EMPIRICAL ANALYSES OF REGULATION AND NEGOTIATED PRICES

João Vitor Granja de Almeida

A DISSERTATION

in

Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2021

Supervisor of Dissertation

Aviv Nevo
George A. Weiss and Lydia Bravo Weiss University Professor

Graduate Group Chairperson

Jesús Fernández-Villaverde
Professor of Economics

Dissertation Committee

Katja Seim, Professor of Economics, Department of Economics and School of Management, Yale University

José Miguel Abito, Assistant Professor of Business Economics & Public Policy

EMPIRICAL ANALYSES OF REGULATION AND NEGOTIATED PRICES

To all those who made it possible.

ACKNOWLEDGMENTS

Writing these acknowledgements marks the end of a long journey. I owe thanks to many people who in one way or another made it possible for me to write these words.

First, I would like to thank those who taught me. I am deeply grateful to my committee members, Aviv Nevo, Katja Seim, and Mike Abito, for their guidance and unabating support, especially during the last few months of my PhD. They were all incredibly generous with their time, patient with my forays into countless dead-ends, and tried to guide my inquiries towards the most fruitful directions. I will do my best to carry those lessons with me and put them to good use.

I thank Leandro Gorno, Humberto Moreira, and Giovanni Maggi for supporting my application to PhD programs. Rudi Rocha helped me through my first experience with research, and convinced me early on that discovery was within reach. I must thank Rolando Gárciga Otero. Without him, this journey might never have begun; at the very least, it would have been much harder. Many at the Federal University of Rio de Janeiro were dedicated teachers, and instilled in me a sense of curiosity that led me down this path.

I feel extremely lucky to have formed close and, I hope, long-lasting friendships during my PhD. Tomás Larroucau, Görkem Bostanci, Sergio Villalvazo, and Gabrielle Vasey were always there to talk about Economics and forget all about it when the time was right. They made the past six years great fun. It would have been much

harder without them, and I am enourmously grateful for their friendship. I also thank Jake Krimmel, Stuart Craig, Ewelina Zurowska, Camille Köenig, and Montserrat Ganderats for their friendship and the many good memories I will carry with me. André Victor Luduvice was always keen on helping and showing me the way, and I thank him for that. I thank the Brazilian group, for providing me with a sense of familiarity and comfort when that was needed. I also thank my friends in Brazil for the many years of friendship, and for being a safe harbor when what I needed was to forget.

Federica De Stefano is the best partner anyone could have. Her unwavering support during the last few years helped tremendously in getting me through the finish line. She somehow always knew what was needed. She gave me a push when one was in order, and was my source of tranquility and enjoyment during the toughest times. I am forever grateful for her being by my side.

Last but not least, I thank my parents and sisters for providing the foundation and stability on top of which this journey was built.

ABSTRACT

EMPIRICAL ANALYSES OF REGULATION AND NEGOTIATED PRICES

João Vitor Granja de Almeida

Aviv Nevo

In the first chapter of this dissertation, I study coverage requirements, a common regulation in the mobile telecommunications industry that intends to accelerate the roll-out of new mobile telecommunications technologies to disadvantaged areas. I argue that the regulation may engender entry deterrence effects that limit its efficacy and lead to technology introduction patterns that are not cost-efficient. To quantify the impact of coverage requirements on market structure and the speed and cost of technology roll-out, I develop and estimate a dynamic game of entry and technology upgrade under regulation. I estimate the model using panel data on mobile technology availability at the municipality level in Brazil. In counterfactual simulations, I find that coverage requirements accelerate the introduction of 3G technology by just over one year, on average, and reduce firms' profits by 24% relative to a scenario with no regulation. I find the entry deterrence effects to be small. Moreover, an alternative subsidization policy leads to a similar acceleration in the roll-out of 3G and substantially higher aggregate profits, likely increasing aggregate welfare relative to coverage requirements. In the second chapter, I investigate how the portfolio of products carried by retailers influences wholesale and retail prices. To this end, I develop and estimate a model of retailer pricing and retailer-manufacturer negotiations over wholesale prices. The estimation approach extends existing econometric tools for multi-product bargaining models to a setting with optimal downstream pricing. I use the estimated model to simulate the effects of counterfactual scenarios in which private label products or the products of a national manufacturer are excluded from retailers' product portfolios. I find that wholesale prices do increase, but those effects are small. Eliminating private label products leads to an average increase in wholesale prices of only 0.10%; retail prices increase by only 0.04%. Eliminating a national manufacturer's products leads to increases in wholesale prices between 0.003% and 0.677%; retail prices decrease by 0.027%-4.210% due to downstream pricing incentives.

TABLE OF CONTENTS

DEDICATION iii

ACKNOWLEDGMENTS iv

ABSTRACT vi

LIST OF TABLES xi

LIST OF ILLUSTRATIONS xiii

1 Regulation and Service Provision in Dynamic Oligopoly: Evidence from Mobile Telecommunications 1

1.1 Introduction 1

1.2 Institutional Setting and Data 9

1.3 Model 22

1.3.1 Symmetric Markov Perfect Equilibria 25

1.3.2 Quasi-Stationary Markov Perfect Equilibria 27

1.3.3 Modeling Flow Profits 28

1.4 Identification and Estimation 33

1.4.1 Estimation of the Flow Profit Function 33

1.4.2 Identification and Estimation of Dynamic Parameters 37

1.4.3 Estimation 39

1.5 Estimation Results 40

1.6 Counterfactual Analysis 42

1.6.1 The Effect of Coverage Requirements 43

1.6.2 Alternative Regulatory Interventions 52

1.7 Conclusion 59

1.A Regulation and Delay in the Fudenberg-Tirole Model 60

1.A.1 The Model 60

1.A.2 Incorporating Regulation 63

1.B Descriptive Models – Alternative Specifications 64

1.C Symmetry Restrictions 69

1.D Conditional Value Functions are Linear in Parameters 69

2 Retailers' Product Portfolios and Negotiated Wholesale Prices 74

2.1 Data 78

2.2 Model 81

2.2.1 Model: Retailer Profit Maximization and Negotiated Wholesale Prices 81

2.2.2 Model: Demand 85

2.2.3 Theoretical Results 86

2.3 Econometrics 89

2.3.1 Demand Estimation 90

2.3.2 Construction of Wholesale Prices 92

2.3.3 Estimation of Manufacturer Cost and Bargaining Parameters . 94

2.4 Counterfactual Exercises 100

2.5 Conclusion 104

2.A Proofs and Auxiliary Results 105

2.B Characterizing Retail Price Changes with Respect to Changes in Costs 112

BIBLIOGRAPHY **113**

LIST OF TABLES

Regulation and Service Provision in Dynamic Oligopoly: Evidence from Mobile Telecommunications **1**

1.1 Summary Statistics – Municipality Characteristics 16
1.2 Summary Statistics – Mobile Expenses and HH Characteristics 17
1.3 Summary Statistics 18
1.4 Entry/Upgrade Models 21
1.5 Static Parameter Estimates 41
1.6 Dynamic Parameter Estimates 42
1.7 Explaining Time to Adoption and the Effect of Regulation 46
1.8 Entry/Upgrade Models – Without group fixed effects 65
1.9 Entry/Upgrade Models – With Neighboring Network Info 67
1.10 Testing for Selection on Infrastructure in Neighboring Municipalities . 68

