the price, monopsony. What about the case when both the buyer and the seller have the power to determine the price? What happens in that case?

In the case where both sides of the transaction have the ability to determine the price, we need a bargaining model. We need a model that can help determine what the outcome will be. This chapter considers three models of bargaining, the **ultimatum game**, the **alternating offer game** and the **Nash bargaining model**.

While the **ultimatum game** is very simple, our predictions of what will happen in the game are both unsatisfactory and do not seem to agree with what happens when actual people play the game. The chapter goes through the various predictions for this game and then looks at what happens in experiments where real people play the game. In order to develop a game that predicts outcomes that seem more reasonable the chapter presents the **alternating offers game**. The chapter shows that while this game is quite complicated, its predictions are quite intuitive. The third model of bargaining is not technically a game. That said, it can be shown that the Nash bargaining model is equivalent to certain alternating offers games. So, although the model itself is not game theoretic, it does have a game theoretic foundation. More importantly, the Nash bargaining model is much simpler than the alternating offers game. This simplicity has allowed the model to become an important tool in empirical analysis of situations where bargaining is used to determine prices.

The chapter uses the Nash bargaining model to understand competition between hospitals in Florida. The empirical estimates are used to predict the effect on hospital prices of a merger between hospitals located in Palm Beach County, the home county for this book's publisher.

DOI: 10.1201/9781003351603-8

8.2 Ultimatum Game

The simplest way to model a bargaining game is with a "take-it-or-leave-it" offer. A TIOLI offer if you will. The game has two periods. In the first period, one player makes an offer on how to split the pie between the two players.[1] In the second period, the other player observes the offer and decides to accept or reject. If the second player accepts, then the two players split the pie as determined by the first player's offer. If the second player rejects the offer, then the pie vanishes and neither players gets anything.

What would you do if you are the first player? How much would you give to the second player? How much would you keep for yourself? What if the pie is \$5.00? Would you give the other player \$2.50? \$2.00? \$1.00? Whoa. \$0.00? How about if the pie is \$5,000,000.00? What if you are the second player? Would you reject an offer that you believe is unreasonably low? Would you reject an offer of \$2.00? What if the pie is \$5,000,000.00, would you reject the \$2.00 offer in that case?

This section considers the prediction when we use **Nash equilibrium** and when we use **subgame perfect Nash equilibrium**. It then brings in some data from experiments where actual people play the game and compares the observed outcomes to the predicted outcomes.

8.2.1 The Game

We can write down the ultimatum game using a formal representation. Odd's strategy is simple. She offers some number between zero and one, including zero or one. This number represents the share of the pie that Odd keeps if the offer is accepted. Even's strategy is more complicated. Even will accept certain offers and will reject other offers. The notation below is completely general. It includes strategies like reject all offers and accept all offers and accept all offers that are multiples of 0.1345888, etc.

- Players: Odd, Even
- Strategies:
  - Odd: Offer $x \in [0, 1]$
  - Even: If $x \in R$ reject, otherwise accept, where $R \subset [0, 1]$.
- Payoffs:
  - Odd offers x, Even accepts: $\{x, 1 - x\}$
  - Odd offers x, Even rejects: $\{0, 0\}$

[1] The term "pie" is used to refer to some outcome that can be divided up between the players. In many cases, this is some fixed amount of money.

where Odd gets the payoff associated with the first element of the set. Odd's strategy is denoted as x, which is the share of the pie that Odd receives. The remainder, $1 - x$, is the share received by Even. Even rejects Odd's offer if x lies in a particular subset of the offers called R, which stands for reject. To make things simpler, we will generally assume that Even plays a "cutoff" strategy. That is R is equivalent to the set x where $x \geq y$, where $y \in [0, 1]$. If x is large, then Even's share is small and he will reject the offer.

8.2.2 Game Tree

While this game is not that complicated, it turns out that presenting it in extensive form is very useful.

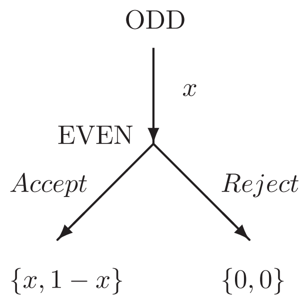


FIGURE 8.1
An extensive form representation of an ultimatum game. Odd makes an offer x which is the share of the pie that they will keep. Even observes the offer and decides whether or not to accept or reject the offer.

Figure 8.1 presents the extensive form representation of the ultimatum game. In the first period, Odd makes an offer of x which is some value between 0 and 1. It is the share of the pie that Odd will keep if Even accepts the offer. After observing Odd's offer, Even decides whether to accept or reject the offer. If Even accepts, then he gets $1 - x$. If he rejects, then no one gets nothing.

8.2.3 Normal Form

As mentioned above, while we generally use extensive form representations for dynamic games, there is nothing to stop us from using a normal form representation. In order to think about some of the predicted outcomes of this game, it is useful to think about the normal form representation.

We will simplify and just consider three strategies by Odd, $x \in \{0, 0.5, 1\}$. We will also limit Even to three strategies. We will further limit Even to cutoff strategies. That is, Even will choose some number such that he will accept any offer less than or equal to that number and reject any offer above that number. We will denote the cutoff value y. So $y = 0$ means that Even rejects any offer that gives him less than the full share. While $y = 0.5$ means that Even will accept an offer that gives him at least half. Lastly, $y = 1$ means that Even accepts any offer that Odd makes.

TABLE 8.1
Normal form representation of simplified ultimatum game. Odd's offer is x which is the share Odd receives, with $1 - x$ going to Even. Even's strategy is denoted y, and it gives the level of the offer that she will accept. The payoffs are in the cells with Odd first.

Odd, Even	$y = 0$	$y = 0.5$	$y = 1$
$x = 0$	0, 1	0, 1	0, 1
$x = 0.5$	0, 0	0.5, 0.5	0.5, 0.5
$x = 1$	0, 0	0, 0	1, 0

Table 8.1 presents a normal form representation of a simplified version of the ultimatum game. What is the Nash equilibrium of the game? Is $x = 0.5, y = 0.5$ a Nash equilibrium? What about $x = 0, y = 0$ or $x = 1, y = 1$?

8.2.4 Nash Equilibrium

This is a pretty simple game. Odd makes an offer, Even observes the offer and decides to reject or accept the offer. What do you think is the likely outcome of the game? Do you think $x = 0.5$ in equilibrium? Do you think there will be an even split? Do you think $x = 1$? Do you think Odd will make an offer where she gets to keep the whole pie? What would Even do if he saw such an offer?

Is $x = 0.5$ a Nash equilibrium? It is if we are a bit more careful in what it is. Remember Even's strategy needs to state what will happen in every possible case, that is, every possible offer Odd could make. Consider the following set of strategies. Odd offers $x = 0.5$ and Even accepts any offer below 0.5 (or equal) and rejects any offer above 0.5. This is a Nash equilibrium.

To see that it is a Nash equilibrium, let's go back to our algorithm. In the algorithm, we first assume Odd plays her strategy. We then determine if it is optimal for Even to play the strategy stated in the proposed equilibrium. If it is, we reverse things and assume Even plays the strategy stated in the proposed equilibrium. Given that strategy, we determine if Odd's optimal strategy is the same as in the proposed equilibrium. If it is, then we have a Nash equilibrium!

The proposed equilibrium is $\{x = 0.5, y = 0.5\}$

- Assume Odd offers $x = 0.5$, what is Even's optimal response?
  - Even's Payoffs for his different strategies:
    * $y = 0.5$: 0.5
    * $y < 0.5$: 0
    * $y > 0.5$: 0.5
  - Even can't be made better off with $y \neq 0.5$, so $y = 0.5$ is optimal.
- Assume Even's cutoff is $y = 0.5$, what is Odd's optimal response?
  - Odd's Payoffs from her strategies:
    * $x = 0.5$: 0.5
    * $x < 0.5$: x (which is less than 0.5)
    * $x > 0.5$: 0
  - Odd is not better off choosing some other offer, so Odd's strategy is optimal and the proposed set of strategies is a Nash equilibrium.

The offer that splits the difference can be part of a Nash equilibrium. Can any split be supported as part of a Nash equilibrium? Yes. All of them.

Any offer by Odd can be supported as a Nash equilibrium of the game. Consider the case where Odd offers a where a is between 0 and 1. For this case also assume that Even's strategy is for $y = a$. That is, Even accepts any offer less than a (or equal to) and rejects any offer above a.

- Assume Odd Offers $x = a$, what is Even's optimal response?
  - Even's payoffs from his strategies:
    * $y = a$: $1 - a$
    * $y < a$: 0
    * $y > a$: $1 - a$
  - Even can't be made better off with $y \neq a$.
- Assume Even will accept any offer below a ($y = a$), what is Odd's optimal response?
  - Odd's payoffs from her strategies:
    * $x = a$: a
    * $x < a$: x
    * $x > a$: 0
  - Odd is not better off choosing some other offer.

There is a set of Nash equilibrium of the ultimatum game such that $\{x = a, y = a\}$, where x represents Odd's offer and y represents Even's cutoff. Even will accept any offer below y and reject any offer above y.

So Nash equilibrium predicts any outcome. That doesn't seem particularly useful nor does it seem intuitive. Do you think Even would really have a strategy that says he will accept any offer below a 10 percent share of the pie?

8.2.5 Subgame Perfection

So Nash equilibrium predicts any thing can happen. What about **subgame perfection**? Remember the definition. An outcome is subgame perfect if it is a Nash equilibrium of every subgame. Our ultimatum game has two (sort of) subgames. There is a subgame after Odd makes the offer x and Even observes the offer. This subgame has one player, Even, and the strategy is just the action accept or reject. The other subgame is the whole game.

What is the Nash equilibrium of the first subgame? Let Odd offer x. Even is best off accepting that offer. If Even accepts he gets $1 - x$, while if he rejects he gets 0. Even is always at least weakly better off accepting Odd's offer.

So going back to the whole game the only strategy of Even's that is subgame perfect is for $y = 1$. That is Even accepts any offer Odd makes. Given that Even will accept any offer, Odd is best off offering $x = 1$.

So subgame perfection substantially reduces the number of outcomes that are an equilibrium. Now we just have one. It is for Odd to get the whole pie! Does that seem reasonable?

8.3 Empirical Analysis: Ultimatum Game in India using R

The ultimatum game may be one of the most studied games in economic experiments. How people play the ultimatum game give us an interesting view into different cultures. What do people think is a fair split? Will people reject splits that they do not think is fair even if it means giving up real money? Do the stakes matter? In particular, do the so called "fair" offers go away when the players are playing with real money? Is subgame perfection actually a better prediction than you thought?

The section analyzes data from an experiment conducted in northern India where the experimenters vary the stakes.

8.3.1 Data

These data are replication data for Andersen et al. (2011). The authors went into a very poor northern Indian village and ran an experiment where villagers played for real money. In terms of hours worked equivalent, the size of the pie varied from \$30 to \$48,000 (assuming \$20 per hour). So yes, real money.

Figure 8.2 presents the density of the offer percentages. These are the amounts that Even receives if he accepts Odd's offer. The modal offer is the 50-50 split and it is actually very rare to offer more than that! The proposer seems to be unwilling to give the responder more than half the pie. That said, there is quite a lot of density less than 50–50 and even some that is pretty low.

```
> file = paste0(dir, "20100982_DATA.dta")
> read.dta(file) |>
+   ggplot(aes(x = percent_offer)) +
+   geom_density(fill = "gray", alpha = 0.5) +
+   geom_vline(xintercept = 0.5, col = "gray") +
+   labs(title = "Density of Offer Shares",
+        x = "Offer Share Responder Receives",
+        y = "") +
+   # remove y-axis
+   theme(axis.text.y=element_blank(),
+         axis.ticks.y=element_blank())
```

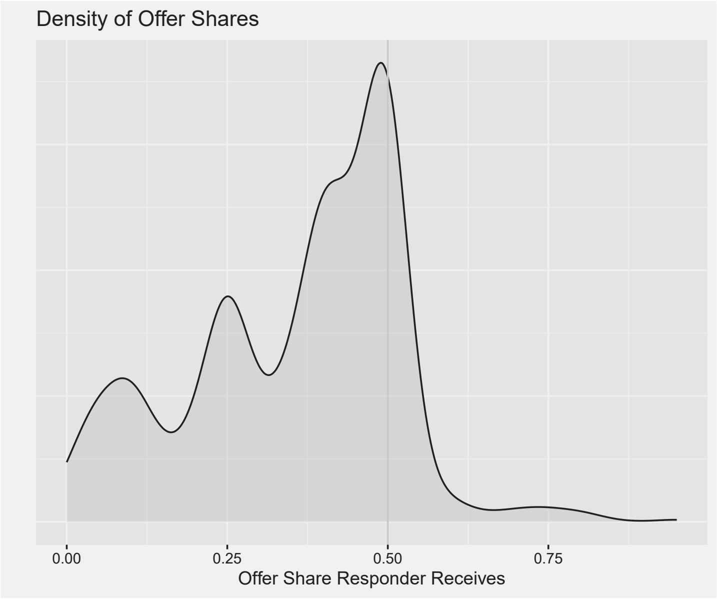


FIGURE 8.2
Density plot of share of the pie offered by the proposer (Odd) to the responder (Even). The bulk of the offers are just below half of the pie going to the responder.

Some responders make pretty lower offers.

8.3.2 Equilibrium Play or Fairness?

We see in Figure 8.2 that a large number of offers are around the 50% mark. Is this just the proposers playing fair or is it part of the equilibrium behavior?

That is, do the proposers believe that if they make an offer below the 50–50 split, it will be rejected?

While we observe the proposer's strategy, we don't observe the strategy for the responder. We could estimate it. If we assume all responders are the same, then we can estimate a model of their strategy.

Assume that we can represent the responder's strategy using a logit. We use `glm()` to estimate the logit. We can determine the probability of accepting the offer as a function of the percent of the pie offered. In order to plot it out, we can determine the predicted probability of accepting the offer for each actual offer we observe in the data.

```
> file = paste0(dir, "20100982_DATA.dta")
> data = read.dta(file)
> glm1 = glm(accept ~ percent_offer,
+             data = data,
+             family = binomial(link = "logit"))
> data = data |>
+   mutate(
+     accept_pred = predict.glm(glm1, type = c("response"))
+   )
```

Figure 8.3 looks nothing like what we would expect an equilibrium strategy to look like. If the strategy is Reject if $x < 0.5$ and Accept if $x \geq 0.5$, then we would expect an S-shape, going to 0 when offers are close to 0 and close to 1 when offers are close to 1, with a cross around 0.50. These players are accepting offers that are much lower than the equilibrium predicts. Similarly, it is not consistent with subgame perfection. In that case, we would expect something like a straight horizontal line at 1.00. There are many more rejections than we would expect if the strategies were part of a subgame perfect Nash equilibrium.

8.3.3 Do Stakes Matter?

What do things look like for a subset of the data where there are very large stakes? Let's restrict the data to when the stakes are 20,000 rupees. In this case, the average offer is just 12 percent of the pie and for the average person the pie is worth 24 days of work.

We can filter the data to only include the large stakes games and then estimate the predicted acceptance rate for these games.

```
> data = data |>
+   filter(
+     stakes_4 == 1
+   )
> glm2 = glm(accept ~ percent_offer,
+             data = data,
+             family = binomial(link = "logit"))
```

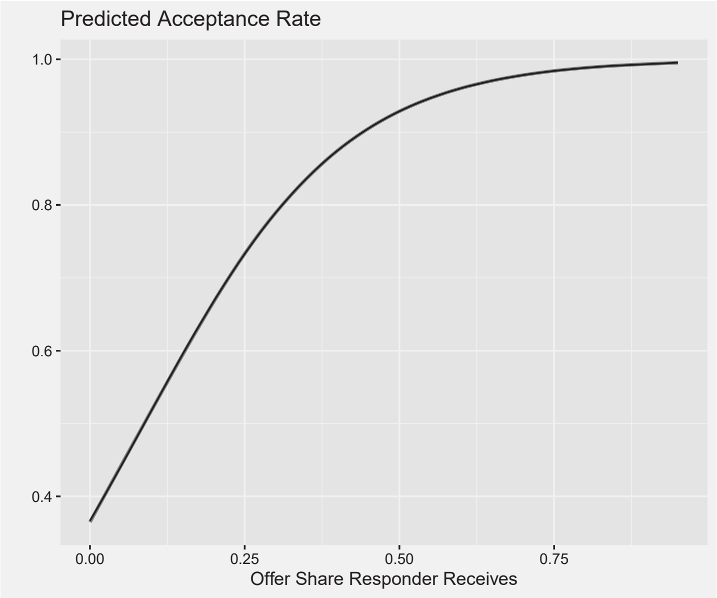


FIGURE 8.3
Plot of predicted acceptance rate as a function of the actual offer shares. The predicted acceptance percentage grows to approach 100 as the offer gets closer to 50 percent.

```
> data = data |>
+     mutate(
+       accept_pred = 7.5*predict.glm(glm2, type = c("response"))
+     )
```

Here is the code to create a `ggplot()` object that shows the density of offers and the predicted acceptance rate for the subset of experiments with large stakes.

```
> ggplot_pred_accept_hs = data |>
+   ggplot() +
+   geom_density(aes(x = percent_offer),
+                fill = "gray",
+                alpha = 0.5) +
+   geom_smooth(aes(x = percent_offer,
+                   y = accept_pred),
+               se = FALSE) +
+   scale_y_continuous(breaks = seq(0, 7.5, by = 7.5/2),
+                      labels = seq(0, 100, by = 50)) +
+   geom_text(aes(x = 0.2, y = 9,
```

```
+                  label = "Predicted Acceptance Rate (percent)"),
+              color = "gray") +
+   geom_text(aes(x = 0.4, y = 2,
+                  label = "Density of Offer Shares"),
+              color = "gray") +
+   labs(title = "",
+        x = "Offer Share to Responder",
+        y = "")

> ggplot_pred_accept_hs
```

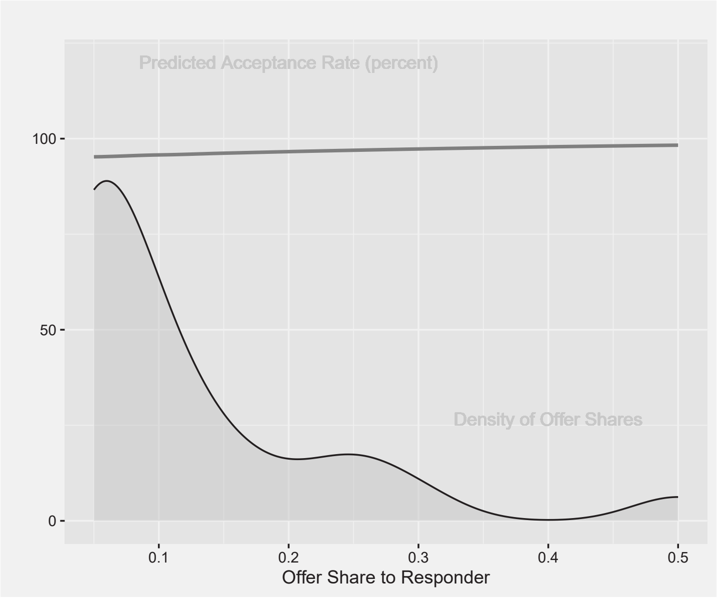


FIGURE 8.4
Density plot of offer shares and predicted acceptance rate for the subset of experiments with large stakes. The proposers makes offers that tend to be close to zero. The receiver's predicted acceptance rate is close to 100 percent.

Figure 8.4 presents the density of offers and the predicted acceptance rates. This graph is consistent with the subgame perfect Nash equilibrium. The weight of offers is 0, and the predicted acceptance rate is consistent with a strategy of accepting any offer.

Do stakes matter? Yes. Apparently they do. When they do, subgame perfection predicts the likely outcome of the game. Is that what you would have thought?

8.4 Two Period Alternating Offers Game

It is nice that the very strong prediction of subgame perfection in the ultimatum game can be born out in some actual games. But it is not that satisfying. Most of the results of the experiment are not consistent with either subgame perfection or Nash equilibrium. Nor are the predictions of the game consistent with our intuition. Why is the 50–50 split so dominant? It is only one of many predictions of the game.

If the simple model is doing a poor job of predicting outcomes of interest, then one solution is a more complicated slash realistic model. This section considers what happens when the model is made more realistic slash more complicated. An alternating offers game is one in which the two players take it in turns to make the offer, where the game only ends if the offer is accepted.

The section presents the game, the extensive form representation and finds the subgame perfect Nash equilibrium.

8.4.1 The Game

We formally represent the game with some compact notation. Compact is code for confusing. As before x_t refers to the share that Odd gets, while Even gets $1 - x_t$. The $t \in \{1, 2\}$ refers to which period we are in, the first or second. The y_t refers to the cutoff strategy. It's exact meaning depends on who the responder is. If the responder is Even (y_1) means that Even will accept any offer where Odd gets less than y_1.

In the second period, Odd and Even switch roles. The x_2 is still the amount that Odd gets, but it is Even that is making the proposal. This may be a function of the offer made by Odd in the first period. It is also a function of whether the offer was accepted or rejected but given that we only see this strategy if it was rejected we can ignore that part of the history. The y_2 refer to Odd's cutoff strategy. This means that Odd will REJECT any offer less than y_2. Remember we are always talking about the share that Odd gets. This is also a function of Odd's offer in period 1.

- Players: Odd, Even
- Strategies:
  - Odd: $\{x_1, y_2(x_1)\}$
  - Even: $\{y_1, x_2(x_1)\}$
- Payoffs
  - $x_1 < y_1$: $\{x_1, 1 - x_1\}$
  - $x_1 > y_1$, $x_2 > y_2$: $\{rx_2, r(1 - x_2)\}$

– $x_1 > y_1$, $x_2 < y_2$: $\{0, 0\}$

where x_t is the offer in period t and the amount received by Odd, y_t is the cutoff amount in period t.

The probability r is meant to capture the possibility that the parties risk failure by extending the negotiations. We will see that the size of this risk bears heavily on the negotiated outcome.

8.4.2 Game Tree

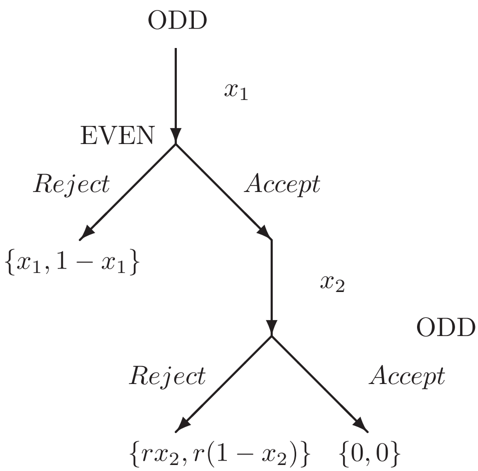


FIGURE 8.5
Two period alternating offers game. Odd makes first offer of x_1. If rejected, Even makes an offer of x_2. With probability r the game continues. Odd's payoff is listed first in brackets.

Figure 8.5 presents the extensive form representation of the two-period alternating offers game. The size of the pie decreases from the first period to the second period. The amount of the decrease is determined by r.

8.4.3 Subgame Perfection

The second period is an ultimatum game with Even as the proposer and Odd as the responder. Given this, $y_2 = 0$ and $x_2 = 0$. That is, Odd will accept any offer and Even will offer 0.

Moving back to the first period, we can take the second period equilibrium as given. If Even accepts, then he gets $1 - x_1$ and if he rejects he gets $r(1 - x_2) = r$, where r is the probability that the second period occurs. Given this, Odd will make an offer such that $1 - x_1 = r$ or $x_1 = 1 - r$.

In this case, the subgame perfect equilibrium is

$$x_1 = 1 - r, y_2 = 0, y_1 = 1 - r, x_2 = 0 \tag{8.1}$$

The outcome we actually see is an offer of $x_1 = 1 - r$. So having this probability that the game ends prior to the second period changes the outcome of the game. Instead of Odd getting the whole pie, they get the pie less a portion equal to the probability that the game continues to the next period.

8.5 Infinite Alternating Offers Game

While we see some evidence that the subgame perfect Nash equilibrium does occur when real people play the ultimatum game, it is not clear that real people actually play the ultimatum game. Maybe there is some other game that more accurately represents what is happening when people bargain.

We saw in the previous section that adding both a second period and the probability that the game ends prior to the second period changes the outcomes. What happens if we add even more periods? What happens if we add an infinite number of periods?

In acknowledgement of the seminal contribution by Israeli economist, Ariel Rubinstein, we generally refer to this as the **Rubinstein bargaining model**.

The section presents the game, the extensive form representation, finds the subgame perfect Nash equilibrium and presents an algorithm for finding that outcome.

8.5.1 The Game

The game is as before, but now with an infinite number of periods. In each period, there is a proposer and a responder. If the period is odd, then the proposer is Odd, while the proposer is Even if the period is even. As before, the responder can either accept or reject the offer. If the responder accepts, the game ends and the players get the payoffs $\{x_t, 1 - x_t\}$. If the responder rejects the offer, then the game goes to the next period and the proposer and responder swap roles.

Similar to the previous game, the size of the pie changes over time. The parameter $r \in (0, 1)$ represents the change in the size of the pie from period to period. We can think of it as standard financial discount rate. Alternatively, we could think about it as representing the probability that the game will continue.

- Players: Odd, Even
- Strategies:

- Odd
  * If the time period t is odd, then given the history of offers up to time period $t-1$, offer x_t to Even.
  * If the time period t is even, then accept or reject Even's offer based on x_t and the history of offers up to $t-1$.
- Even
  * If the time period t is even, then given the history of offers up to time period $t-1$, offer x_t to Odd.
  * If the time period t is odd, then accept or reject Odd's offer based on x_t and the history of offers up to $t-1$.

• Payoffs

- If at time t the an offer x_t is accepted:
  * Odd: $0, 0, 0, ..., r^t x_t$
  * Even: $0, 0, 0, ..., r^t(1-x_t)$
- If at time t the offer x_t is rejected and the offer x_s is accepted in periods $s > t$.
  * Odd: $0, 0, 0, ..., 0, ..., r^s x_s$
  * Even: $0, 0, 0, ..., 0, ..., r^s(1-x_s)$

8.5.2 Game Tree

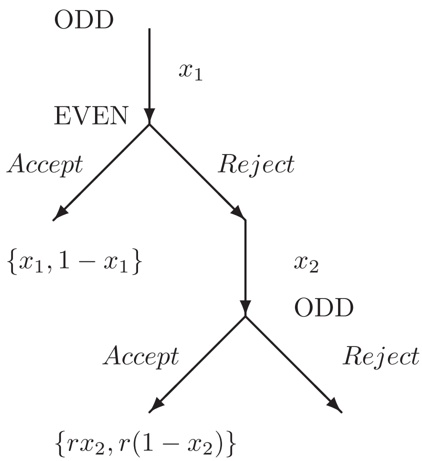


FIGURE 8.6
The first two periods of an infinite period alternating offers game.

Figure 8.6 presents the first two periods of an infinite period alternating offers game. In each period, the size of the pie decreases in proportion to the discount rate r.

8.5.3 Subgame Perfection

So we know how to solve for the subgame perfect Nash equilibrium. Simply go to the last period, work out the equilibrium for that game and then work your way backwards. OK, but what if there is no last period? What if the game has an infinite number of periods?

The standard solution is to approximate our infinite period game with a finite period game. Consider a game that ends at period T (assume odd).

The last period is an ultimatum game where Odd offers x_T and Even observes the offer and chooses whether to accept or reject. In the subgame where Odd has made the offer of x_T, Even's payoffs are

- Accept: $1 - x_T$
- Reject: 0

Even will accept any offer where $1 - x_T \geq 0$ (indifference assume Even accepts). Working backwards, Odd will choose the x_T that is as small as possible, Odd offers $x_T = 1$. If the game gets to period T the payoffs are $\{1, 0\}$.

Now let's do $T - 1$. This is an ultimatum game where Even makes the offer and Odd chooses whether or not to accept or reject. Even offers Odd x_{T-1}. Odd's payoffs are:

- Accept: x_{T-1}
- Reject: $r \times 1$

where r is how much Odd discounts the future. If Odd rejects, she gets nothing immediately, but in one period she knows that she will get the whole pie of 1. But that pie gets discounted in proportion to r.

Odd's best response is to accept any offer such that $x_{T-1} \geq r$ and reject otherwise. Even knows this and wants to make x_{T-1} as small as possible but still have Odd accept the offer. That is where $x_{T-1} = r$. If the game gets to $T - 1$, then the payoffs are $\{r, 1 - r\}$.

Now consider $T - 2$. This an ultimatum game where Odd makes an offer of x_{T-2} to Even and Even decides to accept or reject. Even's payoffs are

- Accept: $1 - x_{T-2}$
- Reject: $r \times (1 - r) = r - r^2$.

Therefore, Even will accept any offer $1 - x_{T-2} \geq r - r^2$. Odd wants to make x_{T-2} as large as possible, so they will choose $x_{T-2} = 1 - r + r^2$. If the game gets to $T - 2$ then the payoffs are $\{1 - r + r^2, r - r^2\}$.

If we let $r = 0.9$, then the payoffs if the game gets to T are $\{1, 0\}$, if it gets to $T - 1$ they are $\{0.9, 0.1\}$ and if it gets to $T - 2$ the payoffs are $\{0.91, 0.09\}$

What happens in $T - 3$?

8.5.4 Game Ends in Period 1

Working all the way back to Period 1, Odd makes an offer x_1 and Even decides to Accept or Reject. In the game with T periods, the offer of x_1 will be the amazingly complicated thing with lots of rs. However, as T gets very large x_1 converges to 0.5. That is, the subgame perfect Nash equilibrium of the game is a 50-50 split! It is a super complicated game that makes a very simple and intuitive prediction.

8.5.5 Infinite Alternating Offers Game in R

We can use the computer to analyze more complicated games than what we looked at above. We can allow the two players to have different beliefs about when the game is going to end and different payoffs if the parties fail to reach a bargain.

Odd makes an offer at time t, assume that the next period the game ends in agreement and Even gets a payoff of $(1 - x_{t+1})V_A$, where V_A is the size of the pie if the offer is accepted. Also assume that Even discounts the future by r_E. In addition, assume that if there is no agreement, this period Even gets v_{EN}.

By having different discount rates and non-agreement values, we can get different bargaining outcomes. You can see that what looks to be small differences lead to quite large differences in the bargaining outcomes.

What offer should Odd make?

Even will accept Odd's offer if and only if

$$1 - x_t \geq v_{EN} + r_E(1 - x_{t+1})V_A \tag{8.2}$$

Assume that the offer is such that the payoff for the responder makes them indifferent between accepting or rejecting the offer. In the function below `x` is the proportion of the pie received by the responder, `V_A` is the size of the pie, `v_N` is the period amount the responder gets if they do not accept the offer and `r` is the discount rate.

```
> offer = function(x, V_A, v_N, r) {
+   v_N + r*x*V_A
+ }
> T = 250
> r_odd = 0.99
> r_even = 0.98
> V_A = 1
> v_EN = 0
> v_ON = 0.002
```

Given the parameter values above the following loop determines the equilibrium of the game.

```
> odd_offers = rep(NA, T)
> even_offers = rep(NA, T)
> odd_offer_old = 0
> for (i in 1:T) {
+   even_offer_old = offer(1 - odd_offer_old, V_A, v_ON, r_odd)
+   odd_offer_old = offer(1 - even_offer_old, V_A, v_EN, r_even)
+   odd_offers[i] = odd_offer_old
+   even_offers[i] = even_offer_old
+   #print(odd_offer_old)
+ }
```

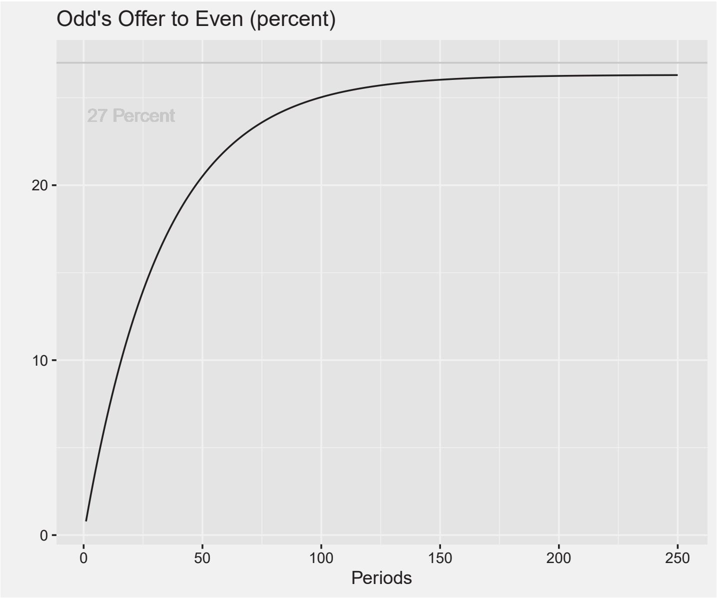


FIGURE 8.7
Line chart of offers to Even as the number of periods gets large. It shows that in equilibrium the offer to Even approaches 27 percent.

Figure 8.7 shows that the equilibrium offer to Even converges to 0.27, with Odd receiving 0.73 of the pie. Why is this split not even? Why is it in favor of Odd? What change could you make to get it more of an even split?

8.6 Nash Bargaining Model

John Nash developed one of the most important ideas in modern game theory - every (finite) game has what we now call a Nash equilibrium. It has an outcome which is "stable" in the sense that no player would want to deviate from that strategy if they knew which strategies all the other players were playing.

Nash was interested in another problem, how to determine the outcome when entities bargain. While the Nash equilibrium became the keystone concept in non-cooperative game theory, Nash himself was not able to work out how to model bargaining as a non-cooperative game. Instead he developed an alternative framework for analyzing bargaining problems known as the Nash bargaining model. The parameterization of the model presented below gives same split of the pie as the infinite period alternating offers game above. This suggests an equivalence between the two models.

The section presents the Nash bargaining model and shows how it is used to analyze competition between hospitals.

8.6.1 The Model

Consider a game where we have two players, say a hospital and an insurance company. The two firms are bargaining over how to pay the hospital for various services that the hospital provides to the insurance companies beneficiaries. The price that the hospital will receive depends on two sets of things. First, it depends upon what both the hospital and the insurance company get if negotiations break down. If there is only one hospital in an area, then the insurance company is not going to be able offer its beneficiaries much of a product if it can't come to a deal with the hospital. If most of the people in the area work for the same firm and are covered by the same insurance then demand for the hospital will drop dramatically if the hospital can't come to a deal with the insurance company. Second, it depends on how good each side is at bargaining, which we will conceptualize as the relative "bargaining weights." These are somewhat amorphous. Practically, these weights are often set to be equal. Below we look at how these weights relate to the rs used in the alternating offers game presented earlier.

$$\max_x \quad (x-a)^{\lambda}(1-x-b)^{1-\lambda} \tag{8.3}$$

where a and b are the alternative outcomes (the outcome if bargaining fails) for Odd and Even, respectively, x is the share of the pie that goes to Odd and λ is the bargaining weight.

Taking first-order conditions and simplifying.

$$\begin{array}{l}\lambda(x-a)^{\lambda-1}(1-x-b)^{1-\lambda} - (1-\lambda)(x-a)^{\lambda}(1-x-b)^{-\lambda} = 0\\ \lambda(1-x-b) - (x-a)(1-\lambda) = 0\\ x = \lambda(1-b) + (1-\lambda)a\end{array} \tag{8.4}$$

We see that Odd's share is increasing in the size of the alternative and decreasing in the size of Even's alternative. The size of Odd's share depends on how much Odd has to lose. If Odd's alternative is good, a is large, then Even will need to offer Odd more to have her accept the bargain.

Definition 14. *A* Nash bargaining model *is an algorithm for determining the outcome from bargaining based on the player's payoffs from agreement and disagreement and from their relative bargaining weights.*

8.6.2 Nash Solution using R

Let's set up the problem so that it is equivalent to the infinite alternating offer model analyzed above.

$$max_x \quad (xV_A - v_{ON})^{\lambda}((1-x)V_A - v_{EN})^{(1-\lambda)} \tag{8.5}$$

where V_A is the value of the agreement, v_{ON} is the value to Odd if there is no agreement, and v_{EN} is the same for Even, x is the proportion of the pie that Odd receives and λ is Odd's bargaining weight.

We can create a little numerical version of the model. You can see how things change when you change various parameters.

```
> lambda = 0.74
> nash_value = function(x) {
+   ((x*V_A - v_ON)^lambda)*(((1 - x)*V_A - v_EN)^(1-lambda))
+ }
> optimize(nash_value, c(0, 1), maximum = TRUE)
  $maximum
  [1] 0.7405195

  $objective
  [1] 0.5626717
```

The resulting share is similar to the results of the alternating offers game analyzed above. There is an equivalence between the bargaining weights in this model and the relative difference in the rs in the infinite period alternating offers model.

8.7 Modeling Hospital Competition and Pricing

How do we work out the effect of hospital mergers on prices when most of the customers don't pay anything or just a small fraction of the actual cost of the services? Insurers pay numbers closers to the actual costs but insurers don't really use the services. A solution is to use the Nash bargaining model to estimate how much the insurer will pay given the choices made by the insurer's beneficiaries.

In the early 2000s, economists of the FTC and in academia began rethinking how pricing worked in the hospital market. They realized that while hospitals provided services to patients, patients were not the ones that determined prices. Prices for hospital services are determined by the interaction of large hospitals bargaining with large insurers.

The section shows how the Nash bargaining model can be used to analyze hospital competition.

8.7.1 Bargaining Model

Consider a hospital and an insurer bargaining over the price of services (p). We have simplified things by assuming that the hospital only has one price and the insurer only has one set of beneficiaries. The insurer's payoff from a successful negotiation is just the value of the hospital to the insurer's beneficiaries ($v(h)$) minus the price paid to the hospital (p). The hospital's payoff is the price paid by the insurer (p) times the number of beneficiaries (q). If the bargaining fails, then the hospital gets no revenue from the insurer and the insurer gets the value of the alternative hospital to insurer's beneficiaries is $v(h')$ and the insurer pays p'.

- Successful:
  - Hospital: pq
  - Insurer: $(v(h) - p)q$
- Failure:
  - Hospital: 0
  - Insurer: $(v(h') - p')q$

The solution to the Nash bargaining model is as follows.

$$\max_p \quad (pq)^{\lambda}((v(h) - v(h') - (p - p'))q)^{1-\lambda} \tag{8.6}$$

For the hospital, the value of agreement is seeing the insurer's beneficiaries. For the insurer, it is the value to their beneficiaries of going to that hospital relative to the alternative less the relative price.

The first-order condition gives the following result. Let $A = pq$ and $B = (v(h) - v(h') - (p - p'))q$.

$$\begin{array}{l} \lambda A^{\lambda-1} q B^{1-\lambda} - (1-\lambda) A^{\lambda} B^{-\lambda} q = 0 \\ \lambda A^{-1} B - (1-\lambda) = 0 \\ \lambda B - (1-\lambda) A = 0 \end{array} \tag{8.7}$$

Substituting in the definitions of A and B, we get the following.

$$\begin{aligned}&\lambda(v(h) - v(h') - (p - p'))q - (1 - \lambda)pq = 0\\&p = \lambda(v(h) - v(h') + p')\end{aligned} \tag{8.8}$$

The price depends on the incremental value of the hospital to the insurer's beneficiaries, $(v(h) - v(h'))$, the price of the alternative hospital (p') and the bargaining weight of the hospital (λ).

To determine the market price, we need to estimate the incremental value of the hospital.

8.7.2 Demand for Hospitals

We don't know a beneficiary's value for a hospital but we can know their revealed preference. We used the same idea in Chapter 4 when looking at the choice of bookstores to enter a market.

A particular person will choose the hospital if the following inequality holds.

$$v_i(h) - v_i(h') > 0 \tag{8.9}$$

Our standard demand model replaces the inequality with the following. Again, this is what we did in Chapter 4. The hospital and the alternative hospital have observed characteristics $\mathbf{X}_h$ and $\mathbf{X}_{h'}$ respectively. The individual values characteristics according to weighting vector β. The unobserved characteristics are ξ_h and $\xi_{h'}$ respectively.

$$(\mathbf{X}_h - \mathbf{X}_{h'})'\beta + \xi_h - \xi_{h'} > 0 \tag{8.10}$$

Under certain assumptions on ξ, we get the logit form for the probability that an individual will choose hospital h.

$$s_h = \frac{\exp(\delta_h)}{1 + \exp(\delta_h)} \tag{8.11}$$

where $\delta_h = (\mathbf{X}_h - \mathbf{X}_{h'})\beta + \xi_h - \xi_{h'}$.

The beneficiary often pays close to nothing for hospital services, so we generally ignore the beneficiaries out of pocket expenses to simplify the problem.

8.7.3 Willingness-to-Pay

From our demand set up and what we know from the bargaining model, we can calculate what an insurance company would be willing to pay to keep a hospital in network. That is, keep the hospital available to its beneficiaries at the discount prices. A common measure is called willingness-to-pay (WTP), equal to $-\log(1 - s_h)$. That is, an insurance company is willing to pay a lot more for the hospital when its beneficiaries are not willing to go to any other hospital.

Why such a weird formula? Remember above we have Equation (8.7), which states the price should be equal to the difference in the relative value of the hospital and the next best alternative plus the price of the next best alternative. Let's not worry about the last part. What is $v(h) - v(h')$? Let's make things simpler and assume that $v(h')$ is just the outside option.

The first thing to note is in our logit world with our assumptions on the distribution of the unobserved values, $v(h) = \log(\exp(\delta_h) + 1)$. Why? Another excellent question. This formula is the expected value of optimal choice.[2] As h' is the outside option, $v(h') = \log(1) = 0$. It is the expected value when hospital h is removed as an option. So $v(h) - v(h') = \log(\exp(\delta_h) + 1)$.

Now, Equation (8.11) tells us what $\exp(\delta_h)$ is. Rearranging that equation, we have the following relationship between it and the **diversion ratio** to the hospital.[3]

$$\begin{array}{l} s_h(1 + \exp(\delta_h)) = \exp(\delta_h) \\ \exp(\delta_h)(1 - s_h) = s_h \\ \exp(\delta_h) = \frac{s_h}{1-s_h} \end{array} \tag{8.12}$$

Plugging this back in we get our WTP formula.

$$\begin{array}{ll} v(h) - v(h') & = \log(\frac{s_h}{1-s_h} + 1) = \log(\frac{s_h+1-s_h}{1-s_h}) = \log(\frac{1}{1-s_h}) \\ & = -\log(1 - s_h) \end{array} \tag{8.13}$$

Now we have a measure of how much the insurer is willing to pay for the hospital as a function of stuff that we observe in the data.

This simple measure has turned out to be an extremely good predictor of what actually happens in hospital markets. We generally find that a 10% increase in WTP is associated with a 2% increase in hospital prices.[4] Below we will estimate this relationship on data from Florida hospitals.

8.8 Empirical Analysis: Hospital Competition with R

By the early 2000s, the two federal antitrust agencies, the FTC and DOJ, had an impressive string of losses in hospital merger enforcement. The FTC's Republican Chairman, Tim Muris, decided to put a large amount of the commission's resources in turning the record around. While most of this work was legal analysis, the FTC's Bureau of Economics became heavily involved in the effort. A lot of the important work in modeling hospital mergers and measuring their potential impact has been done by economists who have been in the Bureau.

[2] See Capps et al. (2003).

[3] The diversion ratio is the share of people who go to the hospital over the share of people who go to all the other options in the market.

[4] See Bob Town's expert report in a Virginia hospital merger case, https://www.vdh.virginia.gov/content/uploads/sites/96/2017/10/Expert-Report-of-Robert-Town.pdf.

The section uses publicly available data from Florida, the US Census Bureau and the Centers for Medicaid and Medicare Services (CMS) to analyze demand for hospitals and uses the parameter estimates to simulate a merger in Palm Beach County.

8.8.1 Data

The section uses publicly available discharge from Florida for 2018. These data provide detailed information on discharges from Florida's hospitals at various demographics and conditions. The analysis uses the demographic data and assume demand is the same across conditions. A more standard analysis would group by condition ("DRG") as well. In addition to this, it is usual in hospital merger cases to have zip code level data. Here we only know the county where the hospital is. We assume that everyone discharged from the hospital lives in the same county and people in the county only choose between hospitals in the county.

Added to this information, we will use the American Community Survey data to determine counts of people in the various demographic groups at the county level. We assume that people who do not visit one of the county's hospitals in a particular year choose the outside option.

WTP is calculated using the observed shares at the demographic levels by county.

To construct these measures, we use weights. These weights are the importance of the group to the hospital. That is if a hospital specializes in a particular disease then that will be captured by the weighting and may lead to a high price even if the hospital doesn't have a large share of the market more generally. For example, the hospital may have a low market share overall but high market share for child birth. In that case, the market share for women aged 25 to 54 will be a lot more important than the same hospital's market share for men 55 or older. In the data we have 6 demographic groups. We label things `_fw` for firm weight and `_cw` for county weight. As we go through each case we find the importance of that demographic to the hospital then we find the share of that demographic in the market for the hospital. Last we calculate the WTP and the hospital share where the weights are the importance of the demographic to the hospital.

```
> ages = c("0_24", "25_54", "55")
> genders = c("female", "male")
> file = paste0(dir, "hospitals.csv")
> df = read.csv(file)[,-1]
> df$WTP = 0
> df$share = 0
> for(i in 1:length(ages)) {
+   for(j in 1:length(genders)) {
+     col = colnames(df)==paste("Discharges_",
```

```
+                               ages[i],"_",
+                               genders[j],"_fw",sep="")
+     weight = df[,col]
+     col = colnames(df)==paste0("Discharges_",
+                               ages[i],"_",
+                               genders[j],"_cw")
+     share = df[,col]
+     df$WTP = ifelse(share==1, df$WTP, df$WTP -
+                       weight*log(1 - share))
+     df$share = df$share + weight*share
+   }
+ }
```

Lastly, we match the data above to hospital cost and pricing reports from the Centers for Medicare and Medicaid (CMS). Not all the hospitals match and so the analysis is limited to the cases where we have matches across the discharge data and the cost reporting data.

8.8.2 Pricing and WTP

Our analysis is limited to estimating the effect of mergers on WTP. The measure could be used in its own right for merger review, similar to the way Herfindahl-Hirschman Index (HHI) or Upward Pricing Pressure (UPP) is used.[5]

Usually in merger cases some sort of relationship between prices and WTP is presented. We combine estimates of WTP and share above with information about prices and costs from CMS for Florida hospitals in 2018. The code brings in the data which combines information on Florida hospitals with prices, demographic and competition measures. It then runs two linear regressions, price on WTP and price on share of market.

```
> require(data.table)
> file = paste0(dir, "hospital3.csv")
> dt = fread(file)


> lm1 = lm(Price ~ Wages + Beds + WTP, data = dt)
> lm2 = lm(Price ~ Wages + Beds + share, data = dt)
```

Table 8.2 presents the linear regressions of price on measures of competition, WTP and share. Both regressions show that there is a positive relationship, although there is a lot of uncertainty. The estimated elasticity of 0.2 between WTP and price seems to be consistent with other estimates.

[5] See the 2010 Horizontal Merger Guidelines, https://www.justice.gov/atr/horizontal-merger-guidelines-08192010 accessed August 5 2023.

TABLE 8.2
Linear regression estimates of the relationship of price on WTP and share using data from Florida hospitals for 2018.

	Dependent variable:	
	Price	
	(1)	(2)
Wages	0.048	0.047
	(0.058)	(0.058)
Beds	0.010	0.007
	(0.067)	(0.067)
WTP	0.223	
	(0.141)	
share		0.357
		(0.221)
Constant	11.288***	11.301***
	(0.468)	(0.468)
Observations	106	106
R^2	0.034	0.035
Note:	*p<0.1; **p<0.05; ***p<0.01	

8.8.3 Mergers and Willingness To Pay

Consider a hospital merger in the home county for this publisher, Palm Beach County. The hospitals are Bethesda East, Bethesda West, and Boca Raton Regional. To estimate the effect of the merger, we recalculate the WTP for the combined hospital. The code is the same as above but recalculated just for the new merged firm.

```
> df2 = df
> merger = c("BETHESDA HOSPITAL EAST", "BETHESDA HOSPITAL WEST",
+            "BOCA RATON REGIONAL HOSPITAL")
> df$merge = ifelse(df$Hospital.Name %in% merger, 1, 0)
> df2$merge = ifelse(df2$Hospital.Name %in% merger, 1, 0)
> df2[df2$merge==1,]$Discharges =
+   sum(df2[df2$merge==1,]$Discharges)
> for(i in 1:length(ages)) {
+   for(j in 1:length(genders)) {
+     col = which(colnames(df)==paste0("Discharges_",
+                                     ages[i],"_",
+                                     genders[j],"_fw"))
+     df2[df2$merge==1,col] = sum(df2[df2$merge==1,col])
+     col = which(colnames(df)==paste0("Discharges_",
+                                     ages[i],"_",
+                                     genders[j],"_cw"))
+     df2[df2$merge==1,col] = sum(df2[df2$merge==1,col])
+   }
+ }
```

It then creates a new data where the shares are passed from the merging hospitals to the new hospital. We drop the merged hospitals from the data.

```
> df3 = df2[-which(df2$Hospital.Name %in%
+                    c("BETHESDA HOSPITAL EAST",
+                      "BETHESDA HOSPITAL WEST")), ]
```

Given the post-merger data, the WTP and share can be recalculated.

```
> df3$WTP = 0
> df3$share = 0
> l = 1
> for(i in 1:length(ages)) {
+   for(j in 1:length(genders)) {
+     col = which(colnames(df3)==paste0("Discharges_",
+                                      ages[i],"_",
+                                      genders[j],"_fw"))
+     weight = df3[,col]
+     col = which(colnames(df3)==paste0("Discharges_",
```

```
+                                 ages[i],"_",
+                                 genders[j],"_cw"))
+     share = df3[,col]
+     df3$WTP = ifelse(share==1, df3$WTP,
+                       df3$WTP - weight*log(1 - share))
+     df3$share = df3$share + weight*share
+     l = l + 1
+   }
+ }
```

Now we can determine the impact of the merger on prices using the WTP change caused by the merger on the price in Palm Beach County.

```
> index_merge = which(df$merge==1)
> index3_merge = which(df3$merge==1)
> a = sum(df$WTP[index_merge], na.rm = TRUE)
> b = sum(df3$WTP[index3_merge], na.rm = TRUE)
> c = (b - a)/a
> lm1$coefficients[4]*c*
+   mean(exp(dt$Price[dt$county.x == "palm beach"]),
+                              na.rm = TRUE)
      WTP
 53583.48
> lm1$coefficients[4]*c
       WTP
 0.4429545
```

This analysis suggests that a merger between these hospitals in Palm Beach County will have a substantive effect on price, a 44 percent increase or $53,583.48 per discharge.[6] Across the Florida hospitals in the sample, the price goes up $333 per discharge or 0.3 of a percent.

8.9 Discussion and Further Reading

The ultimatum game is one of the most common games used in experiments. The results used here suggest that very high stakes games do provide support for subgame perfection as a predictor of the outcome. However, Cameron (2007) suggests that may not always occur.

Bargaining models have become very important in industrial organization, particularly analysis of mergers. They form the heart of antitrust analysis

[6]This should not be consider a legal analysis, but rather an illustration of how these methods are used in legal analysis.

of hospital mergers but have also been used in other mergers where similar dynamics is at play.

Capps et al. (2003) came along at just the right time for US antitrust authorities. The agencies had been on an impressive losing streak with trying to prevent hospital mergers. To the credit of the antitrust agencies, the learnings presented there, in Gaynor and Vogt (2003), Gowrisankaran and Town (2003) and others, were incorporated into antitrust enforcement. The FTC brought a retrospective case against a hospital merger in Chicago and used these methods to prove that the merger was anticompetitive. That is, the FTC showed that the observed post-merger price increases were caused by the merger.[7]

Economists at Federal Trade Commission have made major contributions to this literature. The analysis presented in this chapter is heavily influenced by Raval et al. (2017). Raval et al. (2022) provide a very interesting test of the modeling approach. Garmon (2017) and Balan and Brand (2018) are among other great papers testing and using these methods to analyze the effects of hospital mergers.

[7] `https://www.ftc.gov/sites/default/files/documents/cases/2005/10/051020initialdecision.pdf`.

Part III

Static Games of Incomplete Information

9

Bayes Nash Equilibrium

9.1 Introduction

In the first two parts of the book, the games involved cases where all the players knew everything about what had happened or what was happening. We say that the players have complete information. In Parts III and IV of the book, we drop this assumption. The next two parts of the book consider games where players don't necessarily know who they are playing against, what their opponents payoffs are or what actions other players have played. In the first half the book, we could model chess. In this half of the book, we can model Stratego! If you don't know it, Stratego is a board game with pieces that have different attributes and abilities just like chess. The difference is that the pieces are all the exact same shape and the picture denoting the piece is only printed on one side. This means your opponent knows only that you have a piece in a particular square. They don't observe which piece. Stratego can be modeled as a dynamic game of incomplete information.

The problem with modeling games of incomplete information is that it is not clear how to do it. That we can model these situations is really due to some amazing work by an Hungarian mathematician, John Harsayni, who immigrated to Australia to escape the communists and then immigrated to the United States to escape the Australians.

Harsayni's insight was to think of games of incomplete information as games of complete information. Brilliant! What? Harsanyi realized that we could think of the problem of not knowing the other player's payoffs as a problem of not knowing the other player.

This chapter introduces the idea of **beliefs**. If something is unknown by the players, we need a way to quantify what is unknown. The assumption is that each player of the game places probability weights on the unknown events of the game. It is assumed that these probability weights (beliefs) are known to everyone in the game.

This chapter introduces the idea of a Bayesian game and the Nash equilibrium for such a game. It illustrates the idea using entry games. We saw entry games in Chapters 4, 5, and 6. This time it is assumed that the firms contemplating entering a market do not know about the characteristics of the other firms contemplating entering the market. Interestingly, when players

DOI: 10.1201/9781003351603-9

know about the characteristics of the other firms, the equilibrium is a lot more complicated than when players don't know about the characteristics of the other firms. Making the game more complicated makes the game easier to use for empirical analysis. Yale economist, Katja Seim makes this point in her 2006 *RAND Journal of Economics* paper on the video retail industry (Seim, 2006).

9.2 A Bayesian Game

We call a game with uncertainty over the types, a **Bayesian game** because we require players to use **Bayes rule** to update their **beliefs** (the probability weights). Although we are restricting the information available to players, we still require the players to process a lot of information. Newer game theory models relax some of these assumptions and explore the implications.