Retailers' Product Portfolios and Negotiated Wholesale Prices **74**

2.1 Do Stores Face Lower Marginal Costs for Store Brand Products? . . . 94
2.2 Bargaining parameter estimates 100
2.3 Wholesale Price Inflation: Products Negotiated in Both Scenarios . . 103
2.4 Retail Price Inflation: Negotiated Products 103
2.5 Changes in Consumer Surplus 104

LIST OF ILLUSTRATIONS

Regulation and Service Provision in Dynamic Oligopoly: Evidence from Mobile Telecommunications **1**

1.1 Regulated Carriers – Midwest 13
1.2 Technology availability in the state of Pará 14
1.3 Subscribers by technology over time 16
1.4 Distribution of the Probability of Reaching 3G by December 2019 Without Regulation. 44
1.5 Distribution of the time to introduction of 3G technology or better under alternative regulatory regimes. 45
1.6 How much faster is the introduction of 3G+ under regulation? 47
1.7 Equilibrium Effects 51
1.8 Acceleration of 3G Introduction Under Coverage Requirements and Subsidy 55
1.9 Targeting Properties of Coverage Requirements and a Flat Subsidy . 56
1.10 Targeting Properties of Coverage Requirements and a Municipality-Specific Subsidy 58
1.11 Payoffs in the Fudenberg-Tirole Model. 62
1.12 Payoffs in the Fudenberg-Tirole Model with Regulation. 63

Retailers' Product Portfolios and Negotiated Wholesale Prices **74**

2.1 Distribution of the number of stores across chains 80

2.2 Distribution of the number of suppliers across chains 80

2.3 Distribution of the share of private label products across chains . . . 81

2.4 Price paths for two products (columns) in two different stores (rows). 91

2.5 Distribution of price coefficient estimates under OLS and GMM. . . . 92

Chapter 1

Regulation and Service Provision in Dynamic Oligopoly: Evidence from Mobile Telecommunications

1.1 Introduction

In industries with large fixed costs, firms' failure to appropriate the consumer surplus they generate when they enter new markets and introduce new products may lead to underprovision of goods and services. This possibility is particularly relevant in disadvantaged areas, where the prospects of recouping fixed costs are dim. Concerns regarding service underprovision have led to regulatory oversight and intervention in many industries, such as postal service, healthcare, airlines, and telecommunications.[1] These concerns have historically been particularly salient in the telecommunications industry (Wu, 2010). The substantial investment costs required for network expansion

[1] USPS is subject to a Universal Service Obligation. The HRSA runs the Medicare Rural Hospital Flexibility Program. The DOT runs the Essential Air Service and Small Community Air Service Development Program. The Universal Service Administrative Company spends almost ten billion dollars annually in subsidies for high-speed broadband access.

raise fears that firms will not provide service and bring new mobile telecommunications technologies to low-income, rural, or isolated localities, despite the considerable benefits associated with these services.[2] These concerns have led to the regulation of the roll-out of new mobile telecommunications technologies in countries ranging from Nigeria to the United States. This paper studies the effects of existing regulation on the introduction of new mobile telecommunications technologies, and evaluates the desirability of existing regulation relative to alternative forms of intervention.

Mobile telecommunications markets are typically characterized by a small number of firms. To provide mobile telecommunications services, these firms must acquire from the government licenses to use the radio spectrum. These licenses tyically cover large geographic areas containing many local markets. In the absence of regulation, firms would choose to provide service and introduce new technologies in those markets where variable profits exceed fixed costs, potentially leaving some areas without service or access to new technologies. To avoid this outcome, regulators impose what are called coverage requirements. A coverage requirement tasks a single firm with providing service of a specific technology in a given area by a date set by the regulator.[3]

The goal of this paper is to understand the welfare effects of coverage requirements and alternative regulatory interventions. At first glance, the trade-off faced by regulators when deciding whether or not to impose a coverage requirement is clear. On the one hand, the requirement presumably accelerates the introduction of the new

[2]Telecommunications services have been shown to have positive effects on economic growth (Roller and Waverman, 2001; Czernich et al., 2011); labor productivity (Bertschek and Niebel, 2016; Akerman, Gaarder, and Mogstad, 2015); market efficiency (Jensen, 2007), and risk-sharing (Jack and Suri, 2014). (Aker and Mbiti, 2010) discuss many other potential benefits of mobile telecommunications in developing countries.

[3]Another common form of coverage requirements is that firms are obliged to provide service to at least some fraction of the territory covered by their license by a date set by the regulator. This fraction varies across countries and in some cases is close to 1.

technology in the regulated area, thus increasing the discounted stream of consumer surplus. On the other hand, coverage requirements impose a cost on the regulated firm, for it is required to enter a market or upgrade its technology when it might not have done so in the absence of regulation. The oligopolistic structure of the mobile telecommunications industry overturns this apparent simplicity. A coverage requirement is a credible commitment to provide service on the part of the firm subject to the regulation. This commitment may deter entry by the other firms and lead to further changes in equilibrium behavior that diminish or even reverse the acceleration of the introduction of the new technology alluded to above.

To quantify the effects of coverage requirements and alternative policies, I develop and estimate an empirical dynamic game of firm entry and technology upgrade under regulation. Firms' incentives to enter a market and upgrade their technologies are determined by the incremental variable profit derived from those choices and the associated sunk costs. Therefore, an appropriate empirical model must accurately capture the key features determining those profits and costs. An important characteristic of rapidly evolving industries such as mobile telecommunications is that demand for a new technology tends to increase over time whereas the associated adoption costs tend to decrease. Also important are local market features that shape demand and costs, as well as the local market structure. To account for these key factors, I model firms' flow profits as a time-varying function of market structure and local demographic characteristics. The model also allows the costs of introducing a new technology to vary over time and across local markets.

The other crucial determinant of firms' incentives to introduce the new technology is, of course, the regulation. In the model, as in the data, in each market exactly one firm is required to provide 3G service by a date set exogenously by the regulator. I model the regulation's enforcement by assuming that the regulated firm must pay a fine in every period after the regulation deadline in which it fails to comply with the regulation. There are two dimensions to the incentives stemming from the regulation, given its asymmetric nature. First, the single regulated firm has an added incentive to introduce the new technology, to avoid triggering punishments for non-compliance. Second, the firms that are not subject to the regulation know that the regulated firm will be in the market in the future, and with the new technology. Therefore, they know that the market will be more competitive in the future, and that knowledge negatively affects their incentives to enter and introduce the new technology. The latter mechanism may give rise to a further response by the regulated firm: knowing that the unregulated firms will not enter the market and knowing that adoption costs decrease over time, the regulated firm may have an incentive to wait for costs to fall before introducing the new technology. As this discussion makes clear, capturing these mechanisms requires an equilibrium model of entry and technology adoption.

The question of how much later (or earlier) the introduction of 3G technology would have ocurred in the absence of regulation is a question about time, and thus requires a dynamic model. The nature of the regulation, which sets a deadline for the introduction of the new technology, also makes the problem dynamic (and non-stationary). These aspects justify the dynamic nature of the model.

The time-varying nature of variable profits and technology adoption costs and the regulation deadline make the environment non-stationary, a departure from most of the literature on empirical dynamic games. I also depart from the existing empirical literature on technology adoption, which applies full-solution estimation routines based on backward induction solution algorithms. I instead assume that structural parameters stabilize before the end of the sample and focus on what I call quasi-stationary Markov Perfect Equilibria (QMPE). Essentially, QMPE have a non-stationary phase followed by a stationary phase. This structure allows me to adapt existing estimation methods used in stationary dynamic games to a non-stationary setting.

I estimate the model using new panel data on mobile technology availability at the municipality level in Brazil from June 2013 to June 2020. I analyze firms' entry and technology upgrade behavior in a set of mostly rural municipalities. In each of these municipalities, exactly one of the four major carriers in the country was required to provide 3G service by a date set by the regulator. I call that firm the regulated firm. The identity of the regulated firm varied across municipalities; all of the four major carriers in the country are regulated in some markets but not others. Comparing the behavior of regulated and unregulated firms shows that the latter are less likely to enter a market or upgrade their technology when the regulated firm is yet to satisfy its coverage requirement. This pattern is consistent with the entry deterrence effect outlined above.

The model estimates show that the profits and costs associated with 3G are stable over my sample period. The profits associated with 4G rise sharply, and the costs of 4G installation decrease substantially. The latter inference is driven by a sharp increase in 4G introductions in the final part of the sample. The cost of non-compliance with the regulation is not directly observed, but it is identified from differences in behavior between regulated and unregulated firms. I estimate it to be sizeable: it amounts to about 40% of the median entry cost.

Counterfactual exercises show that in the absence of coverage requirements, 3G technology would have been introduced 1.15 year later, on average. Coverage requirements accelerate the introduction of 3G in almost all municipalities, but there is substantial heterogeneity in the magnitude of that effect. For four markets, equilibrium effects imply that the regulation delays the introduction of 3G, though those effects are quantitatively small. The regulation reduces firms' aggregate expected profits by 1.2 billion 2010 USD, or 24.14% of the profits they obtain in the absence of regulation. I find the entry deterrence effects to be small; the overall effect of the regulation is almost equal to its direct effect on the regulated firm.

I also use the model to evaluate alternative policy interventions. I find that a policy that subsidizes the first firm to introduce 3G technology leads to a slightly larger acceleration of its roll-out. Moreover, firms benefit substantially from the subsidy: their aggregate profits increase by 659 million dollars, or 28% of their earnings with no regulation, after accounting for the financing of the subsidy. These gains stem primarily from a more cost-efficient pattern of technology adoption. The subsidy typically leads an incumbent to introduce the new technology, whereas coverage requirements are imposed on potential entrants in many cases. Incumbents only incur technology installation costs, whereas potential entrants also incur entry costs, which I estimate to be sizeable. This difference drives the cost-efficiency gains. Moreover, subsidy recipients also directly benefit from it. The cost efficiencies associated with the subsidy come at the expense of reduced competition in the market. However, I estimate that one more firm in the market has to generate a gain in consumer surplus that exceeds 40% of consumers' average expenditures for coverage requirements to be preferred to the subsidy. These results suggest that subsidization is a more efficient policy than the current form of regulation.

This paper relates to the literature studying how regulation affects market structure and market outcomes in dynamic environments. (Ryan, 2012) shows that stricter environmental regulation increases entry costs, thus decreasing both the number of firms in the market and consumer surplus. (Gowrisankaran, Lucarelli, et al., 2011) study the effect of the Medicare Rural Hospital Flexibility Program on health care provision in rural America, and show that the program expanded coverage but had a net adverse effect on consumer welfare due to provisions that limited the size and scope of regulated hospitals. (Dunne et al., 2013) study the effects of entry subsidies under the Health Professional Shortage Areas program on local market structure. I contribute to this literature by studying the effect of regulation on the set of products (mobile telecommunications technologies) offered by firms and by studying the effects of asymmetric regulation.

This paper also relates to the empirical literature on technology adoption. (Schmidt-Dengler, 2006) studies US hospitals' decisions to adopt magnetic resonance imaging (MRI). (Igami, 2017) studies how cannibalization, preemption, and incumbents' cost advantages shape firms' adoption of a new generation of hard disk drives. My paper adds to this literature by studying how regulation affects technology adoption. Methodologically, my work departs from the previous literature on technology adoption. Models of technology adoption must allow for time-varying demand and adoption costs. The aforementioned papers accommodate this source of non-stationarity and apply full solution estimation methods, based on backward induction algorithms. Backward induction can be applied in these settings due to a finite horizon assumption (Igami, 2017) or full adoption in finite time (Schmidt-Dengler, 2006). I instead model technology adoption as happening in an infinite horizon and assume that the game has a non-stationary part followed by a stationary part. The aforementioned notion of quasi-stationary Markov Perfect Equilibria allows me to adapt existing iterative estimation methods to this non-stationary setting.

My work also relates to the literature on regulation in telecommunications markets. Most recently, (Björkegren, 2019) has studied the adoption of mobile phones in Rwanda, and in that context evaluated the welfare effect of rural coverage requirements imposed on the dominant mobile network operator. His model is one of consumer choice, not firm rollout. I add to this work by modeling how firms respond to the coverage requirements, and moreover by doing so in an oligopoly context. My work also relates to an earlier, mostly theoretical, literature on universal service obligations, such as (Armstrong, 2001), (Choné, Flochel, and Perrot, 2002), and (Valletti, Hoernig, and Barros, 2002), that was motivated by liberalization in the telecommunications industry (and also in the postal services industry) in the 1990s. My work is the first to empirically quantify the effect of such regulation on service provision and the introduction of new technologies.

Methodologically, this paper is related to a long literature on applied dynamic games, going back to (Ericson and Pakes, 1995). The model I will present below will be a dynamic game with discrete controls. A number of estimators have been proposed for stationary dynamic games with discrete controls, e.g., (Aguirregabiria and Mira, 2007), (Pakes, Ostrovsky, and Berry, 2007), and (Pesendorfer and Schmidt-Dengler, 2008). I will depart slightly from that literature in that my model will feature a non-stationary phase followed by a stationary phase. I show that with a cross-section of markets and the notion of Quasi-Stationary Markov Perfect Equilibria, these estimators can be applied to non-stationary settings.

The rest of the paper is organized as follows. Section 1.2 introduces the institutional setting, the data, and presents some preliminary evidence on the effects of coverage requirements on firm behavior. Section 1.3 introduces a model of entry and technology upgrade with regulated and unregulated firms. Section 1.4 discusses the identification and estimation of the model, and also discusses the parameter estimates. Section 1.6 presents the counterfactual analysis. Finally, section 1.7 provides concluding remarks.

1.2 Institutional Setting and Data

Operators of mobile telecommunications networks transmit data through the radio frequency spectrum, which is a public resource and is subject to government management in most countries. Starting in the 1990s, many countries have adopted auctions as their means of allocating frequency bands to firms, including mobile telecommunications service providers. In these auctions, the government sells licenses to use bands of the radio frequency spectrum. These licenses typically come with a number of conditions, chief among them the coverage requirements that are the focus of this paper.