The section introduces a game based on an actual situation than can happen in undergraduate courses, it then formally defines the game and the equilibrium concept, **Bayes Nash equilibrium**.

9.2.1 A Grading Game

Let's analyze a game in which the players are students in a game theory class. Grading is done on a curve. Everyone who gets below the mean score for the class gets a B and everyone above the class mean gets an A. Assume that the students in the class are grade focused. They would prefer to only get As in their courses if they can help it. Also, students can drop the class at any time without penalty. This last bit is unrealistic, but it makes the analysis simpler.

For each student the problem is to determine whether or not to drop the class. Assume that each student observes their own raw score in the class. For example they may know that their grades add up to a grand total of 33 out of 100. While they don't know raw scores of the other students, they have been told the distribution of raw scores. They can see that the mean is around 50 and scores range from close to 0 up to close to 90.

Assume that each student will stay in the class if they believe that they will get an A but will drop the class if they believe that they will get a B. To simplify the problem, we can assume that each student plays a **cutoff strategy**. A strategy here is a mapping from the player's raw score, to either stay or drop. A **cutoff strategy** states that the player will stay in the course if their score is above some cutoff level and will drop the class if it is below.

9.2.2 Grading Game Simulation using R

It is easier to look at a simulation. Assume we have 100 students in the class and their raw scores are determined by a normal distribution with a mean of

50 and a standard deviation of 20.

```
> set.seed(123456789)
> N = 100
> score = rnorm(N, mean=50, sd=20)
> summary(score)
    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
    4.41   36.61   52.73   50.50   64.98   89.80
```

What will happen in this game? Let's assume that everyone plays a strategy where they never drop. In this case a student in the top half will get an A and a student in the bottom half will get a B.

```
> # Initial grades
> grade = ifelse(score > mean(score), "A", "B")
> table(grade)
grade
 A  B
54 46
```

Assume that students play the strategy that if their scores is above 50 they will stay and students will drop if their score is below 50. That is, if the student thinks they will get an A, they stay in the class. If the student thinks that they will get a B, they will drop the class.

Now, if this is the strategy played by all the players, the grades will be different. Remember the grades are curved based on students who are in the class. The score distribution for those students will be a truncated normal distribution (the top half).

```
> s = score > 50
> score_1 = score[s]
> summary(score_1)
    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   50.18   56.89   64.37   64.95   73.29   89.80
> length(score_1)
[1] 55
```

If students play this strategy, then the mean jumps from 50 to 65 and 45 students drop the class. Given the new reality, the grades will be adjusted.

```
> # New grades
> grade_1 = ifelse(score_1 > mean(score_1), "A", "B")
> table(grade_1)
grade_1
 A  B
25 30
```

```
> mean(score_1)
  [1] 64.9546
```

The number of students who get an A drops to 25 and they have to have a grade above 65.

Assume that the students adjust to this new reality, so their strategy changes. They will stay if their grade is above 65 as this guarantees them an A for the class.

```
> s_1 = score > 65
> score_2 = score[s_1]
> summary(score_2)
     Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
    65.36   71.05   73.49   74.06   78.25   89.80
```

Now the remaining students need to update their beliefs about their grade given the strategies of the other students. Given this update, another 30 students drop the class and the mean increases to 74.

```
> grade_2 = ifelse(score_2 > mean(score_2), "A", "B")
> table(grade_2)
  grade_2
   A  B
  10 15
> s_2 = score > 74
> score_3 = score[s_2]
> summary(score_3)
     Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
    74.39   76.74   78.58   79.56   79.56   89.80
```

Updating again, another 15 students drop the class and the average grade increases to 80.

Will the class have any students in it?

Dartmouth College has a provision to stop this unraveling. While Dartmouth's econ department uses a curve, professors are allowed to include students who dropped the class in calculating the grade distribution, where such students are given the lowest score for the purposes of doing the calculation. A policy like this may still lead students to drop based on beliefs of their grade in the class, but it doesn't lead to the unraveling. What is the outcome of the game used by Dartmouth?

9.2.3 Definitions

Now we have a taste for how these games work, let's get more formal. What do we mean by a **game of incomplete information**?

Definition 15. *In a* game of incomplete information*, players don't necessarily observe the actions of other players or know the payoffs of the other players.*

In Parts I and II of the book, players know exactly what is happening or has happened at every moment. In Parts III and IV of the book, they don't.

Assumption 2. *There are a set of player* types *determining the payoffs the player will get in each outcome. This set of types are known to all the players.*

While players don't know exactly what is going on, we will make the assumption that they do know player **types**. Assumption 2 is a super important idea, it allows us use all the machinery we have developed for analyzing complete information games to analyze incomplete information games. The player's type captures all the information relevant to the game. If each player's type is observed by every other player of the game, then the game is one of complete information.

Assumption 3. *Each player knows their own type and the distribution of types (the probability that the other player is a particular type).*

What makes games of incomplete information hard to think about is the implications of Assumption 3. When analyzing the game, we must be careful to remember what exactly the players know and what they do not know.

9.2.4 The Game

A Bayesian game has the following form.

- Players: Set of players and a set of types.
- Strategies: A function mapping from type to actions.
- Payoffs: For each type and each outcome a payoff.
- Beliefs: A probability distribution over types that is known by each player.

Our strategies are more complicated than for static games of complete information. Again, a strategy is complete plan. In this case, the complete plan is determined prior to the player knowing their type. That is, the plan states what the player will do for each possible type that they could be.

We have added one more piece to the basic game description. We generally refer to this known probability distribution over types as **beliefs**. Players do not know the types of the other players in the game, but they do know probability that another player is a particular type.

9.2.5 Course Grading Game

For the game introduced above, we have the following formal description. In this game, the player's type is the raw score of the player in the class which is denoted θ_i. The raw score is between 0 and 100. The player's strategy states whether the player will stay or leave based on the raw score they observe.

- Players: $N = 100$ students where each observes their score in the class, $\theta_i \in [0, 100]$.
- Strategies: Each player $i \in N$ chooses $s(\theta_i) \in \{0, 1\}$, where 1 means stay.
- Payoffs:
  - Stay and $\theta_i > m$: A
  - Stay and $\theta_i \leq m$: B

  Leave: 0, where $B < 0 < A$

  where m is the mean grade of the students remaining in the class.
- Beliefs: $\theta_i \sim F$

When we write down the game we don't worry about whether there exists an equilibrium of the game. It is also going to turn out that the initial beliefs don't matter that much, so let's just call it F.

9.2.6 Equilibrium

Given our new set up, we need a new equilibrium concept.

Definition 16. *A* Bayes Nash equilibrium *is an outcome where for each type, the outcome cannot be improved upon given the strategies of the other players and beliefs about the distribution of types. Where beliefs are consistent with equilibrium strategies.*

First, we are still assuming Nash equilibrium. Strategies must be optimal given the strategies of the other players. What is new is this idea of beliefs.

Definition 17. *A player's* belief *is what the player knows about the distribution of types playing each strategy.*

The equilibrium concept requires that the beliefs of the players be consistent with the equilibrium strategies. We are assuming that players are choosing their optimal strategies given expected payoffs. The players may not know exactly what payoff will occur because they don't know exactly which type of player they are playing against.

Here it is important to point out the difference between assumptions about the game and assumptions about the predictions of the game (the equilibrium concept). In the game, the players only know their own type and the distribution of types for the other players. In a Bayes Nash equilibrium, it may be that players know exactly the type of the other players because their beliefs must be updated to be consistent with the equilibrium strategies. Hopefully, this distinction will become clearer as we work through examples.

9.2.7 Equilibrium of Course Grade Game

Is there an equilibrium where each student only stays in the class if their grade is an A? No.

We are going to show this using a proof by contradiction. Let c be the cutoff such that if $\theta_i > c$, then student i stays; otherwise, they drop. If c is the same for every student in the class then c is the minimum grade for the students that stay in the class. Given the letter grades are determined by the mean grade of students that stay in the class, then $m = \mathbb{E}(\theta_i | \theta_i > c)$. The curve requires that A is given if $\theta_i > m$ and B is given otherwise.

The proposed equilibrium requires that $m \leq c$. If c and m exists, then $m = \mathbb{E}(\theta_i | \theta_i > c) > c$, a contradiction.

9.3 Empirical Entry Game

We first looked at entry games in Chapter 4. Those were static games of complete information. We revisited entry games in Chapter 5 and again in Chapter 6. The last time we modeled them as dynamic games of complete information. Now we are going to revisit them again. This section models entry games as static games of incomplete information. Comparing our analysis in this chapter to the analysis in Chapters 4 and 5, the assumptions in this chapter make the game itself more complicated, but the equilibrium easier to analyze.

It is worth contemplating the difference between the assumption made here and in Chapter 5. In both cases in equilibrium the firms don't know exactly if the other firm will enter the market. In both cases, they know the probability that the other firm will enter. These models are different. In the model used in Chapter 5, each firm knows the value of the unobserved entry costs, ξ. The econometrician does not know this value, but the firms playing the game do. Here, the value of ξ is unknown to both the econometrician and some of the players of the game. This is a subtle distinction but it substantially changes how we estimate the model.

9.3.1 The Game

- Players: Barnes & Noble (and ξ_{1i}), Borders (and ξ_{2i})
- Strategies:
  - Given ξ_{1i} (and $\mathbf{X}_i$) Barnes & Noble chooses enter or not enter.
  - Given ξ_{2i} (and $\mathbf{X}_i$) Borders chooses to enter or not enter.
- Payoffs:

- Barnes & Noble
  * Enter: $\mathbf{X}_i'\beta_1 - \alpha_1 \Pr(D_{2i} = 1|\xi_{1i}) + \xi_{1i}$
  * Not Enter: 0
- Borders
  * Enter: $\mathbf{X}_i'\beta_2 - \alpha_2 \Pr(D_{1i} = 1|\xi_{2i}) + \xi_{2i}$
  * Not Enter: 0
- Beliefs: $\{\xi_{1i}, \xi_{2i}\} \sim \Phi_2(0, \Sigma)$, where $\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$.

where D_{1i} indicates whether or not Barnes & Noble enters, while D_{2i} indicates Borders entry.

This game is similar to the one we used in Chapter 4. There are unobserved characteristics of the firms and the market, ξ, that are now unobserved by the other player. These are the player types. Entry is determined by observed characteristics $\mathbf{X}_i$ of the market i and the parameter β which is the same across markets. The parameter α determines the extent to which competition from the other firm reduces the benefits of entry into a particular market.

9.3.2 Equilibrium

The difference between this game and what we saw in Chapter 4 is the assumption about what Barnes & Noble knows about D_2 in equilibrium. In Chapter 4 we assumed that in equilibrium, Barnes & Noble knew the exact value of D_2. They knew whether or not Borders was also entering the market. In this chapter we assume that in equilibrium Barnes & Noble only knows the probability that $D_2 = 1$. More precisely, in equilibrium Barnes & Noble accounts for Border's equilibrium strategy in Barnes & Noble's beliefs about D_2. It's value D_2 is determined by the following inequality.

$$\mathbf{X}_i'\beta_2 - \alpha_2 D_{1i} + \xi_{2i} > 0 \tag{9.1}$$

Assume that there exists an equilibrium in cutoff strategies, where the cutoff values are $\{c_{1i}, c_{2i}\}$, respectively. Barnes & Noble's expectation about Border's entry into market i is as follows.

$$\begin{array}{ll} \mathbb{E}(D_{2i}|\mathbf{X}_i, \xi_{1i}) & = \Pr(\xi_{2i} > c_{2i}|\mathbf{X}_i, \xi_{1i}) \\ & = \Pr(\xi_{2i} > -\mathbf{X}_i'\beta_2 + \alpha \Pr(\xi_{1i} > c_{1i})|\xi_{1i}) \end{array} \tag{9.2}$$

In determining the probability that Borders will enter, Barnes & Noble must account for Border's beliefs that Barnes & Noble will enter. Things are a lot simpler if we assume that the unobserved term, the ξ's are independent of each other. That seems pretty unrealistic. We would expect a lot of things about the market to be correlated. While Barnes & Noble doesn't know Border's costs of entering exactly, it does observe its own costs and can make an inference about Border's costs.[1]

Even still, there is a distinct advantage of using games of incomplete information to model entry. In Chapter 4, there were multiple equilibria at certain values of the ξs. That is not the case here. The equilibrium is unique in the cutoff strategies of the firms. The only issue is determining what that equilibrium is!

9.3.3 Estimating Entry Games

Can we back out the distribution of entry costs by looking at the distribution firms in different markets? We can. However, we must solve for the equilibrium in order to do so. It is assumed that the entry costs are distributed bivariate normal. The question is whether we can estimate the correlation coefficient (ρ). In a standard bivariate probit model, it is assumed that the actions are correlated but the optimality of the actions only depend on each other through the correlation. Here the actions themselves are interdependent.

The seminal work of Guerre et al. (2000) suggests that it is not necessary to actually solve for the equilibrium. Instead it is simpler to do a "two-step" procedure and the correlation parameter.

In the first step, we use maximum likelihood to estimate the cutoff value. In this step, we are not making any claims about the parameter values we are estimating. We are assuming that the observed data is being generated by some sort of stationary process. As with any discrete choice type data, our identification is heavily reliant on parametric assumptions. The likelihood function is determined by the standard bivariate normal. The probabilities of the four states are as follows. Note the notation. There is a squiggly line over all the parameters to remind us that these are not the structural parameters but the parameter values coming from the **reduced form estimation** in the first step.[2]

Maximizing the log-likelihood function provides estimates of $\tilde{\beta}_1$, $\tilde{\beta}_2$, and $\tilde{\rho}$. The model for estimating these parameters is presented in Chapter 4.

With this information, we can determine probability of entry for each firm. These probabilities are estimated because they are based on the parameters estimates in the first step.

$$
\begin{aligned}
\Pr(D_{1i} = 1|\mathbf{X}_i, \xi_{2i}) &= 1 - \Phi\left(\frac{-\mathbf{X}_i\hat{\tilde{\beta}}_1 - \hat{\tilde{\rho}}\xi_{2i}}{\sqrt{(1-\hat{\tilde{\rho}}^2)}}\right) \\
\Pr(D_{2i} = 1|\mathbf{X}_i, \xi_{1i}) &= 1 - \Phi\left(\frac{-\mathbf{X}_i\hat{\tilde{\beta}}_2 - \hat{\tilde{\rho}}\xi_{1i}}{\sqrt{(1-\hat{\tilde{\rho}}^2)}}\right)
\end{aligned} \tag{9.3}
$$

Given the estimated probabilities of entry, we can estimate our structural

[1]To be clear, we assume that the firms each know all the observed characteristics, $\mathbf{X}_i$ as well as all the parameter values.

[2]We use the term reduced form to refer to standard empirical estimation techniques where we are not relying on assumptions about actors generating the data are behaving.

parameters.

The probabilities are now without the squiggly lines. This is a reminder that these probabilities are structural estimates. They come from the game theory model. The probabilities with the squiggly lines are observed in the data and don't depend upon any assumptions about how the data are generated.

We also have two extra parameters, α_1 and α_2. Both firms will enter if the following inequality holds. The ξs appears in two different places in the inequalities.

$$\begin{array}{c} \mathbf{X}_i\beta_1 - \alpha_1 \hat{\Pr}(D_{2i} = 1|\mathbf{X}_i, \xi_{1i}) + \xi_{1i} > 0 \\ \mathbf{X}_i\beta_2 - \alpha_2 \hat{\Pr}(D_{1i} = 1|\mathbf{X}_i, \xi_{2i}) + \xi_{2i} > 0 \end{array} \tag{9.4}$$

where $\{\xi_{1i}, \xi_{2i}\} \sim \mathcal{N}_2(0, \Sigma)$ and $\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$.

9.3.4 Estimating Entry Games in R

Similarly, we can write out a function to represent the entry inequalities for the second stage of the estimation (Equation (9.4)). In the code, the extra `t` denotes the parameters estimated from the reduced form model (the squiggly line parameters). Chapter 4 estimates these parameters using `f_entry2()`.

```
> f_biprobit = function(X, beta_1, beta_2,
+                       alpha_1, alpha_2, rho,
+                       beta_1t, beta_2t, rhot) {
+   N = dim(X)[1]
+   xi_1 = Z_1
+   xi_2 = Z_2*sqrt(1 - rho^2) + rho*Z_1
+   Xb_1 = X%*%beta_1
+   Xb_2 = X%*%beta_2
+   Xb_1t = X%*%beta_1t
+   Xb_2t = X%*%beta_2t
+   p_00 = p_01 = p_11 = rep(0, N)
+   for(k in 1:K) {
+     D_1k = 1 -
+       pnorm(0, Xb_1t + rhot*xi_1[k], sqrt(1 - rhot^2))
+     D_2k = 1 -
+       pnorm(0, Xb_2t + rhot*xi_2[k], sqrt(1 - rhot^2))
+     pi_1k = Xb_1 - alpha_1*D_2k + xi_1[k]
+     pi_2k = Xb_2 - alpha_2*D_1k + xi_2[k]
+     p_00 = p_00 + (pi_1k < 0 & pi_2k < 0)
+     p_01 = p_01 + (pi_1k < 0 & pi_2k > 0)
+     p_11 = p_11 + (pi_1k > 0 & pi_2k > 0)
+   }
+   return(list(p_00 = p_00/K,
+               p_01 = p_01/K,
```

```
+                p_11 = p_11/K))
+ }
```

The second step estimator, `f_biprobit`, uses simulation to calculate the probabilities. The code uses `pnorm()` to calculate the probability of entry given the estimated parameters from the first step.

9.4 Empirical Analysis: Mega Bookstore Entry (Again) using R

This section re-estimates the entry of Barnes & Noble and Borders under the assumption that each bookstore chain does not know the exact costs of entry of their competitor. This assumption seems more realistic than the assumption made in Chapter 4. While the assumption makes the estimator somewhat more complicated, it has the nice property that there aren't multiple equilibria.

9.4.1 Estimates

Table 9.1 presents the mean and standard deviation of the coefficient estimates from the first and second stages of the two-step estimator. The estimates show that population, college education and the number of bookstores in 1990 are all important determinants of entry into a county by these two mega bookstore chains. It also shows that Barnes & Noble is much less likely to enter than Borders and that the two firms do not like to compete. It is interesting to compare these results to the results presented in Chapter 4, 5 and 6.

Table **??** presents the model fit excercise comparing the two models with the actual data. The Bayes Nash equilibrium model does a better job of predicting the case where there is only one firm in the market, but does a worse job of predicting the case where there are two firms in the market.

9.4.2 Policy

Given our new estimates, we can reconsider the policy question analyzed in Chapter 4. In that analysis, the merger resolved a coordination problem for the firms. In general, the two firms prefer to have only one of them in the market, but when the firms are independent they cannot coordinate on which one. In Chapter 4, the firms knew whether or not the other firm was going to enter, under these model assumptions they do not. The merger resolves the information problem inherent in the modeling assumptions.

Table 9.2 presents the results of simulating the merger. We see fewer cases where the two firms compete as well as fewer markets with entry. Theoretically this model can have ambiguous predictions about the impact of the merger on consumer welfare, but the empirical estimates suggest that the merger would

TABLE 9.1
Results from estimates of the first and second stage of the estimates assuming a Bayes Nash equilibrium. The first set of columns labeled "First Stage" are estimates assuming that the two firms are making entry decisions that are strategically independent but statistically independent. These are the same as the results presented in Chapter 4 for the same model. The second set of columns labeled "Second Stage" assumes the firm entry decisions are strategically dependent, but each firm does not know exactly if the other firm will enter, but assumes their strategy is consistent with observed entry decisions.

	First Stage	SD	BNE	SD
const_1	−15.03	0.09	−15.03	0.23
Pop_1	1.09	0.01	1.06	0.02
Income_1	−1.04	0.25	−0.75	0.36
College_1	5.51	0.34	5.85	0.40
Stores_1990_1	0.28	0.07	0.44	0.09
const_2	−11.54	0.11	−11.39	0.17
Pop_2	0.66	0.01	0.63	0.02
Income_2	1.74	0.25	1.80	0.39
College_2	2.59	0.33	2.75	0.54
Stores_1990_2	0.49	0.06	0.70	0.10
alpha_1			0.87	0.22
alpha_2			0.64	0.17
rho	−0.08	0.10	−0.01	0.11

TABLE 9.2
Comparison of actual entry in 2000 compared to simulated entry and simulated entry under a merger assuming a Bayes Nash equilibrium of the entry game.

	Actual	Sim	Merger
none	2919	2643	2857
BN	155	284	222
Borders	15	148	71
both	128	104	29

have lowered consumer welfare.

Compare these predictions to the predictions in Chapter 4. It is the same policy analysis. What is different is the assumption about the information that the competitors have.

9.5 Discussion and Further Reading

Modeling entry games as games of complete information leads to all sorts of weirdness (Bresnahan and Reiss, 1990; Tamer, 2003). However, when the game is more realistic by reducing the information available to the players, the game gets a whole lot simpler to analyze. Seim (2006) uses these ideas to analyze entry of retail video stores.

The estimator here is based on Guerre et al. (2000). The argument is that we can assume that the observed data are the result of equilibrium behavior. We can estimate a reduced form model in the first step and then impose equilibrium behavioral assumptions to back out the parameters of the underlying game theory model. We come back to this idea in Chapter 10.

This chapter uses data from Adams and Basker (2025) and their analysis of entry of the mega bookstores.

10

Auctions

10.1 Introduction

According to the travel presenter, Rick Steves, the Aalsmeer auction house is one the largest commercial buildings in the world. Royal Flora Holland, the owner of Aalsmeer, sold 12.5 billion plants and flowers in 2016 through its auction houses. But with $5.2 billion in auction sales, Royal Flora Holland is nowhere near the biggest auction house in the world.[1] That honor may go to Google. Google sold $47.6 billion in search ads using what the economist, Hal Varian, called the biggest auction in the world (Varian, 2007).[2] While that is impressive, a single auction in 2015 almost beat Google's annual number. The US Federal Communication Commission's auction number 97 (AWS-3) raised $44.9 billion dollars for US taxpayers.[3] That pails in comparison to the fact that every week the US Federal government offers billions of dollars in securities auctions. A single 4-week T-bill auction for July 6 2023 was for $70 billion.

Auctions are used to sell and buy a large number of products. Governments use auctions to purchase everything from paper to police body cameras. The US Federal government uses auctions to sell oil drilling rights, FCC spectrum, 10 year bonds and timber access. You can sell and buy items from eBay.com using auctions.

The auctions at Aalsmeer are unique. The auction runs for a short amount of time with a "clock" clicking the price down as the auction continues. As the price falls, the first bidder to hit the button, wins, at whatever price the clock is at. A spokesperson for Aalsmeer stated that because the price falls, it is called a **Dutch auction**. Actually they got the causality backwards. Because the Dutch popularized these types of auctions for selling flowers, we call them **Dutch auctions**.

The auction style you may be most familiar with is called an **English auction**. In this auction, there is an auctioneer who often speaks very very fast and does a lot of pointing while bidders hold up paddles or make hand gestures. In English auctions, the last bidder wins and pays the price at which the bidding stops.

[1] https://www.royalfloraholland.com
[2] eMarketer.com, 7/26/16
[3] https://www.fcc.gov/auction/97/factsheet

DOI: 10.1201/9781003351603-10

Economic analysis of auctions began with William Vickrey's seminal 1961 paper, *Counterspeculation, Auctions, and Competitive Sealed Bid Tenders*. Vickrey pointed out that Dutch auctions and **sealed bid auctions** are **strategically equivalent**. In a standard sealed bid auction each bidder submits a secret written bid. The auctioneer chooses the highest bid, and the bidder pays the number written down in her bid.

Vickrey characterized what a bidder should optimally bid in such an auction. He then showed that the same bidder should bid exactly the same amount in a Dutch auction. That is, in a Dutch auction, the bidder should wait until the price falls to the number written down, and then hit the button. Vickrey showed that while these two auctions formats are strategically equivalent, they are not strategically equivalent to an English auction.

Vickrey wondered if there was a sealed bid auction that is strategically equivalent to an English auction. Vickrey invented a new auction. In a **Vickrey auction**, each bidder writes down a bid like in a standard sealed bid auction and the winner is the person who writes down the highest bid. However, the winner pays the amount written down by the second highest bidder. Vickrey showed that his auction is strategically equivalent to an English auction.

This chapter discusses two of the most important auction formats, sealed bid auctions and English auctions. It presents estimators for both. The sealed bid auction estimation is based on Guerre et al. (2000). The English auction analysis uses the order statistic approach of Athey and Haile (2002). In both cases, it presents analysis of timber auctions. The chapter tests whether loggers are bidding rationally in sealed bid auctions and whether loggers colluded in English auctions.

10.2 Sealed Bid Auctions

Sealed bid auctions are one of the most commonly used auction formats. These auctions are very prominent in procurement, both in government and in the private sector. In a sealed bid auction, each bidder writes down her bid and secretly submits it to the auctioneer. The auctioneer sorts the bids from highest to lowest (or lowest to highest if they are buying instead of selling). The winner is the highest bidder and she pays the amount she wrote down. This is called a **first price auction** because the price is determined by the highest bid or first price.

Vickrey pointed out that sealed bid auctions are strategically complicated. To see this, assume that a bidder's utility for an item is equal to their intrinsic value for the item less the price they pay for the item. For example, a logger bidding in a timber auction will earn profits from the logs less the price paid to the US Forestry service for access to the trees. If a logger bids an amount equal to her expected profits, then if she wins she will earn nothing from the logging. It is optimal for the logger to shade her bid down. The problem is that the more she shades down, the lower her chance of winning the auction. The bidder must calculate the trade off between the probability of winning the auction and the value of winning the auction.

The section presents the model of a sealed bid auction, it then simulates data from such an auction. The section develops an estimator for determining each bidder's type, or valuation for the item.

10.2.1 Sealed Bid Model

The sealed bid game has N bidders and each bidder i knows their own type v_i.

- Players: N bidders each with valuation v_i (type)
- Strategies: For each valuation (type) v_i, for bidder i, she chooses a bid $b_i(v_i)$.
- Payoffs:
  - $b_i > b_j \forall j \neq i$: $v_i - b_i$
  - $b_i < b_j \forall j \neq i$: 0
- Beliefs: $v_i \sim F$

We will ignore ties.[4]

If the bidder has the highest bid she wins and has a payoff $v_i - b_i$, which is her intrinsic value less the amount of the bid. If she loses she gets nothing.

Assumption 4. *Independent Private Values (IPV). Let $v_i \overset{iid}{\sim} F$, where v_i is the value of bidder i and F is the distribution function.*

Assumption 4 makes the exposition a lot simpler. It also seems to be a reasonable approximation for the problems considered in the chapter. It states that a bidder's value for the item is unrelated to the values of the other bidders in the auction, except that they draw their valuation from the same distribution. The next chapter considers an alternative assumption where valuations are associated with each other.

10.2.2 Bayes Nash Equilibrium

In equilibrium, the bidder is assumed to maximize her expected returns from the auction. Assume that the bidder gets 0 if she loses. If she wins, assume she gets her intrinsic value (v_i) for the item less her bid (b_i).

$$\max_{b_i} \quad \Pr(\text{win}|b_i)(v_i - b_i) \tag{10.1}$$

[4]If the bid increments are small and the valuations are continuously distributed, then the probability of a tie is small.