The Brazilian mobile telecommunications market is characterized by 6 mobile network operators (MNO), i.e., carriers that operate their own network infrastructure. There is also a handful of very small mobile virtual network operators (MVNO), which are carriers that do not own their own infrastructure, and instead rent space in one of the MNO's infrastructure. Of the 6 MNOs, four provide service in all of the country and have held licenses covering the entire Brazilian territory since the introduction of mobile telecommunications in the country. The other two MNOs provide more localized service. There has been no entry or exit in this market in the past twenty years.[4]

The Brazilian government conducted its first spectrum auction in 2007 and has since then imposed coverage requirements on the winners of these auctions. For the purpose of this paper, a coverage requirement is an imposition that a firm provide service in some well defined market by a deadline set by the regulator and with a minimum technological requirement (e.g., the firm may be required to provide 4G

[4] In the last couple of years, a process of consolidation has started. Nextel, one of the two small MNOs was sold to Claro, one of the large ones. Oi, one of the big firms, is in the process of being sold, most likely to a consortium formed by the other three large MNOs.

service, or either 3G service or 4G service). In Brazil, the relevant market for the implementation of the regulation is a municipality, and the requirement is considered to be satisfied if that firm provides the designated service in 80% of the municipality's territory. The details of the coverage requirements are a function of municipality population. In municipalities with more than 100,000 inhabitants, 4 MNOs were required to provide 3G service by April 2013; in municipalities with population between 30,000 and 100,000, 3 MNOs were required to provide 3G service by the end of 2017; and in municipalities with population below 30,000, 1 MNO was required to provide 3G service.[5] For the latter group of municipalities, there were four different deadlines: April 2014, April 2016, June 2017, and December 2019.

I focus on the group of municipalities with less than 30,000 inhabitants. The coverage requirements targeting these municipalities are the most likely to influence the availability of service, for in larger municipalities it is probable that firms would have sufficient incentives to enter the market by themselves.[6] I will speak of the single firm in each of these markets that is subject to a coverage requirement as the *regulated firm*; I will refer to the other firms as the *unregulated firms*. All the MNOs are regulated in some markets, but not all. Though these coverage requirements target the introduction of 3G technology, the regulated firm is considered to comply with the regulation if it deploys 4G technology instead. The descriptive analysis in this section uses data from all the municipalities with less than 30,000 inhabitants.[7] The structural analysis will focus on the subset of municipalities with a December

[5]There are also coverage requirements related to 4G technology, but those only apply to municipalities with population above 30,000. There is no 4G coverage requirement in the municipalities with less than 30,000 inhabitants, which are the ones I focus on.

[6]It is likely that the coverage requirements targeting larger municipalities affect the number of firms in the market, but not the availability of service, which is the focus in this paper.

[7]The sample used here is subject to a single sample selection criterion. The regulator has provided me with two different sources of information on the identify of the regulated firm in each market. I keep only the municipalities where these two sources agree with each other.

2019 deadline.[8]

The motivation for coverage requirements rests on two premisses. First, mobile telecommunications services generate substantial welfare gains.[9]. In the words of the Brazilian telecom regulator

> [Mobile telecommunications technologies] *create employment opportunities, improve the education system, increase firm productivity, allow access to public digital services, among other benefits.*[10]

Second, for the intervention to be justified, it must be that firms do not internalize the entirety of the surplus generated by their entry and introduction of new technologies. This seems likely, given the multiple aspects of these benefits and firms' limited ability to price discriminate.

Coverage requirements are enforced by the regulator in a number of ways. First, carriers are required to deposit financial guarantees with the regulator; these guarantees can be executed if the carrier fails to satisfy its coverage requirements. Perhaps more importantly, if a carrier fails to satisfy its coverage requirements, its license can be revoked. In this case, the carrier would also be charged the value paid for its license in proportion to the time used.

The selection of which carrier was to hold the coverage requirement in each municipality was subject to a number of rules. First, the country was divided into 131 "service areas". These varied substantially in size, from a single municipality to an entire code area, which include on average 83 municipalities. Within each of these service areas, one of the four large carriers was required to select 2.5% or 5% of the

[8]This is mostly for computational convenience, as in the structural model the definition of the state space depends on the regulation deadline. A revision of this paper will incorporate data from the other municipalities with less than 30,000 inhabitants.

[9]See, e.g., the references in footnote 2

[10]See https://www.anatel.gov.br/setorregulado/telefonia-movel (last accessed in October 22, 2020).

municipalities in that service area that were subject to the 3G coverage requirements imposed in 2012. The fraction of municipalities to be chosen depended on the license acquired by the firm. The carriers would take turns until all municipalities were chosen. Whenever the number of remaining municipalities in a service area was too small for this rule to be feasible, the regulator decided how many municipalities each carrier would have to choose. Figure 1.1 shows the result of this process. The figure shows a map of the Brazilian midwest, color-coded according to the identity of the regulated carrier. Each subdivision in the map is a municipality. The municipalities with no color were not the subject of the 2012 coverage requirements. All the municipalities in color had to be chosen by some carrier. The noteworthy feature of this figure is that there is no obvious clustering; the municipalities where a firm is regulated are fairly spread out over the map.

The main dataset used in this study comes from ANATEL, the Brazilian telecommunications regulator. The data records at a monthly frequency, for each of the 5,770 municipalities, and for each of the country's mobile network operators whether or not they provide 2G, 3G, and 4G service in that municipality.[11] Figure 1.2 illustrates the structure of the data. The figure shows mobile technology availability in the state of Pará, a relatively poor northern state of Brazil. Each column of the figure corresponds to one of the four major carriers in Brazil and each row corresponds to a year. Within each map, the smaller subdivisions are municipalities in the state of Pará. Municipalities are color coded according to the most advanced technology offered by the corresponding carrier in December of the corresponding year. Therefore, the map in the first row and first column shows the technologies offered in each municipality of the state of Pará by the mobile service provider Claro in December 2013.

The second important piece of data coming from ANATEL is the identity of the

[11] The data does not include MVNOs.

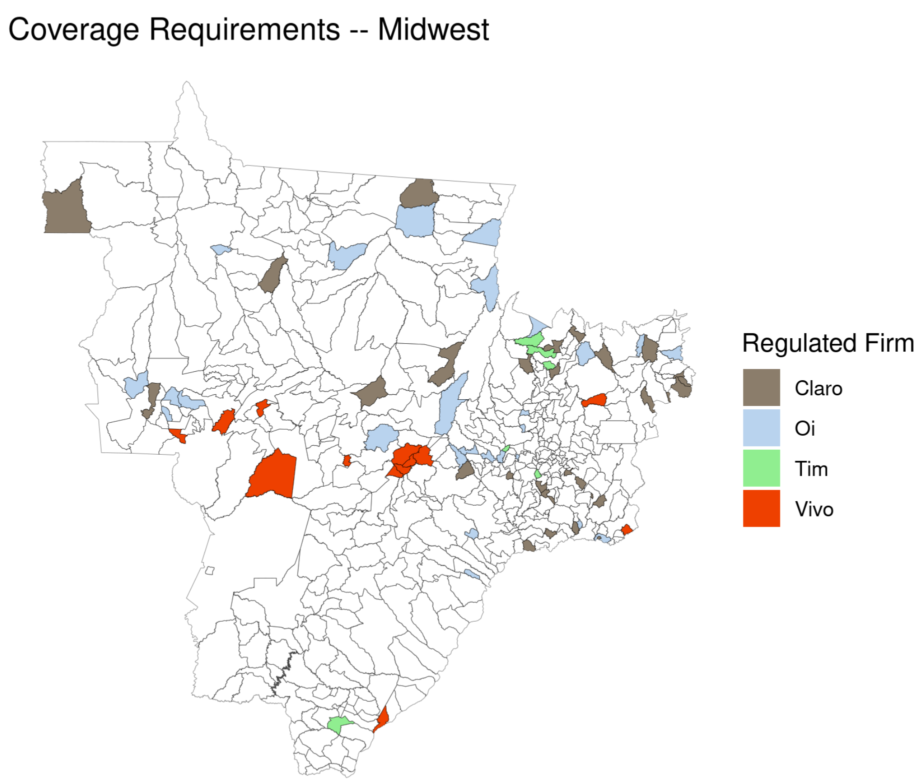

Figure 1.1: Regulated Carriers – Midwest

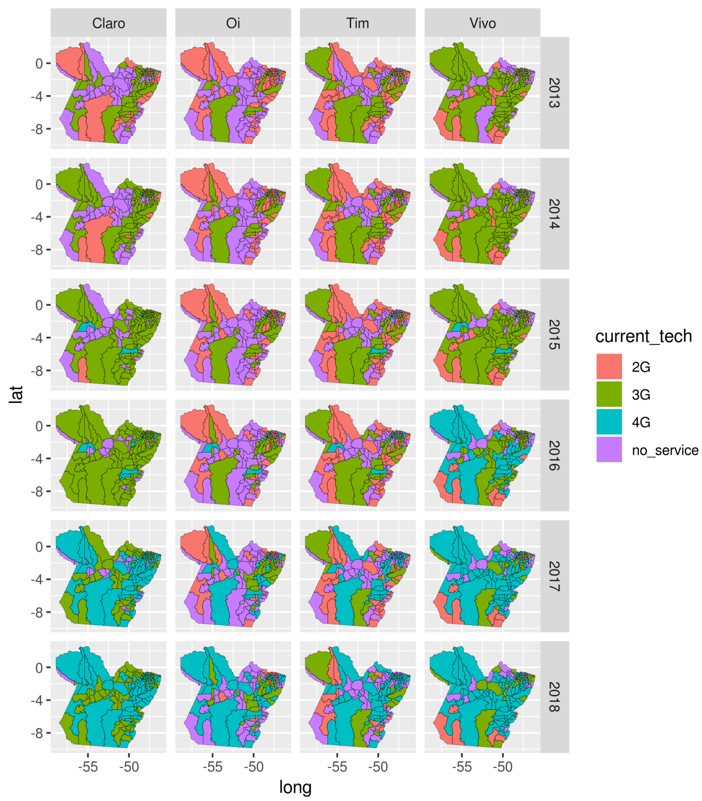

Figure 1.2: Technology availability in the state of Pará

regulated firm in each municipality. Finally, ANATEL also provides data on subscription to mobile telecommunications services. These data are available at the code area-month-carrier-technology level, starting in February 2005 and until December 2018.[12] Figure 1.3 shows the total number of subscribers in the country by technology for the period Jan 2013-Dec 2018. The figure shows that 2G has been in decline over the period, initially being overtaken by 3G. Moreover, 3G reaches a peak in the number of subscribers towards the end of 2015, around the time when the growth of 4G accelerates. To the extent that these patterns are driven by consumer preferences, they shape firms' incentives to introduce new technologies. The empirical model introduced below will account for this pattern in demand by allowing the demand side parameters to vary over time.

I complement the ANATEL data with a number of datasets from the Brazilian Census. First, I utilize municipality demographics and characteristics, such as population, GDP per capita, and area. Summary statistics on these variables are shown in table 1.2. Second, I use the 2017-2018 Family Budget Survey,[13] which provides information on households' income and their expenditure on mobile telecommunications services, among other household characteristics. Third, I use the 2010 Population Census to obtain information on the distribution of individual level demographics at the municipality level.

I drop all code areas where any of the three smaller carriers had a market share of at least 5% at any point in time. I then focus on the four major carriers. Moreover, ANATEL provides two different sources of information on coverage requirements, and I restrict attention to those municipalities for which the two sources of information are consistent with one another. The resulting sample used in the structural analysis

[12] A code area in Brazil is much coarser than a municipality. There are 67 code areas in Brazil, and 5,770 municipalities.

[13] Pesquisa de Orçamentos Familiares.

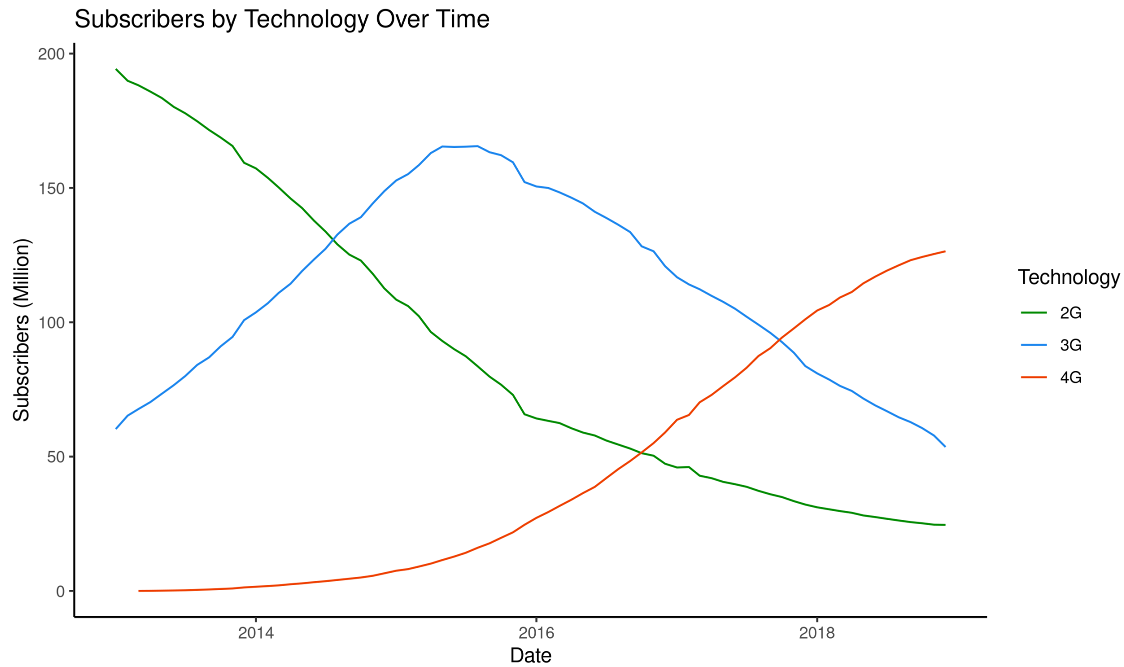

Figure 1.3: Subscribers by technology over time

The figure shows the total number of subscribers in the country, by technology. These quantities are calculated from ANATEL's data on subscription to mobile telecommunications services.