If we take first-order conditions of Equation (10.1), then we get the following expression.

$$g(b_i|N)(v_i - b_i) - G(b_i|N) = 0 \tag{10.2}$$

Let $G(b_i|N)$ denote the probability that bidder i is the highest bidder with a bid of b_i, conditional on there being N bidders in the auction, and $g(b_i|N)$ is the derivative. $G(b_i|N)$ is the probability that she wins the auction.

We can rearrange this formula to show how much the bidder should shade her bid.

$$b_i = v_i - \frac{G(b_i|N)}{g(b_i|N)} \tag{10.3}$$

The formula states that the bidder should bid her value, less a shading factor which is determined by how much a decrease in her bid reduces her probability of winning the auction.

It will be useful for our code to write the probability of winning the auction as a function of the bid distribution as this distribution is observed in the data. Let $H(b)$ denote the distribution of bids in the auctions. Given Assumption 4, the probability of a particular bidder winning the auction is given by the following equation.

$$G(b_i|N) = H(b_i)^{N-1} \tag{10.4}$$

If there are two bidders in the auction, then the probability of winning is simply the probability that her bid is higher than the other bidder. If there are more than two bidders, it is the probability that her bid is higher than *all* the other bidders. The independent private values assumption, Assumption 4, implies that this is the probability that each of the other bidders makes a bid less than hers, all multiplied together.

We can also determine the derivative of this function in terms of the bid distribution observed in the data.

$$g(b_i|N) = (N-1)h(b_i)H(b_i)^{N-2} \tag{10.5}$$

where h is the derivative of the bid distribution H.

10.2.3 Sealed Bid Simulation in R

In the simulated data, each bidder draws their value from a uniform distribution. Vickrey shows that the optimal bid in this auction is calculated using the following formula.

$$b_i = \frac{(N-1)v_i}{N} \tag{10.6}$$

In Vickrey's version of the game, bidders know the function represented by Equation (10.6). The uniform distribution simplifies the problem, which is why it is used. In each simulated auction, there are different numbers of simulated bidders.

```
> set.seed(123456789)
> M = 1000  # number of simulated auctions.
> data1 = matrix(NA,M,12)
> for (i in 1:M) {
+   N = round(runif(1, min=2,max=10)) # number of bidders.
+   v = runif(N) # valuations, uniform distribution.
+   b = (N - 1)*v/N # bid function
+   p = max(b)  # auction price
+   x = rep(NA,10)
+   x[1:N] = b  # bid data
+   data1[i,1] = N
+   data1[i,2] = p
+   data1[i,3:12] = x
+ }
> colnames(data1) = c("Num","Price","Bid1",
+                     "Bid2","Bid3","Bid4",
+                     "Bid5","Bid6","Bid7",
+                     "Bid8","Bid9","Bid10")
> data1 = as.data.frame(data1)
```

The simulation creates a data set with 1,000 auctions. In each auction, there is between 2 and 10 bidders. The bidders are not listed in order.

10.2.4 Sealed Bid Estimator

The estimator uses Equation (10.3) to back out values from observed bids. To do this, we calculate the probability of winning the auction conditional on the number of bidders. It should be straightforward to determine this from the data. Once we have this function, we use the formula to determine the bidder's valuation from their bid.

The first step is to estimate the bid distribution.

$$\hat{H}(b) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}(b_i < b) \tag{10.7}$$

The **non-parametric estimate** of the distribution function, $H(b)$, is the fraction of bids that are below some value b.[5]

[5]A non-parametric estimator makes no parametric assumptions about how the bids are distributed in the data.

The second step is to estimate the derivative of the bid distribution. This can be calculated numerically for some given "small" number, ϵ.[6]

$$\hat{h}(b) = \frac{\hat{H}(b+\epsilon) - \hat{H}(b-\epsilon)}{2\epsilon} \tag{10.8}$$

If there are two bidders, Equation (10.3) determines the valuation for each bidder.

$$\hat{v}_i = b_i + \frac{\hat{H}(b_i)}{\hat{h}(b_i)} \tag{10.9}$$

where $i \in \{1, 2\}$.

10.2.5 Sealed Bid Estimator in R

The estimator limits the data to only those auctions with two bidders. In this special case, the probability of winning is just given by the distribution of bids.[7] In the code, the `epsilon` stands for the Greek letter, ϵ, and refers to a "small" number. See Equation (10.8).

```
> f_sealed_2bid = function(bids, epsilon=0.5) {
+   # epsilon for "small" number for finite difference method
+   # of taking numerical derivatives.
+   values = rep(NA,length(bids))
+   for (i in 1:length(bids)) {
+     H_hat = mean(bids < bids[i])
+     # bid probability distribution
+     h_hat = (mean(bids < bids[i] + epsilon) -
+       mean(bids < bids[i] - epsilon))/(2*epsilon)
+     # bid density
+     values[i] = bids[i] + H_hat/h_hat
+   }
+   return(values)
+ }
```

It is straightforward to calculate the probability of winning, as this is the probability the other bidder bids less. Given IPV (Assumption 4), this is the cumulative probability for a particular bid. Calculating the density is slightly more complicated. We can approximate this derivative numerically by looking at the change in the probability for a "small" change in the bids.[8] The value is calculated using Equation (10.3).

[6] This is a **finite difference estimator**.

[7] The probability of winning is the probability that your bid is higher than the other bidders in the auction.

[8] This is an example of using finite differences to calculate numerical derivatives. What happens with different values of `epsilon`?

The code creates a `ggplot()` object that shows the density of bids and the derived valuations from the two-person auctions.

```
> data2 = data1 |>
+   filter(
+     Num == 2
+   )
> ggplot_sealed = data.frame(
+   "bids" = c(data2$Bid1, data2$Bid2),
+   "values" = f_sealed_2bid(c(data2$Bid1, data2$Bid2))
+ ) |>
+   ggplot() +
+   geom_density(aes(values),
+                fill = "gray",
+                alpha = 0.5) +
+   geom_density(aes(bids),
+                fill="black",
+                alpha = 0.5) +
+   scale_x_continuous(limits = c(-0.5,1.5)) +
+   labs(
+     x = "value",
+     y = "",
+     title = "Density of Bids and Values"
+   ) +
+   geom_text(aes(x = 1.2, y = 0.8, label = "Bids"),
+             color = "gray") +
+   geom_text(aes(x = -0.2, y = 1.2, label = "Values"),
+             color = "gray") +
+   theme(axis.text.y=element_blank(),
+         axis.ticks.y=element_blank())


> ggplot_sealed
```

Figure 10.1 shows that the bids are significantly shaded from the true values, particularly for very high valuations. The figure presents the density functions for bids and derived valuations from the two-person auctions. The true density of valuations lies at 0.5 and goes from 0 to 1. Here the estimated density is a little higher and goes over its bounds. However, part of the reason may be the method we are using to represent the density in the figure.[9] You should try different values of epsilon to see how that changes things.

[9]The kernel density method assumes the distribution can be approximated as a mixture of normal distributions.

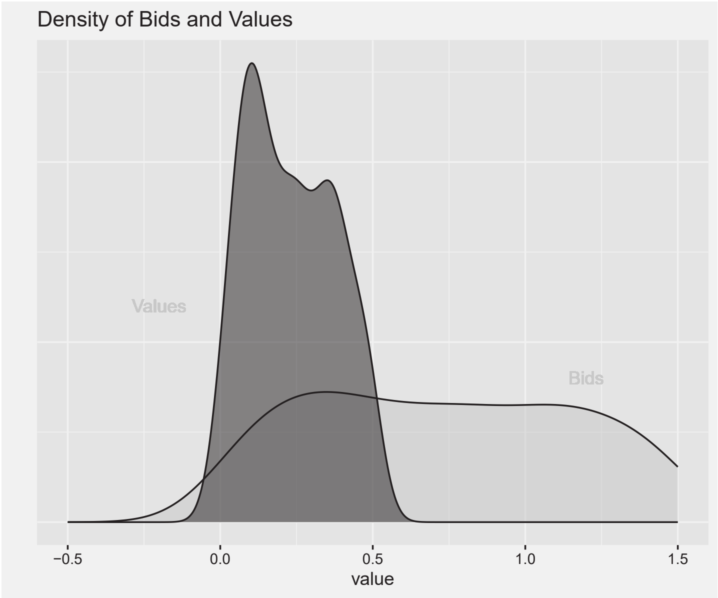


FIGURE 10.1
Plot of the density function for bids and values from 2 person auctions. The true distribution of valuations has a density of 0.5 from 0 to 1.

10.3 English Auctions

The auction format that people are most familiar with is the English auction. These auctions are used to sell cattle, antiques, collector stamps, and houses (in Australia). In the 1970s, they were also the standard format used by the US Forestry Service to sell timber access (Aryal et al., 2018).

Because bidders can observe each other's bid as the auction progresses, it is a dynamic game. To make our life a lot simpler, we will lean heavily on Vickrey's analysis and model these auctions as second price sealed bid auctions.

The section presents the game under the assumption of a second price sealed bid auction and determines the Bayes Nash equilibrium. It then switches focus to estimating valuations from English auctions. Often in these auctions, we only observe the winning bid, or the price, and possibly the number of bidders. Because of this limitation on the data we will use an idea of an order statistic and a result from Athey and Haile (2002).

10.3.1 Second Price Auction Game

The game is a second price sealed bid auction with N bidders and valuations v_i.

- Players: N bidders with valuation v_i
- Strategies: Each bidder i chooses a bid given their valuation for the item, $b_i(v_i)$.
- Payoffs:
  - $b_i > b_j \forall j \neq i$: $v_i - b_2$, where b_2 is the bid of the second highest bidder.
  - $b_i < b_j \forall j \neq i$: 0
- Beliefs: $v_i \sim F$

In addition to assuming away the dynamics, we assume away bid increments that may make our life complicated.

10.3.2 Bayes Nash Equilibrium

Vickrey showed that English auctions are strategically very simple. Imagine a bidder hires an expert auction consultant to help them bid in an English auction.

- Expert: "What is your value for the item?"
- Bidder: "$2,300"
- Expert: "Bid up to $2,300 and then stop."

In sealed bid auctions, there is an optimal trade-off between winning and profiting from the auction. In second price auctions, there is no such trade-off.

The optimal bid for bidder i is to bid her value.

$$b_i = v_i \tag{10.10}$$

Equation (10.10) suggests that empirical analysis of English auctions is a lot simpler than for sealed bid auctions. If only that were so! To be clear, the "bid" in Equation (10.10) means the strategy described by the expert. In the data, we do not necessarily observe this strategy.

If we could observe all the bid strategies in the auction, then we would have an estimate of the value distribution. But that tends to be the problem. Depending on the context, not all active bidders in the auction may actually be observed making a bid. In addition, if the price jumps during the auction we may not have a good idea when bidders stopped bidding (Haile and Tamer, 2003).

Athey and Haile (2002) provide a solution. They point out that the price in an English auction has a straightforward interpretation. When valuations follow Assumption 4, the price is the second highest valuation of the people who bid in the auction. Consider the case when the price is lower than the second highest valuation. How could that be? Why did one of the bidders exit the auction at a price lower than her valuation? What if the price is higher than the second highest valuation? How could that be? Why would a bidder bid more than her valuation?

If the price is equal to the second highest valuation, then it is a particular order statistic of the value distribution. Athey and Haile (2002) show how the observed distributions of an order statistic uniquely determine the value distribution.

10.3.3 Order Statistics

To understand how order statistics work, consider the problem of determining the distribution of heights of players in the WNBA. The obvious way to do it is to take a data set on player heights and calculate the distribution. A less obvious way is to use order statistics.

In this method, data are taken from a random sample of teams, where for each team, the height of the tallest player is measured. Assume each team has 10 players on the roster and you know the height of the tallest, say the center. This is enough information to estimate the distribution of heights in the WBNA. We can use the math of order statistics and the fact that we know both the height of the tallest and we know that 9 other players are shorter. In this case we are using the tallest, but you can do the same method with the shortest or the second tallest, etc.

The price is more or less equal to the second highest valuation of the bidders in the auction.[10] The probability of the second highest of N valuations is equal to some value b which is given by the following formula:

$$\Pr(b^{(N-1):N} = b) = N(N-1)F(b)^{N-2}f(b)(1-F(b)) \tag{10.11}$$

The order statistic notation for the second highest bid of N is $b^{(N-1):N}$. We can parse this equation from right to left. It states that the probability of seeing a price equal to b is the probability that one bidder has a value greater than b. This is the winner of the auction and this probability is given by $1 - F(b)$, where $F(b)$ is the cumulative probability of a bidder's valuation less than b. This probability is multiplied by the probability that there is exactly one bidder with a valuation of b. This is the second highest bidder who is assumed to bid her value. This is represented by the density function $f(b)$.[11] These two values are multiplied by the probability that the remaining bidders

[10] Officially, the price may be a small increment above the bid of the second highest bidder. We will ignore this possibility.

[11] This is the derivative of $F(b)$.

have valuations less than b. If there are N bidders in the auction then $N-2$ of them have valuations less than the price. The probability of this occurring given Assumption 4 is $F(b)^{N-2}$. Lastly, the labeling of the bidders is irrelevant so there are $\frac{N!}{1!(N-2)!} = N(N-1)$ possible combinations. If the auction has two bidders, then the probability of observing a price p is $2f(p)(1-F(p))$.

The question raised by Athey and Haile (2002) is whether we can use this formula to determine F. Can we use the order statistic formula of the distribution of prices to uncover the underlying distribution of valuations? Yes.

10.3.4 Identifying the Value Distribution

Let's say we observe a two bidder auction with a price equal to the lowest possible valuation for the item; call that v_0. Actually, it is a lot easier to think about the case where the price is slightly above the lowest possible value. Say that the price is less than $v_1 = v_0 + \epsilon$, where ϵ is a "small" number. What do we know? We know that one bidder has that very low valuation, which occurs with probability equal to $F(v_1)$. What about the other bidder? The other bidder may also have a value equal to the lowest valuation or they may have a higher valuation. That is, their value for the item could be anything. The probability of value lying between the highest and lowest possible value is 1. So $\Pr(p \leq v_1) = 2 \times 1 \times F(v_1)$ and either bidder could be the high bidder. There are 2 possibilities, so we must multiply by 2. As the probability of a price less than v_1 is observed in the data, we can rearrange things to get the initial probability, $F(v_1) = \Pr(p \leq v_1)/2$.

Now take another value, $v_2 = v_1 + \epsilon$. The probability of observing a price between v_1 and v_2 is as follows.

$$\Pr(p \in (v_1, v_2]) = 2(F(v_2) - F(v_1))(1 - F(v_1)) \tag{10.12}$$

It is the probability of seeing one bidder with a value between v_2 and v_1 and the second bidder with a value greater than v_1. Again, the two bidders can be ordered in two ways.

We can solve $F(v_2)$ using Equation (10.12). We observe the quantity on the left-hand side and we previously calculated $F(v_1)$.

For a finite subset of the valuations, we can use this iterative method to calculate the whole distribution. More generally, we would use differential equations. For this to work, each bidder's valuation is assumed to be independent of the other bidders and comes from the same distribution of valuations (Assumption 4).

10.3.5 English Auction Estimator

The non-parametric estimator of the distribution follows the logic above.

The initial step determines the probability at the minimum value,

$$\hat{F}(v_1) = \frac{\sum_{j=1}^{M} \mathbb{1}(p_j \leq v_1)}{2M} \tag{10.13}$$

where there are M auctions and p_j is the price in auction j.

To this initial condition, we can add an iteration equation.

$$\hat{F}(v_k) = \frac{\sum_{j=1}^{M} \mathbb{1}(v_k < p_j \leq v_{k+1})}{2M(1 - \hat{F}(v_{k-1}))} + \hat{F}(v_{k-1}) \tag{10.14}$$

These equations are then used to determine the distribution of the valuations.

10.3.6 English Auction Estimator in R

We can estimate the distribution function non-parametrically by approximating it at $K = 100$ points evenly distributed across the range of observed values. The estimator is based on Equations (10.13) and (10.14).

```
> f_English_2bid = function(price, K=100, epsilon=1e-8) {
+   # K number of finite values.
+   # epsilon small number for getting the probabilities
+   # calculated correctly.
+   min1 = min(price)
+   max1 = max(price)
+   diff1 = (max1 - min1)/K
+   Fv = matrix(NA,K,2)
+   min_temp = min1 - epsilon
+   max_temp = min_temp + diff1
+   # determines the boundaries of the cell.
+   Fv[1,1] = (min_temp + max_temp)/2
+   gp = mean(price > min_temp & price < max_temp)
+   # price probability
+   Fv[1,2] = gp/2  # initial probability
+   for (k in 2:K) {
+     min_temp = max_temp - epsilon
+     max_temp = min_temp + diff1
+     Fv[k,1] = (min_temp + max_temp)/2
+     gp = mean(price > min_temp & price < max_temp)
+     Fv[k,2] = gp/(2*(1 - Fv[k-1,2])) + Fv[k-1,2]
+     # cumulative probability
+   }
+   return(Fv)
+ }
```

10.4 Empirical Analysis: Testing the Rationality of Loggers using R

In the 1970s, the US Forest Service conducted an interesting experiment. It introduced sealed bid auctions in 1977. Previous to that, most US Forest Service auctions had been English auctions.[12] In 1977, the service mixed between auction formats. As discussed above, bidding in sealed bid auctions is strategically a lot more complicated than bidding in English auctions. In the latter, the bidder simply bids her value. In the former, she must trade off between bidding higher and increasing the likelihood of winning against paying more if she does win.

Because of the experiment, we can test whether the loggers in the sealed bid auctions bid consistently with their actions in the English auctions. Our test involves estimating the underlying value distribution using bid data from sealed bid auctions and comparing that to an estimate of the underlying value distribution using price data from English auctions. These two value distributions are the same under the assumptions of the game theory model.

10.4.1 Timber Data

The data used here are from the US Forest Service downloaded from Phil Haile's website.[13]

In order to estimate the distributions of bids and valuations it is helpful to "normalize" them so that we are comparing apples to apples. The standard method is to use a log function of the bid amount and run a linear regression on various characteristics of the auction including the number of acres bid on, the estimated value of the timber, access costs and characteristics of the forest and species (Haile et al., 2006).[14]

The code brings in the data and uses `lm()` to create the object `lm1`. The regression creates dummy variables for the different tree species, regions, forests and districts using `as.factor()`. It then creates a normalized bid using the residuals from the linear regression.

```
> file = paste0(dir, "auctions.csv")
> df = read.csv(file)
> lm1 = lm(log_amount ~ as.factor(Salvage) + Acres +
```

[12] You may think of this as just some academic question. But the US Senator for Idaho, Senator Church, was not happy with the decision. "In fact, there is a growing body of evidence that shows that serious economic dislocations may already be occurring as a result of the sealed bid requirement." See *Congressional Record* September 14 1977, p. 29223.

[13] http://www.econ.yale.edu/~pah29/timber/timber.htm. The version used here is available from here: https://sites.google.com/view/microeconometricswithr/table-of-contents.

[14] Baldwin et al. (1997) discuss the importance of various observable characteristics of timber auctions.

```
+             Sale.Size + log_value + Haul +
+             Road.Construction + as.factor(Species) +
+             as.factor(Region) + as.factor(Forest) +
+             as.factor(District), data=df)
> # as.factor creates a dummy variable for each entry under the
> # variable name.  For example, it will have a dummy for each
> # species in the data.
> df$norm_bid = NA
> df$norm_bid[-lm1$na.action] = lm1$residuals
> # lm object includes "residuals" term which is the difference
> # between the model estimate and the observed outcome.
> # na.action accounts for the fact that lm drops
> # missing variables (NAs)
```

In general, we are looking for a normal-like distribution. Figure 10.2 presents the histogram of the normalized log bids. It is not required that the distribution be normal, but if the distribution is quite different from normal, you should think about why that may be. Does this distribution look normal?[15]

10.4.2 Sealed Bid Auctions

In order to simplify things, we will limit the analysis to two-bidder auctions. In the data, sealed bid auctions are denoted “S”.

```
> df1 = df |>
+   filter(
+     num_bidders == 2 &
+       Method.of.Sale == "S"
+   ) |>
+   mutate(
+     bids = as.vector(norm_bid),
+     values = f_sealed_2bid(norm_bid)
+   )
> summary(df1$bids)
    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 -4.0262 -0.7987 -0.1918 -0.2419  0.3437  5.0647
> summary(df1$values)
     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
  -4.0262  -0.0218   0.9185   4.8704   2.3108 860.0647
```

Using the same method that we used above, it is possible to back out an estimate of the value distribution from the bids in the data. We see that

[15]It is approximately normal, but it is skewed somewhat to lower values. This may be due to low bids in the English auction. How does the distribution look if only sealed bids are graphed?

```
> df |>
+   ggplot() +
+   geom_histogram(aes(x = norm_bid),
+                  fill = "gray",
+                  alpha = 0.5) +
+   scale_x_continuous(limits = c(-3,2.5)) +
+   labs(
+     x = "bid",
+     y = "",
+     title = "Histogram of Bids"
+   ) +
+   theme(axis.text.y = element_blank(),
+         axis.ticks.y = element_blank())
```

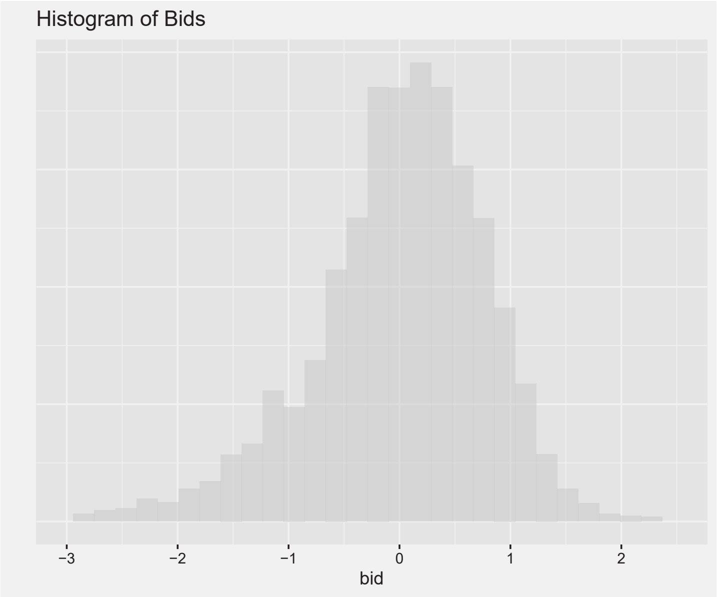


FIGURE 10.2
Histogram of normalized bid residual for US Forest Service auctions from 1977.

comparing the valuations to the bids, the bids are significantly shaded particularly for higher valuations. The negative numbers may seem odd but remember we have normalized the bids in the auction.

10.4.3 Comparing English Auctions to Sealed Bid Auctions

We can back out the value distribution from the English auctions by assuming that the price is the second highest bid, the second highest **order statistic**. The English auctions are denoted “A”.

The function `f_English_2bid` estimates the cumulative probability function for the value distribution.

```
> df2 = df |>
+   filter(
+     num_bidders == 2 &
+       Method.of.Sale == "A" &
+       Rank == 2
+   )
> Fv_english = f_English_2bid(df2$norm_bid)
```

To calculate the equivalent for the sealed bid auctions, we can use the `ecdf()` function.

```
> Fv_sealed = ecdf(df1$values)
```

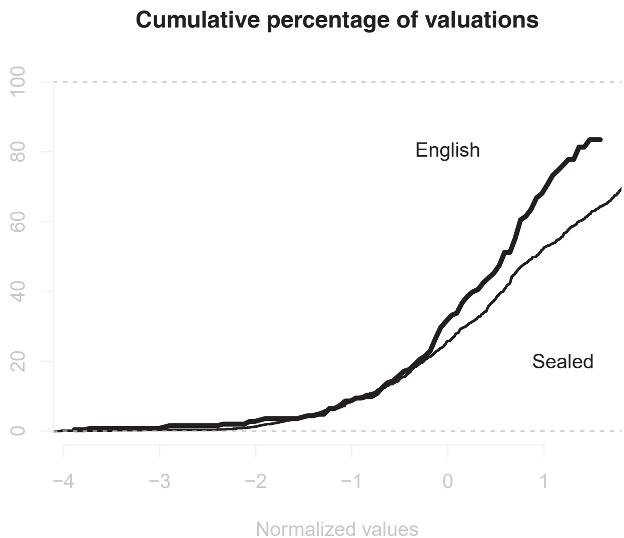


FIGURE 10.3
Comparison of estimated distributions from two bidder English and sealed bid auctions. The estimate from the English auction and the sealed bid auction are similar to about 0, then estimate from English auctions places more weight on lower valuations than the estimate from the sealed bid auctions.

Figure 10.3 shows that there is not a whole lot of difference between the estimate of the distribution of valuations from sealed bid auctions and English auctions. The two distributions of valuations from the sealed bid auctions and

English auctions lie fairly close to each other, particularly for lower values. This suggests loggers are bidding rationally. That said, at higher values, the two distributions diverge. The value distribution from the sealed bid auctions suggests that valuations are higher than the estimate from the English auctions. What else may explain this divergence?

10.5 Empirical Analysis: Testing for Collusion using R

Is there evidence that bidders in English auctions are colluding? This section presents a test of collusion based on using auction theory to back out the implied value distribution of the bidders in the larger English auctions. We can compare the implied distribution to the distribution we have estimated above.

10.5.1 A Test of Collusion

Consider the following test of collusion. Using large English auctions, we can estimate the distribution of valuations. Under the prevailing assumptions of the game theory model, this estimate should be the same as for two-bidder auctions. If the estimate from the large auctions suggests valuations are much lower than for two-bidder auctions, this suggests collusion.

Specifically, if the inferred valuations in these larger auctions look much like auctions with fewer bidders. That is, bidders may behave "as if" there are actually fewer bidders in the auction. For example, if there is an active **bid ring**, bidders may have a mechanism for determining who will win the auction and how the losers may be compensated for not bidding.[16] In an English auction, it is simple to enforce a collusive agreement because members of the bid ring can bid in the auction where their bids are observed.

Can we determine the size of the bid ring? How many people are bidding collusively? What if we have an auction with six bidders? If three of them are members of a bid ring, then those three will agree on who should bid from the ring. Only one of the members of the bid ring will bid their value. If we estimate the model under the assumption that there are four independent bidders, it will match the value distribution we estimated from the two bidder auction.[17]

[16] Asker (2010) presents a detailed account of a bid ring in stamp auctions.

[17] In the bid ring mechanism discussed in Asker (2010), the collusion actually leads to higher prices in the main auction.

10.5.2 "Large" English Auctions

Assume there are six bidders in the auction. From above, the order statistic formula for this case is as follows.

$$\Pr(b^{5:6} = b) = 30F(b)^4 f(b)(1 - F(b)) \tag{10.15}$$

As above, order statistics are used to determine the underlying value distribution (F); however, in this case, it is a little more complicated to determine the starting value.

Think about the situation where the price in a 6 bidder auction is observed at the minimum valuation. What do we know? As before, one bidder may have a value equal to the minimum or a value above the minimum. That is, their value could be anything. The probability of a valuation lying between the minimum and maximum value is 1. We also know that the five other bidders had valuations at the minimum. If not, one of them would have bid more and the price would have been higher. As there are six bidders, there are six different bidders that could have had the highest valuation. This reasoning gives the following formula for the starting value.

$$\Pr(b^{5:6} < v_1) = 6F(v_1)^5 \tag{10.16}$$

Rearranging, we have $F(v_1) = \left(\frac{\Pr(p<v_1)}{6}\right)^{\frac{1}{5}}$.