Table 1.1: Summary Statistics – Municipality Characteristics

	Variable	N	Mean	Std. Dev.	p10	p90
1	GDP Per Capita	972	10,936.05	10,016.44	4,248.89	20,573.72
2	Population	972	4,724.12	2,874.62	2,171.06	8,930.41
3	Area	972	1,029.50	3,799.37	90.34	1,746.49

The data in this table comes from the Brazilian Census Bureau. GDP per capita is in 2010 BRLs. Area is in squared kilometers. The values of GDP per capita are averages of data for 2010-2017, deflated to 2010 BRL. The values of population are averages of 2012-2019 data.

Table 1.2: Summary Statistics – Mobile Expenses and HH Characteristics

	Variable	N	Mean	Std. Dev.	p10	p90
1	Mobile Spending	77,655	88.23	166.48	6.74	259.82
2	HH Income PC	77,655	1,687.10	1,556.25	507.56	3,348.01
3	No. Residents	77,655	2.21	1.04	1	4
4	Urban	77,655	0.81	0.39	0	1

The data in this table comes from the 2017-2018 Family Budget Survey. The unit of observation is an individual. Mobile spending is the total amount the individual spent on mobile telecommunications. It is the sum of expenditures on voice and data plans, pre-paid expenditure, and SIM cards. "HH Income PC" is the per capita income in the individual's household. "No. Residents" is the number of residents in the individual's household. "Urban" is a dummy that is equal to 1 if the individual lives in an urban area.

contains 972 municipalities. Furthermore, because entering a market or upgrading a technology is a non-trivial investment that likely involves some time to build, I use data on a semester frequency rather than monthly. The unit of observation is thus a municipality-carrier-semester; there are 46,656 observations.

Table 1.3 shows summary statistics of the data, measured in June 2013 and December 2018, respectively. The tables show statistics for the number of active firms, the number of (firm, technology) pairs available (labeled "products" in the table), whether 3G and 4G are available, and whether the regulated and some unregulated firm offer 3G or 4G technology. In June 2013, there is on average just over 1 firm per market, and about 1.4 products; 3G is available in 28% of municipalities and 4G is not available anywhere. The regulated firm has adopted 3G technology in just over 20% of cases. In 7% of municipalities, an unregulated firm has adopted 3G.

By December 2018, there are just under 1.7 firms per municipality, with about 3.2 products. By December 2018, 3G has reached 88% of municipalities, whereas 4G has reached just under 60% of municipalities. The diffusion of new mobile technologies is driven mostly by regulated firms, but the contribution of unregulated firms is far from

Table 1.3: Summary Statistics

Panel A – June 2013							
Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Number of Firms	972	1.124	0.469	0	1	1	4
Number of Products	972	1.404	0.730	0	1	2	6
3G Available	972	0.277	0.448	0	0	1	1
4G Available	972	0.000	0.000	0	0	0	0
Regulated 3G+	972	0.212	0.409	0	0	0	1
Unregulated 3G+	972	0.068	0.252	0	0	0	1
Panel B – December 2018							
Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Number of Firms	972	1.665	0.655	0	1	2	4
Number of Products	972	3.188	1.550	0	2	4	10
3G Available	972	0.881	0.324	0	1	1	1
4G Available	972	0.580	0.494	0	0	1	1
Regulated 3G+	972	0.807	0.395	0	1	1	1
Unregulated 3G+	972	0.414	0.493	0	0	1	1

Summary statistics across the 972 municipalities in the sample, measured in June 2013 and December 2018. Number of firms is the number of firms active in a municipality. Number of products is the number of (firm,technology) pairs available in a municipality. 3G Available is a dummy that is equal to 1 if at least one firm provides 3G service in the municipality. 4G Available is defined analogously, but for 4G technology. Regulated 3G+ is a dummy that is equal to 1 if the regulated firm provides either 3G or 4G service in the municipality. Unregulated 3G+ is a dummy that is equal to 1 if some unregulated firm provides either 3G or 4G service in the municipality.

negligible: by December 2018, regulated firms have introduced 3G technology (or 4G) in just over 80% of the municipalities in the sample; in 41% of those municipalities, at least one unregulated firm has introduced 3G technology or better.

The descriptive statistics in table 1.3 suggest an important role for coverage requirements in explaining the diffusion of new mobile telecommunications technologies: regulated firms introduce 3G technology (or better) at a faster pace than unregulated firms. This difference is potentially composed of two different effects of coverage requirements: a positive effect on regulated firms and a negative effect on unregulated firms. Unregulated firms may be less likely to enter new markets or upgrade their technologies because they know that the regulated firm will introduce 3G by the requirement deadline. This implies that the market will be more competitive in the future, reducing the incentives for the unregulated firm to enter the market or upgrade its technology.

The data allow me to investigate these positive and negative mechanisms further. I estimate logit models of entry and technology upgrade decisions, which are reported in table 1.4. These models use data on all municipalities with a 3G coverage requirement. An observation in these models is a firm-municipality-date triple. The key explanatory variables in these models are the dummy variables "Regulated", "Regulated Competitor - Out", and "Regulated Competitor - 2G". The first of these variables is equal to 1 when the firm is regulated, and 0 otherwise. The second variable is equal to 1 when the firm faces a regulated competitor that is out of the market. The third variable is equal to 1 when the firm faces a regulated competitor that has 2G technology. The omitted case is when no firm is subject to the regulation.[14] The

[14] To be precise, the ommitted case pools together observations where either the regulated firm has satisfied its coverage requirement or the regulated firm is one of the small firms. Because I restrict the sample to regions where the small firms have always had negligible market shares, I interpret both situations as there being no firm subject to the regulation.

models also control for the municipality's GDP per capita, population, and area, and also include the number of competitors with each technology.[15] Moreover, to account for unobserved municipality level heterogeneity, these models also include group fixed effects, where the groups are defined by a heuristic approach explained in detail in the Section 1.4.[16]

Each column in Table 1.4 corresponds to a different state for the firms included in the sample. The first column includes only observations such that the corresponding firm is not active and includes only data for the years 2013-2015; the second column includes only observations such that the firm is inactive and only data for 2016-2018; the third column includes observations such that the firm offers only 2G technology and data for 2013-2015; the samples for the remaining two columns are similarly defined. The dependent variable for columns 1 and 2 is a dummy that is equal to 1 if the firm enters the market in the next period; the dependent variable for the remaining columns is a dummy that is equal to 1 if the firm upgrades its technology in the following period. There are two key results in Table 1.4. First, regulated firms that have not satisfied their coverage requirements are more likely to enter the market and upgrade their technologies than unregulated firms. Second, unregulated firms are less likely to enter and upgrade their technologies when the regulated competitor is either out of the market or has 2G technology. These results show that the regulation indeed

[15] It may also be expected that a firm's network infrastructure in neighboring municipalities is important for the their choices. I test for that in Appendix 1.B. There I do find that having service in a neighboring municipality increases the probability of entry and technology upgrade. However, the inclusion of those variables changes the estimated coefficients on the other variables only slightly, if at all. This suggests that the choice of the regulated firm is uncorrelated with their local network infrastructure. Characteristics of a firm's network in neighboring municipalities will not be included in the structural model, as doing so would increase the computational burden by several orders of magnitude. The descriptive results discussed here, however, suggest that this omission will not bias my inference regarding the effect of coverage requirements.

[16] The group fixed effects affect the coefficients on the numbers of competitors the most. The other coefficients change only slightly with their introduction. Appendix 1.B shows the results obtaining estimating these models without the group fixed effects.

Table 1.4: Entry/Upgrade Models

	Dependent variable:				
	Out 13-15	Out 16-18	Upgrade 2G 13-15	2G 16-18	3G
	(1)	(2)	(3)	(4)	(5)
Log GDP PC	1.750***	0.970***	0.685***	0.195***	0.181***
	(0.091)	(0.118)	(0.066)	(0.071)	(0.038)
Log Pop.	2.495***	1.997***	1.324***	0.945***	−0.073
	(0.104)	(0.147)	(0.072)	(0.083)	(0.045)
Log Area	−0.507***	−0.386***	−0.291***	−0.322***	0.018
	(0.037)	(0.050)	(0.026)	(0.030)	(0.019)
Regulated	1.735***	2.192***	2.127***	0.870***	−0.397***
	(0.108)	(0.126)	(0.076)	(0.107)	(0.040)
Regulated Competitor - Out	−0.705***	−1.082***	0.116	−0.341**	−0.162
	(0.172)	(0.284)	(0.151)	(0.165)	(0.133)
Regulated Competitor - 2G	0.103	−0.101	−0.522***	−1.199***	−2.333***
	(0.112)	(0.192)	(0.121)	(0.316)	(0.235)
No. Competitors 2G	−1.345***	−1.043***	−0.422***	−0.238***	−0.064*
	(0.093)	(0.117)	(0.055)	(0.067)	(0.038)
No. Competitors 3G	−1.937***	−2.179***	−0.598***	−0.578***	0.211***
	(0.120)	(0.144)	(0.082)	(0.086)	(0.047)
No. Competitors 4G	−1.472	−1.534***	−2.000***	−0.889***	0.426***
	(1.036)	(0.151)	(0.723)	(0.089)	(0.047)
Group FE	Yes	Yes	Yes	Yes	Yes
Observations	36,230	31,620	24,753	14,002	39,923

Note: *p<0.1; **p<0.05; ***p<0.01

accelerates the introduction of the new technology by regulated firms, but also that it *delays* the introduction of new technologies by unregulated firms, which is evidence of the entry deterrence effects outlined in the introduction. Determining which of these two effects dominates and whether or not coverage requirements accelerate the introduction of new technologies is part of the analysis to follow.

The rest of the paper is concerned with developing tools that allows us to quantify the net effect of coverage requirements on the time to adoption of new mobile telecommunications technologies, as well as the entry deterrence effects alluded to above and the costs that the regulation imposes on firms. This requires, we will need a model of how firms make their entry and upgrade decisions. That is the topic of the next section.

1.3 Model

In this section, I introduce an empirical model of mobile service providers' decisions to enter a market and upgrade their technologies. The model operates at the level of a municipality. In the model, firms' flow profits depend on their own technologies, their competitors' technologies, and the local distribution of consumers' demographic characteristics. Inactive firms make irreversible entry decisions, and both entrants and incumbents choose what technologies to offer in the market; firms incur sunk costs of entry and technology upgrade. Because one of the goals of this paper is to understand the effectiveness of coverage requirements as a tool to accelerate the diffusion of new mobile telecommunications technologies, coverage requirements are explicitly modeled. In each market a single firm is required to provide 3G technology by an exogenously specified deadline. If it fails to do so, it pays a fine every period, until it does introduce 3G technology into the market.

There are four carriers in each market. The four carriers compete by choosing which technology to operate, if any. The available technologies are 2G, 3G, and 4G. I assume that firms offer every technology less advanced than their best technology.[17] Time is discrete and the horizon is infinite. Within a period, the timing of the game is as follows. In the beginning of each period t incumbent firms earn their flow profits. Each firm then privately observes action-specific cost shocks, and firms simultaneously decide which of the available actions to take. Potential entrants can enter with any technology and incumbents can choose to upgrade to any technology that is more advanced than their current technology. After choosing an action, firms pay the associated costs. Technologies change deterministically according to firms' decisions.

Let s_{fmt} denote firm f's technology in market m and period t: $s_{fmt} \in \mathcal{S} := \{0, 2, 3, 4\}$, where $s_{fmt} = 0$ denotes that firm f is out of the market and the other values correspond to each of the available technologies, namely 2G, 3G, and 4G. The market's *technological state* $s_{mt} \in \mathcal{S}^4$ is a vector recording each firm's technology. Firms' flow profits are given by a time-varying function of the market's technological state s and the distribution H_x^m of demographics x in market m: $\pi_t(s, H_x^m)$. The specification of π_t is given in subsection 1.3.3.

Entry and upgrade are costly. I will allow the costs of technology introduction to vary over time in a coarse manner. I group periods into two phases; an early phase denoted E (up to and including December 2015) and a later phase denoted L (after December 2015). This allows the model to capture firms' incentives to wait for costs to decrease before introducing a new technology. I will denote by $p(t)$ the phase associated with period t. I model the costs of deploying each technology as a

[17] This assumption is broadly consistent with the data. Mobile service providers typically keep old technologies in place as a fallback option. This assumption also reduces the dimension of the state space considerably, making the model computationally tractable.

technology-phase specific linear function of market characteristics, z_m. Specifically, costs are modeled as

$$c_t\left(a, s_{fmt}, z_m, \varepsilon\right) = \begin{cases} -\varepsilon(a) & \text{if } a = s_{fmt} \\ \sum_{\{g':g'>s_{fmt}\}}^{a} z_m' \theta_{g',p(t)} + \mathbf{1}\left(s_{fmt} = 0\right) z_m' \theta_e - \varepsilon(a) & \text{if } a > s_{fmt} \end{cases} \tag{1.1}$$

In equation 1.1, $a \in \{s_{fmt}, \ldots, 4\}$ is the action chosen by the firm and $\varepsilon(a)$ is an action-specific cost shock; ε is a vector collecting all the $\varepsilon(a)$. If $a = s_{fmt}$, the firm pays no costs (other than receiving the cost shock). A potential entrant that decides to enter pays an entry cost $z_m'\theta_e$. Moreover, associated with every technology g there are installation costs $z_m'\theta_{g,p(t)}$. One can interpret $z_m'\theta_e$ as the cost of installing basic infrastructure, such as cell phone towers. Because it is associated with basic infrastructure, this cost does not vary over time. The term $z_m'\theta_{g,p(t)}$ captures the cost of installing technology-specific infrastructure, such as radios that only transmit 3G or 4G signal. Because this term is associated with new technologies, it is allowed to vary over time. In equation (1.1), z_m is a vector of observed market characteristics and the θ's are parameters to be estimated. The summation in equation (1.1) reflects the previous assumption that firms offer all technologies less advanced than their best technology. If, for example, a firm's current best technology is 2G, and that firm upgrades to 4G, equation (1.1) says that the firm will pay the costs of installing both 3G and 4G.[18] The cost shocks are assumed to follow a Type 1 Extreme Value distribution with scale parameter σ, and they are iid across firms, periods, and actions.