Given this formula we can use the same iterative method as for two-bidder auctions to solve for the distribution of valuations.

10.5.3 Large English Auction Estimator

Again we can estimate the value distribution by using an iterative process. In this case, we have the following estimators.

$$\hat{F}(v_1) = \left(\frac{\sum_{j=1}^{M} \mathbb{1}(p_j < v_1)}{6M}\right)^{\frac{1}{5}} \tag{10.17}$$

and

$$\hat{F}(v_k) = \frac{\sum_{j=1}^{M} \mathbb{1}(v_k < p_j < v_{k+1})}{30M\hat{F}(v_{k-1})^4(1 - \hat{F}(v_{k-1}))} + \hat{F}(v_{k-1}) \tag{10.18}$$

The other functions are as defined in the previous section.

We can also solve for the implied distribution under the assumption that there are three bidders and under the assumption that there are two bidders.[18] Note in each auction there are at least six bidders.[19]

[18] See Equation (10.11) for the other cases.

[19] For simplicity, it is assumed that all of these auctions have six bidders. Once there are a large enough number of bidders in the auction, prices do not really change with more bidders. In fact, these methods may not work as the number of bidders gets large (Deltas, 2004).

10.5.4 Large English Auction Estimator in R

We can adjust the estimator above to allow any number of bidders, N.

```
> f_English_Nbid = function(price, N, K=100, epsilon=1e-8) {
+   min1 = min(price)
+   max1 = max(price)
+   diff1 = (max1 - min1)/K
+   Fv = matrix(NA,K,2)
+   min_temp = min1 - epsilon
+   max_temp = min_temp + diff1
+   Fv[1,1] = (min_temp + max_temp)/2
+   gp = mean(price > min_temp & price < max_temp)
+   Fv[1,2] = (gp/N)^(1/(N-1))
+   for (k in 2:K) {
+     min_temp = max_temp - epsilon
+     max_temp = min_temp + diff1
+     Fv[k,1] = (min_temp + max_temp)/2
+     gp = mean(price > min_temp & price < max_temp)
+     Fv[k,2] =
+       gp/(N*(N-1)*(Fv[k-1,2]^(N-2))*(1 - Fv[k-1,2])) +
+       Fv[k-1,2]
+   }
+   return(Fv)
+ }
```

10.5.5 Evidence of Collusion

We limit the auctions to ones with more than five bidders.

```
> df3 = df |>
+   filter(
+     num_bidders > 5 &
+       Method.of.Sale == "A" &
+       Rank == 2
+   )
```

We can estimate the value distribution under the assumption that there are six bidders in the auction. We can also estimate the value distribution under the assumption that there are three bidders and two bidders. The results are presented in Figure 10.4.

```
> Fv_6 = f_English_Nbid(df3$norm_bid, N = 6)
> Fv_3 = f_English_Nbid(df3$norm_bid, N = 3)
> Fv_2 = f_English_2bid(df3$norm_bid)
```

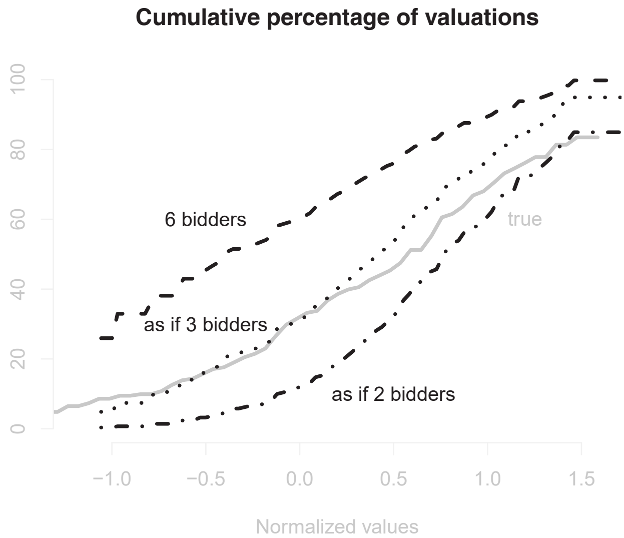


FIGURE 10.4
Comparison of estimated distribution of valuations from English auctions with at least 6 bidders. These estimates are compared to the estimate from 2 bidder auctions which is labeled "true." The estimate of the 6-bidder auctions suggests valuations are much lower than for the 2-bidder auctions.

Figure 10.4 suggests that there is in fact collusion in these auctions! Assuming there are 6 bidders in the auction implies that valuations are much lower than we estimated for 2 bidder auctions from both English auctions and sealed bid auctions. In the chart, the distribution function is shifted to the left, meaning there is greater probability of lower valuations.

In theory, these two estimates should be the same or very close. Remember we are estimating the underlying value distribution which is unrelated to the number of bidders in the auction. The estimate with 6 bidders suggests that bidders value the timber much lower than when there are 2 bidders. If we don't think there is any collusion in 2 bidder auctions, then these estimates provide the ground truth. These estimates tell us how the timber is valued. This implies that the reason the value distribution is lower is not that the values are lower, but that the bids are lower.

If there are 5 bidders in the ring, then bidders will bid as if they are in a 2 bidder auction. We can estimate the value distribution assuming that there are 2 bidders in the auction. The result is that the estimated value distribution is higher than the ground truth. Either the 2 bidders are bidding too much or there are more than 2 independent bidders in the auction.

If there are 4 bidders in the ring, then bidders will bid as if they are in a 3 bidder auction. Estimates assuming there are three bidders and two bidders lie above and below the true value, respectively. This suggests that bidders are behaving as if there are between two and three bidders in the auction. This implies that the bid ring has between four and five bidders in each auction.

These results are suggestive of an active bid ring in these auctions in 1977. It turns out that this was of real concern. In 1977, the United States Senate conducted hearings into collusion in these auctions. In fact, this may be why the US Forestry Service looked into changing to sealed bid auctions. The US Department of Justice also brought cases against loggers and millers (Baldwin et al., 1997). Alternative empirical approaches have also found evidence of collusion in these auctions, including Baldwin et al. (1997) and Athey et al. (2011).

10.6 Discussion and Further Reading

Economic analysis of auctions began with Vickrey's 1961 paper. Vickrey used game theory to analyze sealed bid auctions, Dutch auctions and English auctions. Vickrey also derived a new auction, the sealed bid second price auction.

The chapter considers two of the most important auction mechanisms, sealed bid auctions and English auctions. The sealed bid auctions are analyzed using the two-step procedure of Guerre et al. (2000). The first step uses non-parametric methods to estimate the bid distribution. The second step uses the Nash equilibrium to back out the value distribution.

While we observe all the bids in the sealed bid auctions, we generally only observe the high bids in English auctions. The chapter uses the order statistic approach of Athey and Haile (2002) to estimate the value distribution from these auctions.

Baldwin et al. (1997) and Athey et al. (2011) analyze collusion in US timber auctions. Aryal et al. (2018) uses US timber auctions to measure how decision makers account for risk and uncertainty.

11

Auctions with Affiliated Valuations

11.1 Introduction

The previous chapter assumed that bidder's valuations are independent of each other. This substantially simplified both the game theory and the econometrics. In this chapter we are going to relax this assumption. Now we will allow for bidder valuations to be dependent. These types of auctions are called **affiliated private values auctions**.

The conical example is bidding on the right to drill for oil in the US's outer continental shelf (OCS). The amount of oil that can be pulled out of the ground has nothing to do with the bidder. Similarly, the price of that oil has little or nothing to do with the identity of the bidder.[1]

The chapter analyzes **common value auctions** and takes the game theory model to OCS auctions. The chapters asks whether bid rings should be allowed in these auctions.

11.2 Auctions with Common Values

Here we consider auctions where the bidder's valuations are interdependent. The simplest case of such interdependence is the pure **common value auction**. In this auction, bidder's don't know the exact value of the item they are bidding on. Each bidder draws a signal about the true value of the item such that if all the signals were aggregated, then that would provide a pretty close approximation of the true value of the item.

The classic example is bidders considering how much to bid for the rights to drill for oil in a particular area. The oil field has a particular amount of oil worth a particular value. While unknown, it has nothing to do with who is bidding. The value of the oil field is the same to all bidders. They just don't

[1] Of course, some bidders may have market power in the oil market.

DOI: 10.1201/9781003351603-11

know what it is. Oil companies, the bidders, will hire different geologists with different models, equipment, and expertise to estimate the value of the oil field. These geologists will come up with different guesses. We call these guesses, signals of the oil field's value.

The section presents the game and the **Bayes Nash equilibrium** of the auction. It discusses an idea called the **winner's curse**, which leads bidders to shade their bid down in equilibrium. The section presents an estimator for **common value auctions** and illustrates the estimator using simulated data.

11.2.1 Simple Model

The auction game has N bidders with signals s_i. The signals are drawn from a normal distribution with mean v and variance σ^2. The true value of the item is v. Each bidder's bid is a function of their signal, $b_i(s_i)$. The bidder with the highest bid wins the auction and pays their bid. If the bidder wins, they get the true value of the item, v. If the bidder loses, they get nothing.

- Players: N bidders and signals (types) $s_i \in \Re$.
- Strategies: Each bidder i observes a signal (s_i) and chooses a bid $b_i(s_i)$.
- Payoffs:
  - $b_i > b_j \forall j \neq i$: $v - b_i(s_i)$
  - If $\exists j$ s.t. $b_i < b_j$: 0
- Beliefs: $s_i \sim \mathcal{N}(v, \sigma^2)$, where v is the true value of the item.

If the player i wins, they get v which doesn't have a subscript. This is because it is the same for every bidder. What varies is between bidders, is their signal s_i. Bidders know that their signal is drawn from a normal distribution with mean v but they do not know v.

11.2.2 Winner's Curse

The winner is the bidder with the highest bid. Consider what happens if all bidders bid their valuations. In that case, the winner bids $s^{N:N}$ (using order statistic notation) which is going to be substantially higher than v.

```
> set.seed(123456789)
> N = 10
> v = 5
> s = rnorm(N, v)


> max(s)
  [1] 6.415538
```

In the example with 10 bidders, the winner bids 6.42, which much higher than the true value 5. The winner really loses. The bidders should adjust their strategy by shading their bid down in order to account for the winner's curse.

11.2.3 Bayes Nash Equilibrium

We often see common value auctions that are also **first price auctions**. Bidders are shading their bids for two different reasons. First, as in the previous chapter, they are shading their bid because it is a first price auction and they need to account for winning and paying what they bid. Second, these bidders have to account for the fact that if they win the auction it is because their signal is higher than everyone else's and thus higher than the true value of the item.

To make things simpler, let's assume that bidders bid their expected value for the item. Assume that bidders use **Bayes rule** and their expected value is conditional on both the signal that they observe and on the fact that they won the auction. Of course, they don't actually know whether they win the auction or not, but their bid is of no importance if they lose. The bidder is thinking through the cases. The only case of interest is the case where the bidder won the auction. They realize that if they won the auction, their bid must, by definition, have been the highest bid.

Again, trading off winning and the amount paid is left out for the moment. So how much information does the bidder have? You may think that they don't have much because all they have is one signal, but it turns out that they know a lot more than that. Because they won the auction, they know that their signal must be higher than everyone else's. So that tells them quite a lot about every body else's signal.

Chapter 10 illustrates how order statistics work. The probability that a particular signal s is the highest order statistic ($s^{N:N}$) is given by the following equation assuming that all the signals are drawn independently from the distribution $F(s|\mu, \sigma)$. The probability that a particular s is higher than all the other signals is then $G(s|N, \mu, \sigma) = F(s|\mu, \sigma)^{N-1}$ where μ represents the different possible values of the true value v and there are $N-1$ other signals. The derivative is then $g(s|N, \mu, \sigma) = (N-1)f(s|\mu, \sigma)F(s|\mu, \sigma)^{N-2}$.

We can use this likelihood to determine the bidder's expected value for the item conditional on submitting the winning bid. We just need to use Bayes rule. To determine the expected value, we need the probability that a particular distribution is generating the signals we observe.

$$\gamma(\mu, \sigma|s = s^{N:N}) = \frac{g(s|N, \mu, \sigma)}{\sum_{\mu', \sigma'} g(s|N, \mu', \sigma')} \tag{11.1}$$

If we know all the possible μs and σs and assume a normal distribution and assume that the prior over μs and σs is uniform, then Equation (11.1) gives the probability that the observed signals are generated by a particular μ and σ.

Consider a simple version where $\mu = v$, the true value and $\sigma = 1$. In this case $g(s|N,v) = (N-1)\phi(s-v)\Phi(s-v)^{N-2}$ and $\phi()$ and Φ represent the standard normal's density and probability function, respectively.

The expected value of the item given that the bidder has the highest signal, $s = s^{N:N}$, is then given by following function.

$$\mathbb{E}(v|N,s) = \int_{v'} v'\gamma(v'|s = s^{N:N})d(v') \tag{11.2}$$

where $\gamma()$ is defined in Equation (11.1). Given N bidders and a signal s the expected value of the true bid is the integral over the possible true values weighted by the $\gamma()$ function assuming that s is the highest signal. Remember $\sigma = 1$.

Now we have how much the bidder values the item, the next question is how much to bid. Remember this is a sealed bid auction, so it is still the case that the bidder is trading off the value of winning against the probability of winning.

$$b_i(s_i) = \mathbb{E}(v|N,s_i) - \frac{G(b_i|N,s_i)}{g(b_i|N,s_i)} \tag{11.3}$$

So we need to determine the probability of winning conditional on the bidder's signal. The simplest assumption to make is that equilibrium bids are monotonically increasing in the signal. This doesn't seem unreasonable.

Given our monotonicity assumption, the probability of winning is just the probability of having the highest signal. We have that $G(b_i|N,s_i) = F(s_i|v)^{N-1}$ and $g(b_i|N,s_i) = (N-1)f(s_i|v)F(s_i|v)^{N-2}$, which not coincidentally is similar function that we defined previously.

11.2.4 Common Value Auction in R

Below we create all the probability functions that we need for bidding in first price auctions with common values. It is a little confusing because we are using order statistics for two different things. First, there is the standard method discussed in the previous chapter where the bidder is determining the optimal bid for a first price auction. Second, the bidder is using order statistics to back out their expectation of the value for the item given the signal they observed and conditional upon winning the auction.

The signal distribution is a normal distribution denoted F, where f is the density. We use `log_F()` and `log_f()`. Given `N` bidders, the distribution of the highest signal is denoted by G and we use `log_G()` and `log_g()` for the density. The function `dnorm()` calculates the density of the normal distribution and `pnorm()` calculates the probability of the normal distribution.

```
> log_f = function(s, v, sigma = 1)
+   log(dnorm(s, v, sigma))
> log_F = function(s, v, sigma = 1)
+   log(pnorm(s, v, sigma))
```

```
> log_G = function(s, v, sigma = 1, N) (N-1)*log_F(s, v, sigma)
> log_g = function(s, v, sigma=1, N) log(N-1) +
+   log_f(s, v, sigma) +
+   (N-2)*log_F(s, v, sigma)
```

Given these probabilities, we can determine the bidder's expected value for the item and their bid. The expectation function `E_fun()` takes in two global variables `u` and `sig`. These are the possible parameters of the normal distribution determining the signals observed by the bidders. Lastly, there is the bid function `b_fun()` based on Equation (11.3).

```
> E_fun = function(s, N) {
+   g_u = matrix(NA,length(u),length(sig))
+   u_mat = matrix(NA, length(u), length(sig))
+   for(j in 1:length(sig)) {
+     g_u[,j] = exp(log_g(s, u, sig[j], N))
+     u_mat[,j] = u
+   }
+   sum_g_u = sum(g_u)
+   gamma_u = g_u/sum_g_u
+   mu = sum(u_mat*gamma_u)
+   sigma = sqrt(sum(u_mat^2*gamma_u) - mu^2)
+   return(list(mu=mu, sigma=sigma))
+ }
> b_fun = function(s, N) {
+   v_bar = E_fun(s, N)
+   G = exp(log_G(s, v_bar$mu, v_bar$sigma, N))
+   g = exp(log_g(s, v_bar$mu, v_bar$sigma, N))
+   return(v_bar$mu - G/g)
+ }
```

11.2.5 Simulation of Common Value Auction using R

It helps to work through a simulation. There are 100 auctions and the number of bidders varies. The true value is 0 for each auction, and the signal is distributed standard normal. The function `seq()` calculates a sequence of numbers between the first and second values with the third value as the step. The function `rnorm()` generates random numbers from a normal distribution. The function `sample()` randomly samples from a set of numbers. The function `rep()` repeats a number a certain number of times.

The code uses `sapply()` to loop through the signals and calculate the expected value and bid for each signal.

```
> set.seed(123456789)
> M = 100
```

```
> N = NULL
> bids = NULL
> ids = NULL
> values = NULL
> u = seq(-10, 10, 0.15)
> sig = seq(0.1, 3, 0.15)
> Ns = sample(3:4, M, replace=TRUE)
> v = rep(0, M)
> sigma = 1
> for(i in 1:M) {
+   ids = c(ids, rep(i, Ns[i]))
+   N = c(N, rep(Ns[i], Ns[i]))
+   s_i = rnorm(Ns[i], v[i], sigma)
+   values = c(values,
+              sapply(1:length(s_i),
+                             function(j) E_fun(s_i[j],
+                                                  Ns[i])$mu))
+   bids = c(bids,
+            sapply(1:length(s_i),
+                         function(j) b_fun(s_i[j],
+                                              Ns[i])))
+ }
```

The code below creates a density plot of the bids and the expected values. The bids are shifted down from the expected values. The expected values are shifted down from the signal distribution in equilibrium.

```
> ggplot_sim_cv_bids = data.frame(
+   bids = bids,
+   values = values
+ ) |>
+   ggplot(aes(bids)) +
+   geom_density(alpha = 0.5) +
+   geom_density(aes(values), linetype = 2, alpha = 0.5) +
+   labs(
+     x = "values/bids",
+     y = "",
+     title = "Density of bids and values"
+   ) +
+   geom_vline(xintercept = 0, linetype = 2,
+              color = "gray") +
+   geom_text(aes(x = -5, y = 0.2, label = "bids"),
+             color = "gray") +
+   geom_text(aes(x = 2, y = 0.2, label = "values"),
+             color = "gray") +
```

```
+   theme(axis.text.y=element_blank(),
+          axis.ticks.y=element_blank())
```

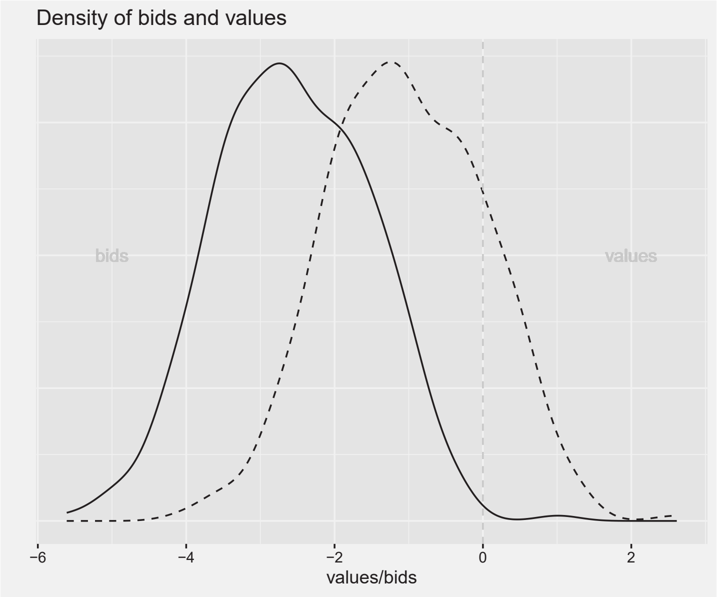


FIGURE 11.1
Plot of the density of the bids and expected values in a first price common values auction. Bids are shifted down from valuations because it is a first price auction.

Figure 11.1 presents the observed bids and the estimated values. As we saw in the IPV case, bids are significantly shaded down from values for first price auctions. Expected values are also significantly shaded down from the original signal observed by the bidder. These expected values are conditional on the signal assuming that signal is highest signal in the Bayes Nash equilibrium. The actual signal is distributed around zero.

11.2.6 Estimator for Common Values Auctions

The estimator reverse engineers the signal distribution from the bid distribution. We know from Equation (11.3) and the discussion in Chapter 10 that we can identify the distribution of the expected values conditional on the signal. Unfortunately, it not generally possible to uniquely determine the signal distribution from the expected value distribution. We will need to rely on parametric restrictions.

11.2.7 Common Values Estimator in R

To estimate the underlying signal distribution from the observed bids, we will combine maximum likelihood with simulation. The estimator chooses the μ and σ of the signal distribution that maximizes the likelihood of the observed bids given the game theory assumption that bidders bid their expected value conditional on their signal being the highest.

The estimator works by taking a set of parameter values for the distribution of signals, `mu` and `sigma`, and simulating the resulting bids, `b_sim`. It simulates the signals, then loops through the simulated signals and creates the corresponding simulated bids. It then calculates the log likelihood of observing the observed bids (`bids_temp`) given the derived parameters from the bid distribution. It does this for each size of auction in the data.

```
> f_bid_ml = function(mu, sigma, bids_temp, N_temp, s) {
+   Ns = unique(N_temp)
+   log_lik = rep(NA, length(bids_temp))
+   for(i in 1:length(Ns)) {
+     N_i = Ns[i]
+     index = which(N_temp==Ns[i])
+     b_sim = sapply(1:length(s), function(i) b_fun(s[i], N_i))
+     mu_i = mean(b_sim, na.rm = TRUE)
+     sigma_i = sd(b_sim, na.rm = TRUE)
+     z_i = (bids_temp[index] - mu_i)/sigma_i
+     log_lik[index] = log(dnorm(z_i)) - log(sigma_i)
+   }
+   return(log_lik)
+ }
> f_bid_ml_int = function(par, bids_temp, N_temp) {
+   set.seed(123456789)
+   mu = par[1]
+   sigma = exp(par[2])
+   s = U*sigma + mu
+   return(-sum(f_bid_ml(mu, sigma, bids_temp, N_temp, s)))
+ }
```

This estimator requires three global variables `U`, `u`, and `sig`.

```
> U = rnorm(1000)
> a = optim(par = c(0, log(sigma)), f_bid_ml_int,
+            bids_temp = bids, N_temp = N,
+           control = list(trace=0, maxit=100000))
```

The code below estimates the signal distribution from the observed bids.

```
> ggplot_sim_cv_signals =
+   data.frame(
```

```
+     signals = rnorm(length(values),
+                     a$par[1],
+                     exp(a$par[2])),
+     values = values
+   ) |>
+     ggplot(aes(values)) +
+     geom_density(alpha = 0.5) +
+     geom_density(aes(signals), linetype = 2, alpha = 0.5) +
+     labs(
+       x = "values/signals",
+       y = "",
+       title = "Density of signals and values"
+     ) +
+     geom_vline(xintercept = 0, linetype = 2, color = "gray") +
+     geom_text(aes(x = -3.5, y = 0.2, label = "values"),
+               color = "gray") +
+     geom_text(aes(x = 2.5, y = 0.2, label = "signals (est.)"),
+               color = "gray") +
+     theme(axis.text.y=element_blank(),
+           axis.ticks.y=element_blank())
```

```
> ggplot_sim_cv_signals
```

Figure 11.2 presents the density of the expected values and the estimated signals in the simulated data. The estimated signals are pretty close to the true distribution, which is a standard normal distribution. The figure shows that the bidders substantially discount their bids from the observed signals. We see that this occurs for two reasons. First, their expected value (conditional on winning) is discounted from their signal. Second, because it is a first price auction, bidders discount their bid from the expected value of the item (see Figure 11.1).

11.3 Empirical Analysis: Signal Distribution from OCS Auctions using R

This section uses data on outer continental shelf (OCS) oil and gas tracts off Texas and Louisiana from 1954 to 1979.[2]

[2]These data are available from Penn State https://capcp.la.psu.edu/data-and-software/outer-continental-shelf-ocs-auction-data/.

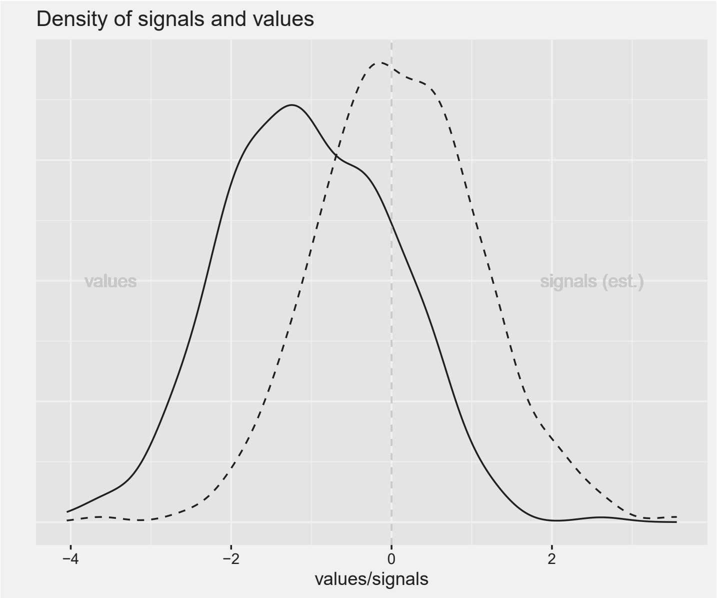


FIGURE 11.2
Plot of the density of expected values and signals (estimated) in a first price common values auction. The bidder's valuations are shifted down from the estimated signal distribution.

11.3.1 Data

The code brings in the data set, `book_ocs_ch11.csv`. Next we do the trick of creating a residual auction value by regressing bids on observed characteristics of the auctions. The function `as.factor()` is used to create dummy variables for the block code and date of the auction. The function `lm()` is used to estimate the linear regression. The residuals are then calculated, and the data set is converted to a data frame.

```
> file = paste0(dir, "book_ocs_ch11.csv")
> df = read.csv(file) |>
+   select(
+     lbid,
+     lvalue,
+     lcost,
+     BlockCode,
+     Date,
+     TractNumber,
+     nCompany
+   ) |>
```

```
+   na.omit()
> lm1 = lm(lbid ~ lvalue + as.factor(BlockCode) +
+             as.factor(Date) + lcost, data = df)
> df$res = lm1$residuals
> dt = setDT(df)

> ggplot_ocs_bids =
+   df |>
+   ggplot(aes(res)) +
+   geom_density(alpha = 0.5) +
+   geom_density(aes(rnorm(length(res),
+                        mean(res),
+                        sd(res))),
+                linetype = 2, alpha = 0.5) +
+   scale_x_continuous(limits = c(-4,4)) +
+   labs(
+     x = "residuals",
+     y = "",
+     title = "Density of residuals"
+   ) +
+   geom_vline(xintercept = 0, linetype = 2, color = "gray") +
+   geom_text(aes(x = 2, y = 0.2, label = "residuals"),
+             color = "gray") +
+   geom_text(aes(x = -3, y = 0.2, label = "normal"),
+             color = "gray") +
+   theme(axis.text.y=element_blank(),
+         axis.ticks.y=element_blank())

> ggplot_ocs_bids
```

Figure 11.5 presents the normalized bids for the OCS auctions. The figure also shows simulated bids from a normal distribution to suggest that a normal distribution is a reasonable approximation.

11.3.2 Estimating the Signal Distribution

We restrict the sample to just those auctions without coalitions in them.[3] We can use the estimator above to estimate the signal distribution for these auctions. In addition, the code creates an index of observations that have less than 3 bidders, more than 10 bidders and missing residuals. The code then calculates the initial values for the optimization routine and runs the routine. The code creates an object `index` that determines the auctions with less than 3 bidders and more than 10 bidders. It then uses `-index` to drop those auctions.

[3]Coalitions are discussed in detail in the next section. They are legal bid rings.

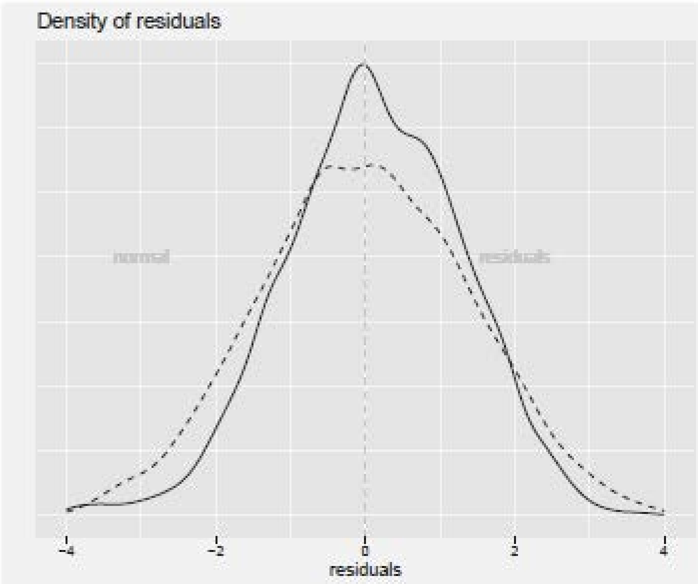


FIGURE 11.3
Plot of the density of the residual bids against a simulated data set drawn from a normal distribution with the mean and variance equal to the mean and variance for the normalized bids. This indicates that the normal distribution is a reasonable approximation of the bids.

```
> dt2 = dt[num_coy == N]
> index = c(which(dt2$num_coy < 3 | dt2$num_coy > 10),
+           which(is.na(dt2$res)))
> init = c(mean(dt2$res[-index]), sd(dt2$res[-index]))
> b1 = optim(par = init,
+            f_bid_ml_int,
+            bids_temp = dt2$res[-index],
+            N_temp = dt2$num_coy[-index],
+            control = list(trace = FALSE,
+                           maxit = 1000000))
```

The code then creates a plot of the estimated signal distribution and the observed bids.

```
> ggplot_est_cv_signals =
+   data.frame(
+     bids = dt$res,
+     signals = rnorm(length(dt$res),
+                     b1$par[1],
```

```
+                      exp(b1$par[2]))
+   ) |>
+     ggplot(aes(bids)) +
+     geom_density(alpha = 0.5) +
+     geom_density(aes(signals), linetype = 2, alpha = 0.5) +
+     scale_x_continuous(limits = c(-5,8)) +
+     labs(
+       x = "bids/signals",
+       y = "",
+       title = "Density of signals and bids"
+     ) +
+     geom_vline(xintercept = 0, linetype = 2,
+                color = "gray") +
+     geom_text(aes(x = -3.5, y = 0.2, label = "bids"),
+               color = "gray") +
+     geom_text(aes(x = 6.5, y = 0.2, label = "signals (est.)"),
+               color = "gray") +
+     theme(axis.text.y=element_blank(),
+           axis.ticks.y=element_blank())

> ggplot_est_cv_signals
```

Figure 11.4 presents the density of the normalized bids and the estimated signals in the OCS auctions. The figure shows that the bidders substantially discount their bids from the observed signals. We see that this occurs for two reasons. First, their expected value (conditional on winning) is discounted from their signal because of winner's curse. Second, because it is a first price auction, bidders discount their bid from the expected value of the item.

11.4 Auctions with Coalitions

The classic data set for considering common values auctions is the US federal government's off-shore drilling auctions. One surprising fact is that the auctions include "coalitions" of bidders. Basically, legal bid rings. Given the discussion about bid rings in the Chapter 10, it seems very odd that the government would allow collusion in these auctions.

An obvious policy question is whether the government should in fact allow bid rings in OCS auctions. You are probably thinking that the answer is obviously no. Actually it is not that obvious in the case of common value auctions.

The section works through how bidding in coalitions can be estimated. It then uses the estimated parameters from above and some characteristics of OCS auctions to simulate the policy.

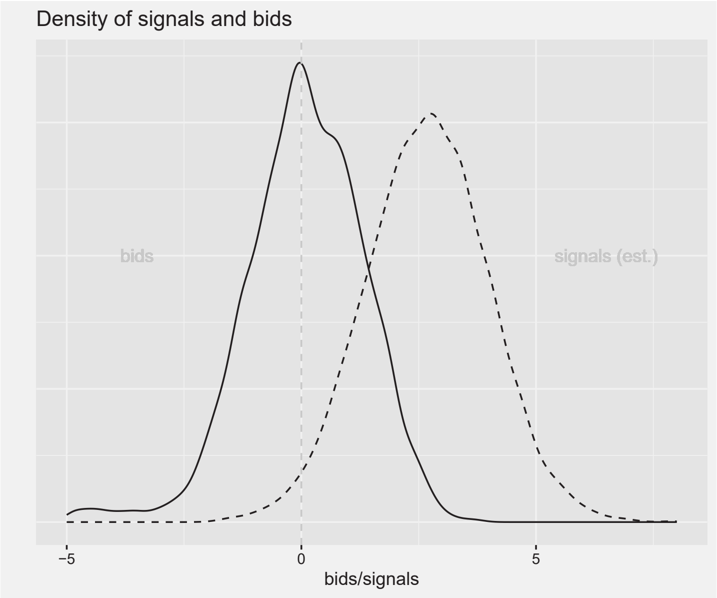


FIGURE 11.4
Plot of the density of normalized bids from all auctions and signals (estimated) in OCS auctions without coalitions that have between 3 and 10 bidders. The bidder's valuations are shifted down from the estimated signal distribution.

11.4.1 The Benefit of Coalitions

The reason for allowing coalitions is that coalitions allow bidders to pool their information about the item's value. By increasing the amount of information available to the bidders, the coalitions may lead to higher bids!

Remember we said that there are two reasons for bidders shading their bids. The first, discussed in Chapter 10, states that bidders shade in order to account for the probability of winning and trade off the probability of winning against how much they pay if they win. As the number of bidders increases, the probability of any particular bidder winning decreases, and so the smaller trade off leads to higher bids. As we said previously, a bid ring allows bidders to reduce their bids because their probability of winning is higher. The second reason for shading the bids is because of the information problem. Here the bid ring works in the opposite direction, by pooling information the bidders have a more precise signal of the value of the item which allows them to bid more.

How does a bidder's valuation change with more signals? Assume that the expected valuation for the ring will be the mean of the signals, conditional upon that mean being greater than all the other signals. From statistics, we know that we can approximate the distribution of this sample mean as a normal distribution with a mean equal to the true mean and the variance equal to true variance divided by the sample size.

For the J members of the coalition, the probability of winning the auction with a particular average of signals ($\bar{s}$) is as follows.

$$\Phi\left(\frac{\bar{s}-\mu}{\sigma}\right)^{N-J-1} \tag{11.4}$$

where this gives the probability that the J members of the coalition will observe a particular average of their signals multiplied by the probability that the other bidders outside the coalition will have signals below coalition's average.

What about for the bidders outside the coalition?

$$\Phi\left(\frac{s-\mu}{\sigma}\right)^{N-J-1} \Phi\left(\frac{s-\mu}{\frac{\sigma}{J}}\right) \tag{11.5}$$

The probability is the probability that a signal s is observed and is greater than all the other bidders outside the coalition and greater than average of the signals that are in the coalition.

11.4.2 Estimating Coalitions in R

The probabilities with coalitions allowed are the following. For bidders in the coalition, their signal has less noise than for bidders outside the coalition. For all bidders, the number of independent bidders is lower. The code uses `_in` to refer to the bidders in the coalition and `_out` for the bidders outside the coalition.

```
> log_G_in = function(s, v, sigma=1, N, J)
+   (N-J-1)*log_F(s, v, sigma/sqrt(J))
> log_g_in = function(s, v, sigma=1, N, J) log(N-J-1) +
+   log_f(s, v, sigma/sqrt(J)) +
+   (N-J-2)*log_F(s, v, sigma/sqrt(J))
> log_G_out = function(s, v, sigma=1, N, J)
+   (N-J-1)*log_F(s, v, sigma)
> log_g_out = function(s, v, sigma=1, N, J) log(N-J-1) +
+   log_f(s, v, sigma) + (N-J-2)*log_F(s, v, sigma)
```

The expected values and bids for bidders in and out of the coalition are as you would expect.

```
> E_in = function(s, N_i, J_i) {
+   g_u = matrix(NA,length(u),length(sig))
+   u_mat = matrix(NA, length(u), length(sig))
+   for(j in 1:length(sig)) {
+     g_u[,j] = exp(log_g_in(s, u, sig[j], N_i, J_i))
+     u_mat[,j] = u
```

```
+   }
+   sum_g_u = sum(g_u)
+   gamma_u = g_u/sum_g_u
+   mu = sum(u_mat*gamma_u)
+   sigma = sqrt(sum(u_mat^2*gamma_u) - mu^2)
+   return(list(mu=mu, sigma=sigma))
+ }
> E_out = function(s, N_i, J_i) {
+   g_u = matrix(NA,length(u),length(sig))
+   u_mat = matrix(NA, length(u), length(sig))
+   for(j in 1:length(sig)) {
+     g_u[,j] = exp(log_g_out(s, u, sig[j], N_i, J_i))
+     u_mat[,j] = u
+   }
+   sum_g_u = sum(g_u)
+   gamma_u = g_u/sum_g_u
+   mu = sum(u_mat*gamma_u)
+   sigma = sqrt(sum(u_mat^2*gamma_u) - mu^2)
+   return(list(mu=mu, sigma=sigma))
+ }
> b_in = function(s, N, J) {
+   v_bar = E_in(s, N, J)
+   G = exp(log_G(s, v_bar$mu, v_bar$sigma, N-J+1))
+   g = exp(log_g(s, v_bar$mu, v_bar$sigma, N-J+1))
+   return(v_bar$mu - G/g)
+ }
> b_out = function(s, N, J) {
+   v_bar = E_out(s, N, J)
+   G = exp(log_G(s, v_bar$mu, v_bar$sigma, N-J+1))
+   g = exp(log_g(s, v_bar$mu, v_bar$sigma, N-J+1))
+   return(v_bar$mu - G/g)
+ }
```

11.4.3 Policy Simulation

In the mid-1970s, the Department changed the policy to make illegal for larger bidders to join forces, but that still allowed small bidders to join with big bidders or with other small bidders.

Assume that auctions with coalitions have the same signal distribution as auctions without coalitions. This allows us to use the estimates from the previous section above in the policy simulations.

This analysis is restricted to cases where there is just one coalition and there are more than two bidders. When the number of bids is smaller than the number of bidders, we have coalitions in the auction. This information is then merged back into the main data set. The code uses the estimates of `mu` and `sigma` from the previous section to simulate bids in and outside of the coalition. The number of bidders in the auction is from the data. It simulates the auctions allowing for a coalition of bidders and if the coalition is not allowed.

```
> dt1 = dt[, .(num = .N,
+              num_coy = sum(as.numeric(nCompany))),
+          by = TractNumber]
> dt = merge(dt, dt1, by="TractNumber")
```

```
> dt2 = dt[num_coy - num == 1 & num > 2]
> M = length(unique(dt2$TractNumber))
> N = dt2$num_coy
> mu = rep(b1$par[1], M)
> sigma = exp(b1$par[2])
> bids_sim = NULL
> bids_cf = NULL
> set.seed(123456789)
> for(i in 1:M) {
+   N_i = N[i]
+   s_i = rnorm(N_i, mu, sigma)
+   bids_i = sapply(1:N_i, function(j) b_fun(s_i[j], N_i))
+   bids_i_in = b_in(mean(s_i[1:2]), N_i, 2)
+   bids_i_out = sapply(1:(N_i-1), function(j)
+     b_out(s_i[j], N_i, 2))
+   bids_sim = c(bids_sim, bids_i_in, bids_i_out)
+   bids_cf = c(bids_cf, bids_i)
+   # print(i)
+ }
```

The code then creates a plot of the density of the bids in auctions with and without coalitions.

```
> ggplot_ocs_bids =
+   data.frame(
+     rings = bids_sim,
+     no_rings = bids_cf
+   ) |>
+     filter(
+       is.finite(rings) & is.finite(no_rings)
+     ) |>
+     ggplot(aes(rings)) +
+     geom_density(alpha = 0.5) +
+     geom_density(aes(no_rings), linetype = 2, alpha = 0.5) +
+     labs(
```

```
+       x = "Normalized bids",
+       y = "",
+       title = "Density of bids"
+     ) +
+     geom_text(aes(x = 3, y = 0.2, label = "rings"),
+               color = "gray") +
+     geom_text(aes(x = -3, y = 0.2, label = "no rings"),
+               color = "gray") +
+     theme(axis.text.y=element_blank(),
+           axis.ticks.y=element_blank())

> ggplot_ocs_bids
```

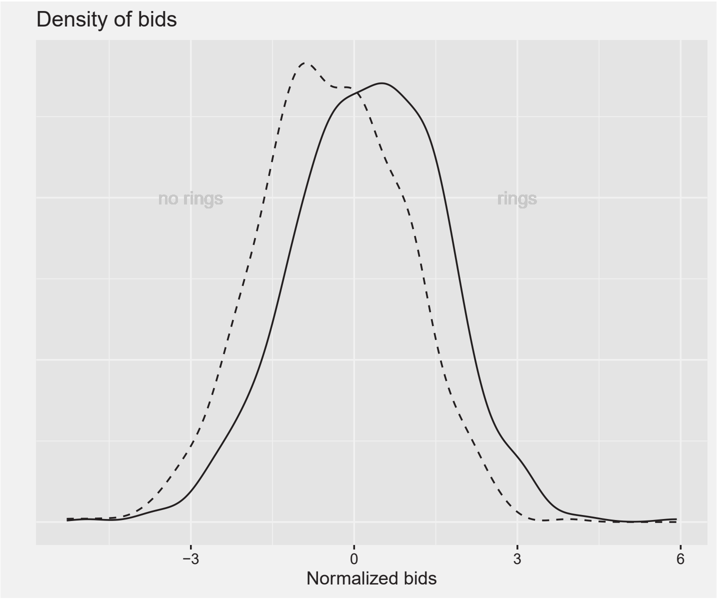


FIGURE 11.5
Plot of the density of the amount bid where coalitions are not allowed ("no rings") and coalitions are allowed ("rings"). The value of the bids have been normalized. The bids are higher when coalitions are allowed in these OCS auctions

Figure 11.5 shows that allowing bid rings (coalitions) tends to lead to higher bids! This analysis accounts for the fact that bidders will bid lower because the number of independent bidders is lower. Despite that, the bids are higher showing the advantage of aggregating signals in common values auctions.

11.5 Discussion and Further Reading

Chapter 10 made a simplifying assumption called independent private values (IPV). This assumption rules out the common values model. Laffont and Vuong (1996) present the main negative result of the common values literature. It states that without strong parametric assumptions it is not possible to identify the exact model generating the data in this setting. Despite this negative result, we could test for whether the auction is a common values auction (Haile et al., 2006).

The analysis in this chapter uses a parametric model. The chapter suggests that allowing cooperation among competitors may lead to better outcomes for the government in the sale of oil drilling rights. See Paarsch and Hong (2006) for more detailed analysis of the econometrics of auctions, including common value auctions. The OCS data used here have been analyzed in a number of papers, for example, Hendricks et al. (2003).

Part IV

Dynamic Games of Incomplete Information

12

Moral Hazard

12.1 Introduction

The book has four parts covering complete information static games, complete information dynamic games, incomplete information static games, and incomplete information dynamic games. This part is the last.

This part considers two types of dynamic games of incomplete information. The first are games in which one player does not observe the other player's action before making their choice. What makes these games dynamic is that the first player's action is associated with a signal that is observed by the second player before making his choice. The second set of games are ones where the action of the first player is observed by the second player, but the second player does not know the first player's payoffs. The second set of games are discussed in the next chapter.

The first set of games are often given the moniker, moral hazard or principal-agent games. Consider a game between an employer and employee, where the principal is the employer and the agent is the employee. The employee takes actions like working hard and thinking carefully, but the employer doesn't get to observe this. The employer may observe signals of these things like the number of hours worked or profits of the firm. In these games, we are generally interested how the employer can get the employee to take actions that the employer prefers even when those actions are not directly observed.

This chapter presents the principal-agent game and applies it to the problem faced by financiers of whaling expeditions in New England in the 1830s. A new, at the time, form of company organization, the corporation, allowed many more people to invest in whaling. Ventures funded through corporations tended to give managers less of a stake in the outcome of the venture and tended to have worse outcomes.

DOI: 10.1201/9781003351603-12

12.2 Principal-Agent Game

In the 1830s, whaling was an important industry in the United States, particularly in communities in New England. The industry developed to hunt and kill sperm whales for oil and other products. The ships sailed literally around the world and the voyages would last years. To hunt and kill whales, you need a ship with the necessary equipment, supplies, a captain, and a crew of around 30. Obviously you need money for all of this. The investment will payoff when you sell the oil and other products rendered from the whales.

How do investors insure that they get a good return on their investment? How do they know that they have a good captain and crew? How do they know that the captain and crew are doing a good job when they are literally on the other side of the world and it is the 1830s?

These operations generally relied on a number of incentive mechanisms. The captain and crew were paid a small share of the profits. The investors hired an agent who was responsible for hiring the captain and crew, determining the voyage's route and communicating with the captain during the voyage. Under the law, the owner of the vessel controlled everything on the vessel including any equipment used. Because of this, investors generally bought an ownership share of the vessel itself.

Two organizational structures were used to finance these hunts. In one case a small number of investors hired an agent. The agent was generally given a large share of the ownership of the vessel and the investors may include family members or members of the local community. An alternative method for raising funds from investors was to create a corporation. These organizations provided some legal recourse for investors and made it clearer who owned what. Corporate whaling enterprises had a much larger set of owners investing smaller amounts of money. Like unincorporated ventures, they still hired an agent who was responsible for overseeing the operation. One big difference between the two organizations seems to be share of ownership given to the agent. For the corporations the evidence suggests that the agent earned a small fraction of the profits.

This section works through a formal game of the investor–agent relationship in 1830s whaling. It then goes through an example with simulated data.

12.2.1 Simple No Contract Game

Consider a simplified version of the game faced by whaling ventures. In this game, the agent chooses her effort level and the investor pays her based on what the outcome they observe.

- Players: Investor, Agent
- Strategies:

  - Agent: $e \in \{e_L, e_H\}$
  - Nature: $y \in \{y_L, y_H\}$, where $y_H > y_L$, as a function of e, $\Pr(y_H|e_H) = p_H > \Pr(y_H|e_L) = p_L$
  - Investor: $w(y) \in \{w_L, w_H\}$

- Payoffs:

  - Agent: $w(y) - c(e)$, where $c(e_H) = c$ and $c(e_L) = 0$
  - Investor: $y - w(y)$

- Beliefs: $p = \Pr(e = e_H)$

In this game, the agent moves first and chooses how much "effort" to put into the venture. If they choose the high effort level e_H, then it costs them c, while the low effort level cost them nothing. Nature observes the effort level and chooses the outcome of the venture. The higher effort level increases the probability of the better outcome y_H. Lastly, the Investor observes the choice of Nature and pays the Agent w. The more she pays the agent, the worse off the investor is.

What is the Bayes Nash equilibrium of the game? Is there an equilibrium where the Agent chooses e_H? What is the subgame Perfect Nash equilibrium?

The last is a trick question. There is only one subgame, the whole game.[1]

The Agent's strategy is to choose an effort level e_L or e_H. The Investor's strategy is to choose a payment level given the observed outcome y, $w(y)$ and the Investor's beliefs about the choice of the Agent.

12.2.2 Bayes Nash Equilibrium

One proposed equilibrium is for the Agent to choose the high effort level e_H and for the Investor choose a payment that pays more if the outcome is y_H than if the outcome is y_L. Such payments could make it worthwhile for the Agent to choose the higher cost effort level because that effort level increases the probability of getting a higher payment.

This is not an equilibrium. It is not optimal for the Investor to pay the Agent anything. What ever effort level the Agent chooses, the Investor prefer not to pay the Agent anything.

Assume the equilibrium is for the Agent to choose e_H and the Investor pays the Agent $w(y_H) = w_H$ and $w(y_L) = w_L$ where the following inequality holds.

$$p_H w_H + (1 - p_H) w_L - c \geq p_L w_H + (1 - p_L) w_L \tag{12.1}$$

[1] A **perfect Bayesian Nash equilibrium** has a similar flavor to a subgame perfect Nash equilibrium. In a perfect Bayesian Nash equilibrium, the players' strategies must be optimal given their beliefs and the beliefs must be consistent with Bayes' rule at each information set along each equilibrium path that can be reached with positive probability.

The expected payoff from choosing the high effort level is higher than the expected payoff from choose the low effort level. We call this the incentive compatibility (IC) constraint.

If in the proposed equilibrium the payment $w(y)$ is such that the IC holds, then the Agent will choose the high effort level. What if Agent chooses e_H. What contract should the Investor offer? Let w^* denote the Investor's payment if they observe y_H. In equilibrium, the Investor's expected profits are as follows.

$$p_H(y_H - w^*) + (1 - p_H)(y_L - w_L) \tag{12.2}$$

The Investor should offer $w^* = 0$.

In the proposed equilibrium, the Agent chooses e_H and the Investor must believe that is true. Given all of that, why should the Investor pay the Agent anything?

12.2.3 Simple Contract Game

The problem in the previous game is that the Investor cannot commit to paying the Agent for their high effort level. The result is that the Agent is not willing to put in the effort. We need a contract. In general, we assume that such a contract can be enforced, for example in a court of law. Alternatively, it could be enforced in the court of public opinion or the court of the back alleys of the New England port town. If the Investor doesn't pay the Agent what is stated in the contract then the Investor gets punished. Making the contract enforceable limits what is contractable. We cannot contract on the effort level of the Agent because that is not observed by anyone but the Agent, certainly not in a court of law. Potentially, we can contract on the outcome (y) as that is observed. We will assume that the y is both observable and verifiable by a court of law.

- Players: Investor, Agent
- Strategies:
  - Investor: Offer contract $w(y)$
  - Agent: Accept or Reject given offer $w(y)$
  - Agent: Given Accept of $w(y)$ choose $e \in \{e_L, e_H\}$
  - Nature: Choose $y \in \{y_L, y_H\}$ given e, $\Pr(y_H|e_H) = p_H > \Pr(y_H|e_L) = p_L$
- Payoffs: If the contract is accepted:
  - Investor: $\mathbb{E}(y(e) - w(y(e)))$
  - Agent $\mathbb{E}(w(y(e))) - c(e)$

  Assume that both get 0 if rejected.
- Beliefs: $p = \Pr(e = e_H)$

In the Bayes Nash equilibrium, the Investor offers a contract such that the following equality holds.

$$p_H w_H + (1 - p_H)w_H - c = 0 \tag{12.3}$$

The Agent is indifferent between their expected wage if they choose the high effort level and their outside option (which is assumed to be 0). We call this the individual rationality (IR) constraint. In this set up, w_L will be negative, but you should just think of this as a value lower than Agent's outside option or alternative if they reject the contract.

Also, the Agent prefers the high effort level to the low effort level. The IC is as follows.

$$p_H w_H + (1 - p_H)w_L - c \geq p_L w_H + (1 - p_L)w_L \tag{12.4}$$

The difference between pay given the good outcome w_H and pay given the bad outcome w_L is large enough to induce the Agent to choose the high effort level despite the higher cost c.

The Investor then chooses w_H and w_L such that both the IR and IC constraints hold. Moreover, for it to be an equilibrium, it must be that $p_H(y_H - w_H) + (1 - p_H)(y_L - w_L) \geq 0$. We know from the IR constraint that it must be that the following in equality holds for the contract to be profitable.

$$p_H y_H + (1 - p_H)y_L - c \geq 0 \tag{12.5}$$

So the contract is profitable if the expected output from the project is greater than the cost to the Agent of completing the project.

12.2.4 More Complicated Contract Game

Let's model the principal-agent problem associated with investors contracting with an agent to run the whaling venture.

- Players: Investor, Agent
- Strategies:
  - Investor: Offer $w(y)$
  - Agent: Accept or Reject offer
  - Agent: Choose $e \geq 0$.
  - Nature: Choose y given e, $y \sim F(e)$
- Payoffs
  - Agent: $\mathbb{E}(u(w(y))|e) - c(e)$
  - Investor: $\mathbb{E}(y(e) - w(y(e)))$
- Beliefs: $e \sim P$.

The Agent chooses an effort level, e, which incurs some cost $c(e)$. The contract is based on the outcome, y, where this is a random number drawn from a distribution that is determined by Agent's effort choice $F(e)$. Based on this outcome, the Agent is paid $w(y)$. One difference between this model and what we have seen previously is that the Agent's payoff is determined by their utility function $u(y)$. Here we are going to assume that the agent is risk-averse while the principal is risk-neutral. The principal would like to incentivize the agent by giving them a large share of the outcome. This implies that the agent is taking on a lot of risk. The principle will need to compensate the agent by paying them a very large amount that does not vary with the outcome of the venture. We will talk more about this in a bit.

12.2.5 Bayes Nash Equilibrium

One of the issues highlighted in this book is the difference between assumptions about the game and assumptions about equilibrium. This issue comes up again here. Look at the payoffs for the Investor. Her expected payoff is not conditional on the effort level of the Agent. In a Bayes Nash equilibrium, we assume that the Investor's beliefs are consistent with equilibrium behavior. In equilibrium, the Agent will choose a particular effort level e^*, so *in equilibrium* the Investor's expected payoff will be $\mathbb{E}(y(e) - w(y(e))|e = e^*)$.

The Agent will choose this effort level to optimize his expected payoff.

$$\max_e \quad \mathbb{E}(u(w(y(e)))) - c(e) \tag{12.6}$$

The Agent knows e and so the cost function is not inside the expectation.

The Investor will choose $w(y)$ so as to maximize her profits.

$$\begin{array}{ll} \max_{w(y)} & \mathbb{E}(y(e) - w(y(e))|e = e^*) \\ s.t. & \mathbb{E}(u(w(e^*))) - c(e^*) \geq 0 \\ & e^* = \arg\max_e \mathbb{E}(u(w(y(e)))|e) - c(e) \end{array} \tag{12.7}$$

In words, the Investor will choose a payment as a function of the outcome $w(y)$, that maximizes her expected return subject to the Agent choosing to accept the contract and choosing their optimal effort level conditional on the payment offer.

12.2.6 Parameterized Model

Assume that the Agent's expected utility is a mean-variance utility. That is, the Agent's utility is increasing in the mean of his payoffs and decreasing in the variance of his payoffs. The agent's dislike of risk is represented by the parameter r.

$$\mathbb{E}(u(x)) = \mu_x - r\frac{\sigma_x^2}{2} \tag{12.8}$$

where $x \sim \mathcal{N}(\mu_x, \sigma_x^2)$. This utility function is a substantial simplification. It is often justified using a particular utility function and assuming the outcome is normally distributed.

The Agent doesn't like producing effort and the cost of the effort is governed by the parameter k and the cost is increasing in effort at a geometric rate, $c(e) = -k\frac{e^2}{2}$. For a particular incentive rate, b, the Agent's expected utility is as follows.

$$b(\mu_x + e) - r\frac{b^2\sigma_x^2}{2} - k\frac{e^2}{2} \tag{12.9}$$

The variance of bx is $b^2\sigma_x^2$. Also the dividing by 2 thing is useful once we get to the first-order condition.

$$\begin{aligned} &b - ke = 0 \\ &e = \tfrac{b}{k} \end{aligned} \tag{12.10}$$

The agent's optimal effort level is increasing in the power of the incentives and decreasing in the cost.

The production function is simply that $y(e) \sim \mathcal{N}(\mu + e, \sigma)$. The mean of output is increasing 1 to 1 with the effort level e.

The incentive contract used by the Investor is a linear function of output, $w(y) = a + by$, where a is a constant and b is the fraction of the output received by the Agent. When you look below you see there is no a. This is because we have implicitly solve the individual rationality constraint by choosing a such that the Agent is indifferent between accepting the contract and rejecting the contract. We can also substitute in the Agent's optimal effort level.

$$\pi_I(b) = \mu + \frac{b}{k} - \frac{rb^2(\sigma^2)}{2} - \frac{b^2}{2k} \tag{12.11}$$

The first-order condition is then given by the following equation.

$$\begin{aligned} &\tfrac{1}{k} - rb\sigma^2 - \tfrac{b}{k} = 0 \\ &1 - krb\sigma^2 - b = 0 \\ &b = \tfrac{1}{1+kr\sigma^2} \end{aligned} \tag{12.12}$$

The power of the incentive contract is decreasing in the effort cost of the Agent, the risk aversion of the Agent and the variance in the output.

12.2.7 Simulation with R

In the code, the function in Equation (12.11) is as follows. The parameters `mu` and `sigma` are global variables determined below.

```
> Pi_I = function(b, r, k) {
+   mu + b/k - r*(b^2)*(sigma^2)/2 - (b^2)/(2*k)
+ }
```

The parameters of the simulation have been calibrated such that the incentive contract is similar to the average for unincorporated ventures in the data.

```
> mu = 90000
> sigma = 40000
> k = 0.000005
> r = 0.0002
```

Given the set up we can solve for the optimal share of the output given to the agent and the optimal effort level.

```
> b1 = optimize(Pi_I, c(0, 1), maximum = TRUE, r=r, k=k)
> b1$maximum
  [1] 0.3846154
> b1$maximum/k
  [1] 76923.08
> Pi_I(b1$maximum, r, k)
  [1] 128461.5
```

The firm makes $128,000 from the venture with the agent taking 39%. The agent's effort cost is $77,000.

Now consider what happens if the power of the incentive contract is significantly reduced. Let $b = 0.05$, rather than the optimal level.

```
> 0.05/b1$maximum
  [1] 0.13
> Pi_I(0.05, r, k)/Pi_I(b1$maximum, r, k)
  [1] 0.7733832
```

The new level of optimal effort for the Agent is 13% of the optimal contract and the Investor's profits also falls but to just 77% of what they would be with the optimal contract.

12.3 Empirical Analysis: Whaling Corporations in the 19th Century using R

Wellsley College professor, Eric Hilt, documents the surprising failure of the corporate structure in whaling. In the first half of the 19th century, corporations were a relatively new type of institution in the United States. They provided a legal structure for people to raise money from investors where it was clear what rights investors did and did not have. Whaling in New England was generally financed by small groups of investors, many of whom knew each other or were from the same family. Corporations opened up whaling to a much broader range of investors. Given these advantages it is surprising that the corporate structure performed so poorly.

This section using data from whaling venture financial records to understand the contracts with agents used by whaling corporations. Were the contracts responsible for the poor performance of corporations?

12.3.1 Whaling Data

Bringing in the data called `whaling.csv`.

The code below reads in the data and plots average output by year and by how the venture was financed.

```
> file = paste0(dir, "whaling.csv")
> dt = fread(file)
> dt1 = dt[, .(lprod_wb = mean(lprod_wb,
+                                 na.rm = TRUE)),
+           by = .(ayear, corp)]
> ggplot_log_output = setDF(dt1) |>
+   mutate(
+     corp = as.factor(corp)
+   ) |>
+   ggplot(aes(ayear, lprod_wb, corp,
+                   linetype = corp)) +
+     geom_line() +
+     labs(
+       x = "Year",
+       y = "",
+       title = "Average log output by year"
+     ) +
+     theme(axis.text.y=element_blank(),
+           axis.ticks.y=element_blank())


> ggplot_log_output
```

Figure 12.1 presents a plot of the productivity of the whaling ventures. It is the ratio of the output of the venture in terms of the value of the oil and other products rendered from the whale to the size of the ship multiplied by the length of the journey.

The figure shows two things. First, productivity is falling dramatically over time. It is becoming harder and harder to find whales to kill. Second, the corporate ventures tend be less productive than closely held ventures.

Why are corporations doing so poorly?

12.3.2 Regressions

Does the pattern from Figure 12.1 hold when we are more careful about accounting for various factors determining the outcome. One of the big variables

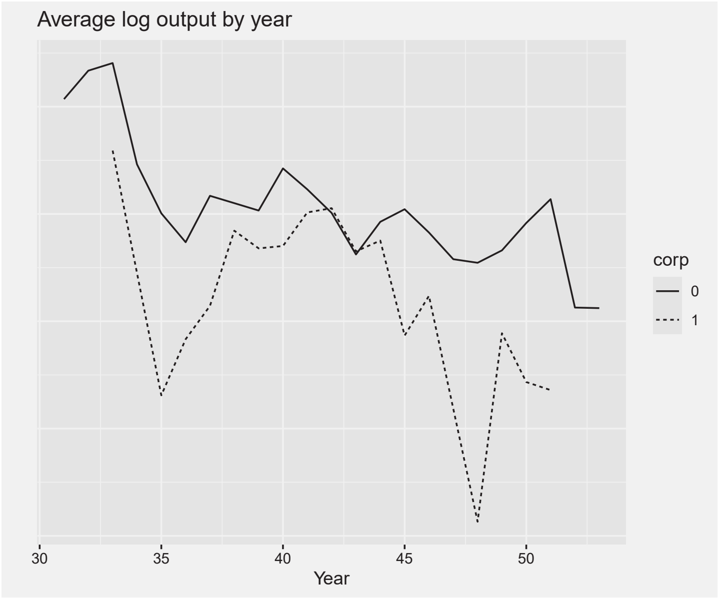


FIGURE 12.1
Line chart of average log of output by year and corporation. It shows output decreasing from 1835 to 1850, but corporate (corp = 1) output is generally lower.

is the agent themselves. The data include cases where the same agent is used by a closely held firm and by a corporation. Using fixed effects, we can account for differences across agents. We can hold the "agent-effect" fixed.

Table 12.1 presents the results from the **fixed effects analysis**. The empirical model is generally called two-way fixed effects. We have fixed effects for the agent (the first way) and fixed effects for the year (the second way) which are not presented in the table. The idea is that we can account for the agent and the year to isolate the effect of the financial structure. It shows that corporations have lower productivity even accounting for individual agent effects. Specifications (3) and (4) suggest that it may not be the corporate entity itself but due to the larger number of owners associated with the corporate structure. It also shows that having the captain die, is not good for the success of the hunt.

Why may this be happening? Corporations provided agents with a share of the output, but shares were substantially lower than what we see for the closely held ventures. What are the implications of this difference in compensation to the agent's effort level and profitability of the venture?

TABLE 12.1
OLS regressions of output on corporate form with agent and year fixed effects. Ownership structure is accounted for either through the dummy variable for corporation or by the number of investors.

	Dependent variable:			
	lprod_wb			
	(1)	(2)	(3)	(4)
corp	−0.455**	−0.411**		−0.279
	(0.178)	(0.179)		(0.442)
I(owners/10)			−0.126*	−0.126*
			(0.065)	(0.065)
I((owners/10)ˆ2)			0.014	0.014
			(0.012)	(0.012)
atlantic		0.048	0.074	0.074
		(0.053)	(0.060)	(0.060)
pacific		−0.107**	−0.102**	−0.102**
		(0.044)	(0.049)	(0.049)
tons		−0.001***	−0.001**	−0.001**
		(0.0003)	(0.0003)	(0.0003)
vesselage		−0.003	−0.004	−0.004
		(0.002)	(0.002)	(0.002)
capexp		0.006	0.008	0.008
		(0.008)	(0.010)	(0.010)
capdied		−0.338***	−0.407***	−0.407***
		(0.100)	(0.121)	(0.121)
Observations	831	809	671	671
R^2	0.300	0.339	0.343	0.343
Note:				*p<0.1; **p<0.05; ***p<0.01

12.3.3 Calibrating the Model

In order to get some sense of what happens when incentives of the agent are changed, we can use the observed data to calibrate the game presented above.

In order to get apples to apples, we can use the same trick that we used in Chapter 10. Regress the output measure on various observed characteristics and then use the residual as the normalized output.

```
> df1 = setDF(dt) |>
+   filter(
+     corp == 0
+   ) |>
+   na.omit()
> lm5 = lm(output_wb ~ atlantic + pacific + tons +
+             vesselage + dyear, data = df1)
> df1$res = lm5$residuals + lm5$coefficients[1]
> mu = mean(df1$res, na.rm = TRUE)
> sigma = sd(df1$res, na.rm = TRUE)
```

Given the normalized output, we can use the observed contracts from the closely held firms to back out the parameters on agent's costs (k) and risk-preferences (r).

```
> e_1 = mean(df1$res, na.rm = TRUE)
> sigma_1 = sd(df1$res, na.rm=TRUE)
> b_1 = mean(df1$agt_shr, na.rm = TRUE)
> k_1 = b_1/e_1
> r_1 = e_1*(1 - b_1)/((b_1^2)*sigma_1^2)
```

```
> Pi_I(b_1, r_1, k_1)
  [1] 269567.9
> e_1*(1 - b_1)
  [1] 73007.96
```

The closely held firm makes \$270,000, and the agent's effort costs are \$73,000. Looking at the corporations, we can again normalize output and estimate the effort level of the agent given the share received by the agent.

```
> df2 = setDF(dt) |>
+   filter(
+     corp == 1
+   ) |>
+   select(
+     output_wb,
+     atlantic,
+     pacific,
```

```
+      tons,
+      vesselage,
+      dyear,
+      agt_shr
+    ) |>
+    na.omit()
> df2$res = df2$output_wb -
+    predict.lm(lm5, df2) +  lm5$coefficients[1]
> e_2 = mean(df2$res, na.rm = TRUE)
> b_2 = mean(df2$agt_shr, na.rm = TRUE)
> e_2_pred = b_2/k_1
```

```
> Pi_I(b_2, r_1, k_1)
  [1] 183437.4
> e_2_pred*(1 - b_2)
  [1] 3718.121
```

In the case of the profit sharing under the corporate form, the venture's profits fall to $183,000 and the agent's effort would be only $3,700.

```
> e_2_pred/e_1
  [1] 0.02094994
> Pi_I(b_2, r_1, k_1)/Pi_I(b_1, r_1, k_1)
  [1] 0.680487
```

Given the share offered to agents working for corporations, the predicted effort level falls to just 2% of the unincorporated effort level. That said, the predicted profits only fall to 68% of the unincorporated profits. The reason is that while the output falls, the share of output received by the investors is higher. In addition, the investors don't have to compensate the agent for risk.

```
> e_2/e_1
  [1] 0.6883835
```

While our model predicts that the Agent's effort level will drop precipitously. The effort level does fall by a large amount to 69% of the unincorporated level. The predicted output is much lower than what we actually observe in the corporate ventures. This discrepancy suggests that corporations are providing incentives. They are just not providing incentives in the form of share of output. They must be using other mechanisms like direct supervision and the threat of firing, rather than giving high powered incentives. Those incentives may leading to significantly lower effort levels by the Agent, but the returns to the Investors are not necessarily that bad.

12.4 Discussion and Further Reading

Wellsley College professor, Eric Hilt, has created an amazing data set on how whaling firms actually worked. It is very cool to bring modern economic theory to bear on why closely held firms worked so well in the 1830s (Hilt, 2006).

Theoretical work on incentives and the principal-agent problem substantially improved our understanding of contracts and performance pay systems (Lazear and Oyer, 2013). The classic paper on the principal-agent problem is Holmström (1979).

13

Adverse Selection

13.1 Introduction

This chapter considers **adverse selection** problems. These games are generally split between **signalling** and **screening** games. In a **signalling game**, the player with an unknown type chooses an action and that action choice may provide information to the other player. In a **screening game**, the uninformed player goes first and offers some choices to the player whose type is not known. The observed choice may provide information to the uninformed player.

In either case, the question is whether the observed action provides information about the player's type. We may have equilibria in which the action provides information about the player's type, we call these **separating equilibria**. We also may have equilibria where the action does not provide information about the player's type. We call these **pooling equilibria**.

The chapter presents the most famous game of unknown types, George Akerlof's model a used car market and the lemons problem. It then uses these ideas to think about the market for health insurance. It considers whether government subsidies or taxes can be used to solve the lemons problem in health insurance. The chapter analyzes these policies with a model calibrated using data from the Medical Expenditure Panel Survey (MEPS).

13.2 Akerlof's Lemons

Outside economics, there is a perception that economists are gungho believers in free market capitalism. For economists that study markets and how markets work or don't work, it is odd to be called "believers." George Ackerlof is an economist through and through, he even won the Nobel prize in economics. He is also married to one of the most influential economists in the world, former Treasury Secretary Janet Yellen. Akerlof showed how a simple misallocation of information in a market, would cause that market to fail.

This section introduces Akerlof's lemons market and provides a formal game-theoretic version of the problem.

DOI: 10.1201/9781003351603-13

13.2.1 The Lemons Market

Consider a car market. In this market, there are two types of cars, good cars and lemons (bad cars). Unfortunately, consumers cannot observe the type of car prior to purchase. They can, however, resell the car on the used car market after observing the car's type.

Now consider a bunch of new cars become available. The key thing about new cars is that the seller of the cars will sell the car no matter if it is good or a lemon. Consumers buy the car and learn the car's type. After the consumers learn the car's type they can sell the car on the used market or keep it. If prices in the used car market are roughly similar to prices in the new car market, then those customers who purchased lemons will dump them on the used car market and go back to the new car market hoping to get a good car.

Customers on the used car market are going to notice that the proportion of lemons just got a lot higher and so the price in the used car market will fall. Where previously, people that owned good cars may have been willing to sell their car on the used car market, the price dropped because of all the lemons entering the market. So these sellers are going to keep their good cars. With owners of lemons dumping them and owners of good cars keeping them, the used car market becomes increasingly filled with lemons. In response, the price in the used car markets drops. This only exacerbates things. As the price falls more, the only cars that are profitable sell in the used car market are lemons. Finally, the used car market is just lemons.

The market fails.

13.2.2 Lemons Game

Akerlof's original argument was not game theory. We can reorient it as a game.

- Players: Seller of car type Bad (b) or Good (g), Buyer
- Strategies:
  - Buyer: Offer price, p.
  - Nature: Determine car type, $\{b, g\}$
  - Seller: Observe car type and offered price, then Sell or Keep car.
- Payoffs:
  - Seller: Sell: $\mathbb{E}(p|Sell)$, Keep: V
  - Buyer: $\mathbb{E}(V|Sell, p) - p$
  - $V \in \{b, g\}$
- Beliefs: $\Pr(V = g) = \pi$

The Buyer offers a price for the used car. The Seller observe her type, actually the car's type. Given that observation and the observed price offered by the Buyer, the seller chooses to sell or keep the car. Again it is important to separate out what happens in the game and what happens in equilibrium. In equilibrium, the Buyer's price needs to adjust for the strategy of the Seller.

13.2.3 Bayes Nash Equilibrium

There is no Bayes Nash equilibrium in which there is a transaction of good cars.

Remember when determining a Bayes Nash equilibrium, we need to update the player's beliefs about the other player's type. Another way to say it, is that beliefs have to be consistent with player strategies in equilibrium.

Consider the case where only b type sellers sell. In this case, the Buyer updates his beliefs on the type of Seller that sells a car on the used car market. Given these updated beliefs, the Buyer will offer a price of b. Bad car sellers will sell, but good car sellers won't. This is an equilibrium. A Seller with car of type b is willing to sell at price b. A Seller of car type g, is not willing to sell at price b. The Buyer believes that all cars being sold are type b. The buyer offers b. We have a **separating equilibrium**.

What if both types of Sellers sell on the market. Both types sell, so the Buyer does not update their beliefs. In this case, the price offered by the buyer will be $\pi g + (1 - \pi)b$ where π is the probability that the Seller is has a good car. As this price is greater than b, sellers of cars of type b will also sell. The price offered is $\pi g + (1 - \pi)b < g$. Because this low price is below the value of the good car, the good-type seller will not want to sell. That is to say, this is not an equilibrium. There is no **pooling equilibrium**.

In the Bayes Nash equilibrium of the lemon's market, all cars sold in the used car market are lemons. Is this an accurate representation of the actual used car market? What mechanisms exist that encourage sellers of good cars to sell in the used car market?

13.3 Insurance

The most well-known example of **adverse selection** is in the context of insurance. Insurance is the idea that two agents can trade because they have different risk preferences. In actual fact, the insurance company is able to be less risk-averse because it is able to diversify across a large number of bets.

This section presents a model of insurance where buyers of insurance are risk-averse. It then considers what happens when buyers are of different unobserved types. Some buyers tend to be sicker. These buyers will value insurance more. There will be a tendency for the market to behave like Akerlof's lemons market.

The sicker types buying insurance and the healthy types not buying insurance. The question is what types of policies could we implement to encourage healthier people to buy insurance.

13.3.1 Model

Assume there are two possible states of the world, sick and healthy. In the healthy state, an individual earns a certain income y, but in the sick state, the individual earns y and pays h in costs. These may be medical expenses or loss of income, etc. The individual's utility in the two states are $u(y)$ and $u(y-h)$. Importantly, the individual is better off receiving the expected outcome rather than expected utility over the two outcomes. That is the following inequality holds.

$$\begin{aligned} u(\pi y + (1-\pi)(y-h)) &= u(y-(1-\pi)h) \\ &> \pi u(y) + (1-\pi)u(y-h) \end{aligned} \tag{13.1}$$

where $(1-\pi)$ is the probability that the sick state occurs for this individual.

If the inequality holds, we say that the individual is risk-averse. Mathematically, their utility function is concave.

Because of the higher utility of the expected value than the separate states, the individual is willing to buy a product that pays less in the healthy state but more in the sick state.

Consider a product that pays h in the sick state at a cost (premium) of p that is paid in every state.

$$\pi u(y-p) + (1-\pi)u(y-h+h-p) = u(y-p) \tag{13.2}$$

If the premium is low enough, then the individual will prefer to be insured. If $p = (1-\pi)h + \epsilon$, then the individual strictly prefers insurance from Equation (13.1) (if ϵ is small enough).

For the insurance company, they would like to offer the product if $p - (1-\pi)h \geq 0$. So there exists a positive ϵ where the insurance company finds it profitable to offer insurance and the individual is willing to purchase the insurance.

13.3.2 Game with Unknown Types

So in a world with perfect information (albeit with uncertainty), it is profitable for insurance companies to offer insurance. What if the riskiness of the person is not observed by the insurer? We have two types $\pi \in \{\pi_l, \pi_h\}$, where π_l has a much higher probability of being in the sick state. The individual gets to observe their type prior to buying insurance, while the insurance company does not get to observe the individual's type.

In this situation, while both types may like to buy insurance, it is difficult to offer insurance to people with a low probability of being sick. This is basically the same problem as Akerlof's lemons market.

- Players: Buyer of type (Low (π_l), High (π_h)), Insurer
- Strategies:
  - Insurer: Offer a contract that pays h in the sick state at a premium p (which paid in all states).
  - Nature: Choose Buyer type $\{\pi_l, \pi_h\}$.
  - Buyer: Buy (or not) insurance at the offered premium after observing their type.
- Payoffs:
  - Buyer: Buy: $u(y - p)$, Don't Buy: $\pi u(y) + (1 - \pi)u(y - h)$
  - Insurer: Buy: $p - (1 - \pi)h$, Don't Buy: 0
  - where $\pi \in \{\pi_l, \pi_h\}$
- Beliefs: $\Pr(\pi = \pi_l) = q$

13.3.3 Bayes Nash Equilibrium

Consider the case where both types are willing to buy insurance and let q be the proportion of "sick" individuals. In this case, the insurer will want the premium to be such that $p-(1-q\pi_l-(1-q)\pi_h)h \geq 0$. For the "sick" individual, they will purchase if the following inequality holds.

$$u(y - p) \geq \pi_l u(y) + (1 - \pi_l)u(y - h) \tag{13.3}$$

Similarly, for the healthy individual.

Is there a Bayes Nash equilibrium of this game? Let use simulation to see what happens.

13.3.4 Simulation using R

Consider the following simulated game. The buyers are risk averse, with a constant absolute risk-aversion utility function, with the parameter r determining the extent of their aversion.

$$u(C) = \frac{C^{1-r}}{1 - r} \tag{13.4}$$

This is a relatively simple utility function that is used a lot in the literature.

Healthy individuals only get sick 1% of the time, while sick individuals get sick 10% of the time. The economy has a population where 90% are the healthy type. If an individual gets sick, they lose half of their income.

One difference between the simulation and the set up of the game above is that premiums are assumed to be set at the actuarially fair rate. This is equivalent to saying that the Insurer makes zero profits which is equivalent to saying that there is perfect competition in the insurance market.

```
> u_f = function(C, r) {(C^(1 - r))/(1 - r)}
> Eu_f = function(r, x, p) sum(p*u_f(x, r))
> prem_f = function(pi, h) (1 - pi)*h
```

The function `u_f()` is the utility function with constant absolute risk aversion. The function `Eu_f()` is the Agent's expected utility and `prem_f()` calculates the actuarially fair premium.

The parameters for the simulation are as follows. There is no particular rhyme or reason for the choice. The sick type has a 10 percent probability of being in the sick state, while the health type has a 0.1 percent probability of being in the sick state.

```
> r = 0.7
> pi_l = 0.9
> pi_h = 0.999
> q = 0.9
> y = 1
> h = 0.5
```

Utility for the uninsured for each type of individual.

```
> Eu_f(r, c(y, y-h), c(pi_l, 1 - pi_l))
  [1] 3.270751
> Eu_f(r, c(y, y-h), c(pi_h, 1 - pi_h))
  [1] 3.332708
```

Assume we have a **pooling equilibrium**. That is, both types purchase insurance.

```
> prem = prem_f(q*pi_l + (1 - q)*pi_h, h)
```

Given this premium of half of one percent of income, would both types purchase insurance?

```
> Eu_f(r, c(y - prem, y - prem),
+      c(pi_l, 1 - pi_l)) >
+   Eu_f(r, c(y, y-h), c(pi_l, 1 - pi_l))
  [1] TRUE
> Eu_f(r, c(y - prem, y - prem),
+      c(pi_h, 1 - pi_h)) >
+   Eu_f(r, c(y, y-h), c(pi_h, 1 - pi_h))
  [1] FALSE
```

No. The healthy type is not willing to purchase insurance at that premium.

Is there an equilibrium where only the sick types are insured? Is there a **separating equilibrium**?

```
> prem = prem_f(pi_l, h)
> Eu_f(r, c(y - prem, y - prem),
+      c(pi_l, 1 - pi_l)) >
+   Eu_f(r, c(y, y-h), c(pi_l, 1 - pi_l))
  [1] TRUE
> prem
  [1] 0.05
```

Yes. The sick individual is willing to pay a premium of 5% of her income. But 90% of the population is uninsured.

13.3.5 Mandatory Insurance

In the simulation above, only 10% of the population purchases insurance and the premiums are very high.

One solution to this problem is to make insurance mandatory. Or more accurately have some sort of fine or tax for those that don't take up insurance. That is, the payoff to the individual is decreased by the amount of the tax when the individual chooses not to purchase health insurance.

```
> tax = 0.05
> prem = prem_f(q*pi_l + (1 - q)*pi_h, h)
> Eu_f(r, c(y - prem, y - prem),
+      c(pi_l, 1 - pi_l)) >
+   Eu_f(r, c(y - tax, y-h - tax),
+        c(pi_l, 1 - pi_l))
  [1] TRUE
> Eu_f(r, c(y - prem, y - prem),
+      c(pi_h, 1 - pi_h)) >
+   Eu_f(r, c(y - tax, y-h - tax),
+        c(pi_h, 1 - pi_h))
  [1] TRUE
> prem
  [1] 0.04505
```

By adding a tax of 5%, we make the non-insurance expected utility lower, so the high type is willing to pay a higher premium to be insured. This allows risks to be pooled and makes it profitable for the insurance company to offer insurance that the healthy individuals are willing to accept.

Under this policy, the sick types do very well. Their premiums drop from 5% of income to 4.5% of income.

13.4 Empirical Analysis: Health Insurance using R

One of the concerns policy makers have with the health insurance market is that many people do not carry insurance. People who are generally healthy don't carry insurance. This tendency means that Akerlof's lemons problem leads to a market failure. People who tend to be sicker will have insurance, while healthier people will not have insurance.

This section looks at the actual health insurance market in the United States using the Medical Expenditure Panel Survey (MEPS). This data set provides information on how much health costs people actually have, how much income they have, and whether or not the people actually buy insurance.

13.4.1 Willingness to Pay

For each subgroup, we can calculate the expected utility in the case where they are uninsured, the case where they are insured in separating equilibrium (the baseline case), and for the case when they are insured under a pooling equilibrium (counterfactual case). To be clear, the analysis assumes that the current observed data are from a separating equilibrium. In each subgroup, there are healthy types that are choosing not to get insurance. The insurance company is assumed to offer a premium that is actuarially fair given the types that purchases insurance.

The utility is a constant absolute risk aversion (CARA) function with risk-parameter `r`. To derive expected utility, we take the average income for the subgroup, the probability of having medical expenditure, the average expenditure, and the standard deviation of expenditure. In the code, the expected utility is calculated numerically. The code uses a trick of creating a global variable and using transformations of the uniform and the standard normal distributions.

```
> set.seed(123456789)
> K = 1000
> U = runif(K) #uniform distribution
> Z = qnorm(U) #transform to standard normal
> EU_exp = function(r, income, exp_pos, exp_mean, exp_sd) {
+   exp = income - (U < exp_pos)*(Z*exp_sd + exp_mean)
+   return(Eu_f(r, ifelse(exp > 0, exp, 0), rep(1/K,K)))
+ }
```

13.4.2 Premiums

There are three premiums we can calculate in the data. First, there are the actual premiums paid by beneficiaries in the data. These are called "out of pocket" premiums. Many Americans have their premiums subsidized. For many working Americans, the premiums are paid in part or full by the firm that they work for. These workers are accepting at least some amount of lower salary in order to get the subsidy on the insurance premium from their employer. There is also a substantial tax benefit to workers who get health insurance, which is paid for by the American tax payer. Second, we can calculate the actuarial fair premium for each subgroup that has insurance. Lastly, we can calculate what the actuarial fair premium would be if the uninsured became insured for each subgroup.

The actuarial fair premium is calculated as the expected expenditure for the subgroup conditioning on having insurance (`UNISURD == 2`). The out of pocket premium is read from the data.

```
> file = paste0(dir, "meps_full.csv")
> dt1 = fread(file)
> dt_ins = dt1[UNINSURD == 2,.(premium = premium,
+                              premium_alt = exp_pos*exp_mean),
+             by = c("age_group", "SEX", "edu_group")]
```

MEPS data are used to calculate average out of pocket premiums and health expenditures for each subgroup.

The pooled premium is calculated as the average expected expenditure for each subgroup where both insured and unsured individuals are included in the average.

```
> dt_pool = dt1[, .(premium_pool = mean(exp_pos*exp_mean)),
+                  by = c("age_group", "SEX", "edu_group")]
```

The two data sets are merged back into the original data.

```
> dt2 = merge(dt1, dt_ins,
+              by = c("age_group", "SEX", "edu_group"))
> dt2 = merge(dt2, dt_pool,
+              by = c("age_group", "SEX", "edu_group"))
> dt2$premium = dt2$premium.y
```

The code below calculates the premiums by age group and plots the results.

```
> ins = dt2[,.(premium = mean(premium),
+              premium_alt = mean(premium_alt),
+              premium_pool = mean(premium_pool)),
+                   by = age_group]
> ins = ins[order(age_group)]
> line_prems = setDF(ins) |>
+   ggplot(aes(x = age_group)) +
+   geom_line(aes(y = premium_alt), linetype = 2) +
+   geom_line(aes(y = premium)) +
```

```
+   geom_line(aes(y = premium_pool), linetype = 3) +
+   labs(x = "Age",
+        y = "",
+        title = "Premium ($)") +
+   scale_y_continuous(limits = c(0,5000)) +
+   annotate("text", x = 50, y = 4000, label = "Seperating") +
+   annotate("text", x = 62, y = 3000, label = "Pooling") +
+   annotate("text", x = 60, y = 500, label = "Actual")
```

```
> line_prems
```

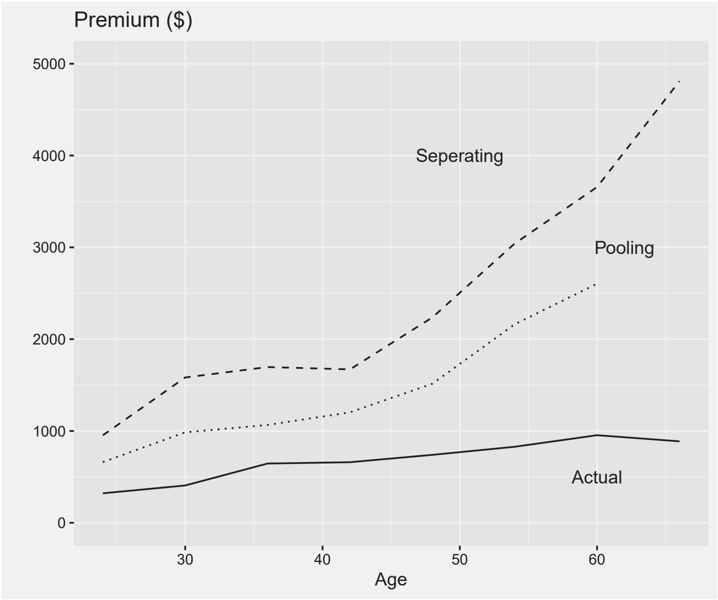


FIGURE 13.1
Line graph of average premiums for different age groups in the MEPS data. The lines show that premiums are increasing with age. The actual out of pocket premiums is substantially lower than the actuarial fair premiums for the insured population. The actual premiums increase from about \$200 to \$1,000 from 25 to 65. The actuarial fair premiums increase from \$1,000 to \$4,500 over the age-range. Pooled premiums increase from about \$750 to \$3,000.

Figure 13.1 presents line charts for average premiums across the different age groups. It presents the actual out of pocket premiums paid by insured beneficiaries, as well as the actuarially fair premium for the observed case and for the case where everyone is insured. It shows that the actual premium paid is substantially lower than the actuarially fair premium, particularly for the older beneficiaries. It also shows that premiums would come down in a pooling equilibrium.

13.4.3 Tax Policy

We may have too many people who are choosing to not have health insurance. The uninsured can be a burden on society because when they get sick they may not be able to cover the medical expenses, which moves those expenses to the hospitals or the government. There are a couple of policies that we could use to encourage a higher uptake of health insurance. We could either make choosing to have health insurance cheaper or we can make choosing not to have health insurance more expensive. The US uses the former. Taxes are not paid on the health insurance portion of income, giving a tax subsidy to employed people for purchasing insurance. An alternative policy of taxing those that don't choose to get insurance was originally planned for the insurance exchanges created as part of the Affordable Care Act.

The tax policy charges a tax on people who choose not to buy health insurance. If the tax is high enough, then most people will choose to have health insurance and the insurance premiums will fall.

To determine how behavior changes, calculate the expected value for being insured and being uninsured for the people in the data who are currently insured and for people who are currently uninsured. Below these are denoted `dt2_in` and `dt2_un`, respectively.

```
> r = 0.7
> dt2_in = dt2[UNINSURD == 2,
+             .(EU_uns = EU_exp(r,
+                               income + premium_alt - premium,
+                               exp_pos,
+                               exp_mean,
+                               exp_sd),
+               EU_ins = EU_exp(r,
+                               income - premium,
+                               0,
+                               0,
+                               0)),
+             by = "id"]
> dt2_un = dt2[UNINSURD == 1, .(EU_uns = EU_exp(r,
+                                               income,
+                                               exp_pos,
+                                               exp_mean,
+                                               exp_sd),
+               EU_ins_sep = EU_exp(r,
+                                   income - premium_alt,
```

```
+                                   0,
+                                   0,
+                                   0),
+              EU_ins_pool = EU_exp(r,
+                                    income - premium_pool,
+                                    0,
+                                    0,
+                                    0)),
+          by="id"]
```

The function tax_policy() calculates the percentage of the population that is insured given a tax rate. The code is a bit ugly. It calculates the expected utility if the uninsured individual remains uninsured with the new tax. It then mergers, the results back into the original data. The function reports the percent of the population that choose to be insured under the policy. It sums the proportion of the population already insured and then for the population that is uninsured it determines whether the expected utility under a pooling equilibrium is greater then the tax.

```
> r = 0.7
> tax_policy = function(tax) {
+   dt21 = dt2[UNINSURD == 1,
+              .(EU_tax = EU_exp(r,
+                                (1 - tax)*income,
+                                exp_pos,
+                                exp_mean,
+                                exp_sd)),
+             by = "id"]
+   dt2_un1 = merge(dt2_un, dt21, by = "id")
+   dt2_un2 = merge(dt2_un1, dt2, by = "id")
+   return((sum(dt2$count[dt2$UNINSURD==2], na.rm = TRUE) +
+             sum((dt2_un2$EU_tax < dt2_un2$EU_ins_pool)*
+                   dt2_un2$count, na.rm = TRUE))/
+            sum(dt2$count, na.rm = TRUE))
+ }
```

The code below calculates the percentage of the population that is insured under different tax rates and plots the results.

```
> line_ins_prop = data.frame(tax = seq(0, 1, 0.01),
+            ins_prop = sapply(1:101, function(i)
+              tax_policy(i/100))) |>
+   ggplot(aes(x = tax, y = ins_prop)) +
+   geom_line() +
+   labs(x = "Tax rate",
+        y = "",
```

```
+         title = "Percent insured") +
+    scale_x_continuous(limits = c(0, 0.4)) +
+    scale_y_continuous(limits = c(0.75, 1)) +
+    geom_hline(yintercept = 1, linetype = 2)

> line_ins_prop
```

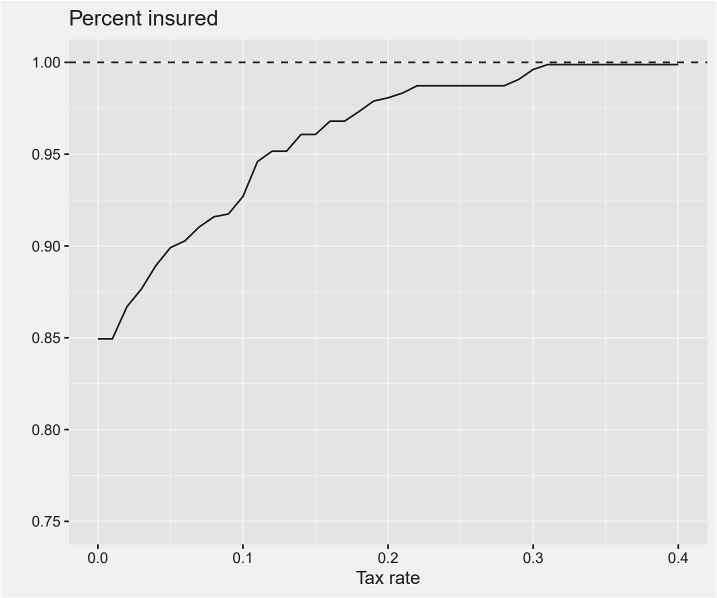


FIGURE 13.2
Line graph of the percentage of the population insured against the tax rate. Around 85% are insured without any tax, the proportion reaches 100% with a tax rate over 30%.

Figure 13.2 suggests that it is not that easy to encourage people to purchase insurance by imposing a tax penalty on the uninsured. The relationship is non-linear meaning that a small tax can get a large proportion of people onto insurance, but the relative effectiveness falls the higher the tax. The analysis suggests taxes 30 percent rage are needed to get everyone insured.

13.5 Discussion and Further Reading

Adverse selection is most famously illustrated by Akerlof (1970) and the idea of a lemons market. Adverse selection and moral hazard have become key to understanding health insurance markets. These ideas formed the basis for policies introduced in the Affordable Care Act that aimed to reduce the number of uninsured. See Pauly (1974). The tax on the uninsured, called the "individual mandate" was subject to lawsuits and was eventually repealed. So did it work? A recent survey by Brookings economist, Matt Fiedler finds that the tax *may* have reduced the uninsured rate. It turns out to be very tricky to determine what happened given all the changes that occurred with the introduction of the ACA and the complexity of the policy (Fiedler, 2020).

Bibliography

Christopher P. Adams, *Learning Microeconometrics with R*, The R Series, CRC Press, December 2020.

Christopher P. Adams and Emek Basker. The Battle of the Big-Box Bookstores. January 2025.

Victor Aguirregabiria. Empirical industrial organization: Models, methods, and applications. University of Toronto, June 2021.

George A. Akerlof. The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84(3):488–500, August 1970.

Steffen Andersen, Seda Ertaç, Uri Gneezy, Moshe Hoffman, and John A. List. Stakes matter in ultimatum games. *American Economic Review*, 101(7): 3427–3439, December 2011. https://doi.org/10.1257/aer.101.7.3427.

Gaurab Aryal, Serafin Grundl, Dong-Hyuk Kim, and Yu Zhu. Empirical relevance of ambiguity in first-price auctions. *Journal of Econometrics*, 204(2):189–206, 2018.

Orley Ashenfelter, David Ashmore, Jonathan Baker, Suzanne Gleason, and Dan Hosken. Empirical Methods in Merger Analysis: Econometric analysis of pricing in FTC vs Staples. *International Journal of the Economics of Business*, 13(2):265–279, 2006.

John Asker. A study of the internal organization of a bidding cartel. *American Economic Review*, 100(3):724–762, 2010.

Susan Athey and Philip A. Haile. Identification in standard auction models. *Econometrica*, 70(6):2170–2140, 2002.

Susan Athey, Jonathan Levin, and Enrique Seira. Comparing open and sealed bid auctions: Evidence from timber auctions. *The Quarterly Journal of Economics*, 126:207–257, 2011.

Matthew Backus, Christopher Conlon, and Michael Sinkinson. Common ownership and competition in the ready-to-eat cereal industry. NBER Working Paper No. 28350, Janurary 2021.

David J. Balan and Keith Brand. Simulating Hospital Merger Simulations. *The Journal of Industrial Organization*, 71(1):47–123, 2018.

Laura H. Baldwin, Robert C. Marshall, and Jean-Francois Richard. Bidder collusion at forest service timber sales. *Journal of Political Economy*, 105 (5):657–699, 1997.

Steven Berry, James Levinsohn, and Ariel Pakes. Automobile prices in market equilibrium. *Econometrica*, 60(4):889–917, 1995.

Timothy F. Bresnahan and Peter C. Reiss. Entry in monopoly markets. *Review of Economic Studies*, 57(4):57–81, 1990.

Timothy F. Bresnahan and Peter C. Reiss. Empirical models of discrete games. *Journal of Econometrics*, 48:57–81, 1991a.

Timothy F. Bresnahan and Peter C. Reiss. Entry and competition in concentrated market. *The Journal of Political Economy*, 99(5):977–1009, 1991b.

David P. Byrne and Nicolas de Roos. Learning to Coordinate: A Study in Retail Gasoline. *The American Economic Review*, 109(2):591–619, 2019.

Colin F. Camerer. *Behavioral Game Theory: Experiments in Strategic Interaction.* Princeton University Press, 2003.

Lisa A. Cameron. Raising the Stakes in the Ultimatum Game: Evidence from Indonesia. *Economic Inquiry*, 37(1):47–59, 2007.

Cory Capps, David Dranove, and Mark Satterthwaite. Competition and Market Power in Option Demand Markets. *The RAND Journal of Economics*, 34 (4):737–763, Winter 2003.

George Deltas. Asymptotic and small sample analysis of the stochastic properties and certainty equivalents of winning bids in independent private values auctions. *Economic Theory*, 23:715–738, 2004.

Matthew Fiedler. The ACA's Individual Mandate In Retrospect: What Did It Do, And Where Do We Go From Here? *Health Affairs*, 39(3): 429–435, March 2020.

Drew Fudenberg and Jean Tirole. *Game Theory.* MIT Press, 1991.

Christopher Garmon. The accuracy of hospital merger screening methods. *The RAND Journal of Economics*, 48(4):1068–1102, 2017.

Martin Gaynor and William B. Vogt. Competition among hospitals. *The RAND Journal of Economics*, 34(4):764–785, 2003.

Robert Gibbons. *Game Theory for Applied Economists.* Princeton University Press, 1992.

Gautam Gowrisankaran and Robert J. Town. Competition, Payers, and Hospital Quality. *Health Services Research*, 38(6):1403–1422, 2003.

Emmanuel Guerre, Isabelle Perrigne, and Quang Vuong. Optimal nonparametric estimation of first-price auctions. *Econometrica*, 68(3):525–574, 2000.

Philip A. Haile and Elie T. Tamer. Inference with incomplete model of English auctions. *Journal of Political Economy*, 111(1), 2003.

Philip A. Haile, Han Hong, and Matt Shum. Nonparametric tests for common values in first-price sealed-bid auctions. Yale, November 2006.

Lars Peter Hansen. Large sample properties of generalized method of moments estimators. *Econometrica*, 50(4):1029–1054, 1982.

Kenneth Hendricks, Joris Pinkse, and Robert H. Porter. Empirical implications of equilibrium bidding in first-price, symmetric, common value auctions. *Review of Economic Studies*, 70:115–145, 2003.

Eric Hilt. Incentives in corporations: Evidence from the american whaling industry. *Journal of Law & Economics*, 49(1):197–227, 2006.

Bengt Holmström. Moral Hazard and Observability. *The Bell Journal of Economics*, 10(1): 74–91, Spring 1979.

Ali Hortaçsu and Joonhwi Joo. *Structural Econometric Modeling in Industrial Organization and Quantitative Marketing: Theory and Applications*. Princeton University Press, 2023.

Harold Hotelling. Stability in competition. *The Economic Journal*, 39(153): 41–57, March 1929. https://www.jstor.org/stable/2224214.

David E. Kalist. Data from the television game show friend or foe? *Journal of Statistics Education*, 12(3), 2004.

Kate Ho and Ariel Pakes. Hospital Choices, Hospital Prices, and Financial Incentives to Physicians. *The American Economic Review*, 104(12), December 2014.

Jean-Jacques Laffont and Quang Vuong. Structural analysis of auction data. *The American Economic Review*, 86(2):414–420, May 1996.

Edward P. Lazear and Paul Oyer. *Personnel Economics*, chapter 12, pages 479–519. The Handbook of Organizational Economics. Princeton University Press, Princeton, 2013. Robert Gibbons and John Roberts (eds).

Matthew S. Lewis. Price leadership and coordination in retail gasoline markets with price cycles. *International Journal of Industrial Organization*, 30(4): 342–351, July 2012.

John A. List. Friend or foe? a natural experiment of the prisoner's dilemma. *Review of Economics and Statistics*, 88(3):463–471, August 2006.

Mark D. Manuszak and Charles C. Moul. Prices and endogenous market structure in office supply superstores. *The Journal of Industrial Economics*, 56(1):94–112, March 2008.

Eric Maskin and Jean Tirole. A theory of dynamic oligopoly, ii: Price competition, kinked demand curves and edgeworth cycles. *Econometrica*, 56(3): 571–599, May 1988.

Nathan H. Miller and Matthew C. Weinberg. Understanding the price effects of the millercoors joint venture. *Econometrica*, 85(6):1763–1791, November 2017.

Nathan H. Miller, Gloria Sheu, and Matthew C. Weinberg. Oligopolistic price leadership and mergers: The united states beer industry. *The American Economic Review*, 111(10):3123–3159, October 2021.

Aviv Nevo. Measuring market power in the ready-to-eat cereal industry. Econometrica, 69(2):307–342, 2001.

Harry J. Paarsch and Han Hong (with M. Ryan Haley). *An Introduction to the Structural Econometrics of Auction Data*. MIT Press, 2006.

Mark V. Pauly. Overinsurance and public provision of insurance: The roles of moral hazard and adverse selection. *Quarterly Journal of Economics*, 88(1): 44–62, February 1974.

Martin Puterman. Markov Decision Processes: Discrete Stochastic Dynamic Programming. *John Wiley and Sons*, 1994.

Devesh Raval, Ted Rosenbaum, and Steve Tenn. Semiparametric discrete choice model: An application to hospital mergers. *Economic Inquiry*, 55(4): 1919–1944, 2017.

Devesh Raval, Ted Rosenbaum, and Nathan E. Wilson. Using disaster induced closures to evaluate discrete choice models of hospital demand. *The RAND Journal of Economics*, 53(3):561–589, 2022.

Elie T. Tamer. Incomplete simultaneous discrete response model with multiple equilibria. *The Review of Economic Studies*, 70(1):147–165, January 2003.

Katja Seim. An empirical model of firm entry with endogenous product-type choices. *The Rand Journal of Economics*; 37(3): 619-640, Autumn 2006.

Raphael Thomadsen. The effect of ownership structure on prices in geographically differentiated industries. *The RAND Journal of Economics*, 36(4): 908–929, 2005.

Jean Tirole. *The Theory of Industrial Organization*. MIT Press, 1988.

Hal Varian. Position auction. *International Journal of Industrial Organization*, 25(6):1163–1178, December 2007.

Zhongmin Wang. (Mixed) Strategy in Oligopoly Pricing: Evidence from Gasoline Price Cycles Before and Under a Timing Regulation. *Journal of Political Economy*, 117(6):987–1030, December 2009.

Subject Index

A

adverse selection xviii, 233, 235, 246
affiliated values auctions 197; *see also* common, values auctions
Akerlof's lemons market 234–236, 240, 246
alternating offers bargaining game, *see* Rubinstein alternating offers bargaining game

B

Bayes/Bayesian
  game 164–165, 167
  Nash equilibrium 163–164, 168–169, 173, 178, 183–184, 198–199, 203, 221, 223, 224, 235, 237
  rule 164, 99
beliefs 148, 163–171, 178, 184, 198, 221–224, 234–235, 237
Bertrand
  game xvii, 39–40, 45, 47, 62, 118, 120, 122, 124–125, 130–132
  Nash equilibrium 120, 122, 125–126, 128–129
best response function 40–44, 48–49, 51, 115, 119, 147
bid ring 192, 196–197, 207, 209, 211, 214
biprobit/bivariate probit 75, 77–78
bootstrap 76–77, 93, 95

C

common
  knowledge 3
  values auction 197, 200, 203–204, 206, 209, 214; *see also* affiliated values auctions
coordination games xvii, 79
Cournot game xvii, 27, 39–40, 42–45, 62, 120
cutoff strategy 143, 164

D

dominant strategy xvii, 11, 13, 18–22
  equilibrium xvii, 19–20, 22
Dutch auction 176–177, 196

E

endogeneity 25
English auction xviii, 176–177, 183–196; *see also* Vickrey auction
entry game 24–27, 29–30, 38, 63, 67–68, 71, 76, 79, 100, 103–104, 163, 169, 171–172, 175
equilibrium 19–20, 22, 24, 26, 28; *see also* Nash, equilibrium
  refinement 24, 103, 107
extensive form representation xv, 103–104, 106–107, 135, 143–145; *see also* game tree; normal form representation

F

field experiment 17
finite difference estimator 181
first

price auction 177, 199–200, 203, 205, 209; *see also* sealed bid auctions
mover advantage 108
fixed effects, *see* two-way fixed effects

G
game tree 106–108, 135, 144, 146; *see also* extensive form representation
generalized method of moments (GMM) 108
Grim-trigger strategy 117

H
Hotelling game/line xvii, 39, 47–48, 61–62

L
lemons market, *see* Akerlof's lemons market
linear regression xv, 15, 17, 33, 51, 156, 157, 188, 206
logit xv, 15–17, 118–119, 122, 126, 140, 153–154; *see also* probit; softmax

M
marginal cost 27, 39, 40, 42, 44–45, 47–48, 52–54, 121–122, 125–126
maximum likelihood estimation xv, 68, 74, 171, 204
method of moments 53, 92; *see also* generalized method of moments
mixed strategy xv, 79–80, 82–86, 92, 95–98, 100, 132; *see also* pure strategy
moral hazard xviii, 133, 150–152

N
Nash
bargaining model/solution xviii, 133, 150–152
equilibrium xv, xvii, 18–19, 22–24, 26, 28–29, 38–41, 43, 47–49, 51, 58–59, 79–85, 92, 95–100, 105–108, 110–111, 114–117, 119–126, 128–129, 134, 136–138, 140, 142–143, 145, 147–148, 150, 163, 168, 196
non-parametric estimate 180, 186–187, 196
normal form representation xv, 10, 12, 20–21, 26, 81, 83, 85, 89, 96, 103, 115, 135–136; *see also* extensive form representation

O
oligopoly xvii–xviii, 39, 61, 113, 124–125
Order statistic 177, 190, 193, 196, 198–200
Ordinary least squares, *see* linear regression

P
Pareto optimal 11, 114
pipe operator 13
pooling equilibrium 233, 235, 240, 243–244; *see also* separating equilibrium
principal-agent problem xviii, 219–220, 223, 232
prisoner's dilemma game xvii, 3, 10, 12–13, 15, 20–23, 114–117, 124–125
probit 15–17, 65, 67, 74–75, 77–78, 171–173; *see also* logit
pure strategy 82–84; *see also* mixed strategy

R
reduced form estimation 171–172, 175; *see also* structural modeling
Rubinstein alternating offers bargaining game xviii, 133, 143–151

S
screening game 233; *see also* signaling game
sealed bid auctions xviii, 177–178, 183–184, 188–192, 195–196, 200; *see also* first, price auctions
second-price sealed bid auction, *see* Vickrey auction
signaling game 233; *see also* screening game
softmax 71; *see also* logit
strategically equivalent 177
structural modeling xiv, 31–33, 35, 38, 59, 100, 113, 171, 172; *see also* reduced form estimation
subgame 108–109, 138, 147, 221
    perfection xv, xviii, 103–104, 107–108, 110, 138, 140, 142–144, 147, 159
super markups (smu) 125–128

T
two-step estimation xv, 171, 173, 196
two-way fixed effects xv, 228
type 164, 167–170, 178, 198, 233–240

U
ultimatum game xviii, 133–138, 143–145, 147, 159

V
Vickrey auction 177; *see also* English auction

W
weakly dominant strategy 13, 19; *see also* dominant strategy
    equilibrium 19
Winner's curse 198–199, 209

Z
zero-sum game xvii, 80–82, 89, 96

Author Index

A

Adams, Christopher P. 73, 79, 99, 112, 175
Akerlof, George A. 233–236, 240, 246
Andersen, Steffen 138
Aryal, Gaurab 183, 196
Ashenfelter, Orley 25, 38
Ashmore, David 25, 38
Asker, John 192
Athey, Susan 177, 183, 185, 186, 196

B

Backus, Matthew 131
Baker, Jonathan 25, 38
Balan, David J. 160
Baldwin, Laura H. 188, 196
Basker, Emek 73, 79, 112, 175
Berry, Steven 62
Bertrand, François 42, 45, 47
Brand, Keith 160
Bresnahan, Timothy F. 24, 30, 32, 38, 79, 175
Byrne, David P. 113, 131

C

Cameron, Lisa A. 159
Capps, Cory 154, 159
Conlon, Christopher 131
Cournot, Augustine 40

D

Deltas, George 193
de Roos, Nicolas 113, 131
Dranove, David 154, 159

E

Ertaç, Seda 138

F

Fiedler, Matthew 246

G

Garmon, Christopher 160
Gaynor, Martin 159
Gibbons, Robert xvii, xxi
Gleason, Suzanne 25, 38
Gneezy, Uri 138
Gowrisankaran, Gautam 159
Grundl, Serafin 183, 196
Guerre, Emmanuel 171, 175, 177, 196

H

Haile, Philip A. 177, 183–186, 188, 196, 214
Hansen, Lars Peter 100
Harsayni, John 163
Hendricks, Kenneth 215
Hilt, Eric 226, 232
Hoffman, Moshe 138
Holmström, Bengt 232
Hong, Han xxi, 188, 214–215
Hosken, Daniel 25, 38
Hotelling, Harold 47, 61

K

Kalist, David E. 13, 17
Kim, Dong-Hyuk 183, 196

L

Laffont, Jean-Jacques 214
Lazear, Edward P. 232
Levin, Jonathan 196
Levinsohn, James 62
Lewis, Matthew S. 131
List, John A. 17, 138

M
Marshall, Robert C. 188, 196
McFadden, Daniel 63
Miller, Nathan H. 113, 131
Morgenstern, Oskar 18

N
Nash, John C. 58
Nash, John F. 18, 22, 58, 61, 80, 82, 150
Nevo, Aviv 131

O
Oyer, Paul 232

P
Paarsch, Harry J. xxi, 215
Pakes, Ariel 62
Pauly, Mark V. 246
Perrigne, Isabelle 171, 175, 177, 196
Pinske, Joris 215
Porter, Robert H. 215

R
Raval, Devesh 160
Richard, Jean-Francois 188, 196
Reiss, Peter C. 24, 30, 32, 38, 79, 175
Rosenbaum, Ted 160
Rubinstein, Ariel 145

S
Satterthwaite, Mark 154, 159
Seim, Katja 164, 175
Seira, Enrique 196
Sheu, Gloria 113, 131
Shum, Matt 188, 214
Sinkinson, Michael 131

T
Tamer, Elie T. 79, 175, 184
Tenn, Steve 160
Thomadsen, Raphael xx, 49, 51
Town, Robert J. 154, 159

V
Vickrey, William 177, 179, 183, 184, 196
Varian, Hal 176
Vogt, William B. 159
von Neumann, John 18
Vuong, Quang 171, 175, 177, 196, 214

W
Weinberg, Matthew C. 113, 131
Wilson, Nathan E. 160

Z
Zhu, Yu 183, 196