In each market m, exactly one firm is required to provide 3G service or better by a date T_m exogenously specified by the regulator.[19] I will call that firm the *regulated*

[18]Note that this implies that an entering firm will always offer 2G. Because the cost of installing 2G is only paid by an entering firm, θ_e and θ_{2G} will not be separately identified. Therefore, in estimation I drop θ_{2G}. The estimate of θ_e thus includes both entry costs and 2G installation costs.

[19]In the empirical application, T_m is always equal to December 31, 2019.

or *committed* firm and the other firms the *unregulated* or *uncommitted* firms. If the regulated firm fails to provide at least 3G service by the date T_m, it pays a fine φ every period, starting in $T_m + 1$ and until it deploys either 3G or 4G.

Firms choose their actions to maximize their discounted expected profits, taking their competitors' behavior as given. I focus on Markov Perfect Equilibria (MPE), as is typical in empirical applications of dynamic games. I allow regulated and unregulated firms to behave differently, but beyond that I impose symmetry.

There are two sources of non-stationarity in this environment. First, flow profits and entry and technology upgrade costs vary over time. Second, coverage requirements also imply that firm behavior depends on the date. Suppose that the regulated firm has not satisfied its commitment and $t < T_m$; as time goes by, the regulated firm gets closer to being fined and therefore should become more likely to introduce 3G technology. Conditional choice probabilities thus change over time. I now discuss symmetry and non-stationarity in turn.

1.3.1 Symmetric Markov Perfect Equilibria

A Markov Perfect Equilibrium is a strategy profile $(\sigma_1, \ldots, \sigma_4)$, such that σ_i is a function that maps a firm's state variables into a feasible action. In a symmetric Markov Perfect Equilibrium, strategies don't depend on firms' identities. Instead, I define value and policy functions for regulated and unregulated firms. To simplify the notation, I subsume all the market-specific variables that do not vary over time in a superscript. The state of an unregulated firm is $(s_1, s_r, s_-, t, \varepsilon)$, where s_1 is that firm's technology, s_r is the technology of the regulated firm, and s_- is a vector with the technologies of the other two firms. The state of a regulated firm is $(s_1, s_-, t, \varepsilon)$ where now s_- denotes the technologies of the three remaining firms. Let Ω_0, Ω_1 denote the state space for unregulated and regulated firms, respectively, with typical element

$\omega_r, r \in \{0,1\}$. A strategy is a function $\sigma_r : \Omega_r \to \{0,2,3,4\}$ satisfying the restriction that $\sigma_r(\omega_r) \in A(s_1(\omega_r)) := \{s_1(\omega_r), \ldots, 4\}$, where $s_1(\omega_r)$ is the first coordinate of ω_r.[20]

Let $\sigma^m = (\sigma_0^m, \sigma_1^m)$ be a symmetric strategy profile. Define the implied ex-ante value function

$$V_{r,\sigma}^m(s,t) := \mathbb{E}_\varepsilon \Big\{ \sum_{\tau=t}^{\infty} \delta^{\tau-t} \Big[\pi_\tau^m(s_{f\tau}, s_{-f,\tau}) - c_\tau^m\left(a_{f\tau}, s_{f\tau}\right) + \\ + \varepsilon_{f\tau}\left(a_\tau\right) - \varphi r \mathbf{1}\left(T_m < \tau, s_{f\tau} < 3\right) \Big] \Big| r, s, t; \sigma \Big\}$$

where $\mathbb{E}_\varepsilon$ indicates that the expectation is taken over the sequence of ε's for all firms; firms' states evolve according to (σ_0^m, σ_1^m).

Symmetry implies restrictions on $\sigma_0, \sigma_1, V_0, V_1$. For example, for a regulated firm, it is irrelevant whether $s_- = (3,2,1)$ or $s_- = (1,2,3)$. Therefore $V_1(s_1,3,2,1) = V_1(s_1,1,2,3)$ and similarly for the policy function σ_1. Similar restrictions apply to unregulated firms. Furthermore, symmetry implies that V_0 and V_1 are equal for some states. For example, suppose that $s_j = s_r \geq 3$. Then $V_0(s_j, s_r, s_-) = V_1(s_r, s_j, s_-)$. Symmetry implies further restrictions on value and policy functions. Appendix 1.C presents all of those restrictions and how they're used to efficiently represent firms' state spaces.

Finally, note that the recursive characterization of Markov Perfect Equilibria (e.g.,

[20]The idiosyncratic nature of the regulated firm's technology is the reason why I don't define the state variable to be given by the number of competitors with each technology. The model could be equivalently represented in that way, but given that it is necessary to keep track of the regulated firm's technology, it is simpler to keep track of all firms technologies and impose the appropriate symmetry conditions.

(Doraszelski and Escobar, 2010)) implies that $\{\sigma_0^m, \sigma_1^m\}$ is a MPE if and only if

$$\sigma_r^m(s,t,\varepsilon) = \underset{a \in A\left(s_f\right)}{\operatorname{argmax}} \left\{ \pi_t^m(s_{ft}, s_{-f,t}) - c_t^m(a, s_f) + \delta \mathbb{E}_{\varepsilon_{-f}} \left[V_{r,\sigma}^m \left(a, s'_{-f}, t+1\right) | r, s, t \right] + \varepsilon(a) \right\} \tag{1.2}$$

where, for firms $h \neq f$, $s'_h = \sigma_{r_h}^m(s,t,\varepsilon_h)$ and the expectation is with respect to the shocks ε_h of firms $h \neq f$.

1.3.2 Quasi-Stationary Markov Perfect Equilibria

As discussed above, there are two sources of non-stationarity in this environment. First, flow profits and cost parameters change over time. Second, coverage requirements imply that firms' policy functions respond to the proximity of the requirement expiration date T_m. In this subsection, I introduce assumtions that accommodate these two sources of non-stationarity, but impose a degree of stationarity.

The specification of entry and technology upgrade costs assumes that those costs eventually stabilize.[21] I will assume the same of flow profits. Specifically, I assume that flow profits vary in a way known to the firms from the start of my sample until the beginning of 2018, after which they stabilize. I then make two assumptions regarding equilibrium behavior. First, after parameters have stabilized and the expiration date of the coverage requirement has passed, behavior doesn't depend on the date anymore. Second, the same is true if parameters have stabilized and the committed firm has satisfied its commitment.

Formally, I focus on *Quasi-stationary Symmetric Markov Perfect Equilibria*, defined below. Let T_θ denote the earliest time period such that flow profits and costs do not vary after T_θ.

[21] Entry and technology upgrade costs vary between the early and the late phases, but do not change after that.

Definition 1. *A Symmetric Markov Perfect Equilibrium* (σ_0, σ_1) *is said to be quasi-stationary if there exist functions* $\tilde{\sigma}_r(s, \varepsilon), r \in \{0, 1\}$*, such that, if either*

(i) $t \geq \max\{T_m + 1, T_\theta\}$*, or*

(ii) $t \geq T_\theta$ *and* $s_r \geq 3$*,*

then $\sigma_r(s, t, \varepsilon) = \tilde{\sigma}_r(s, \varepsilon)$*.*

I assume throughout that the data is generated by a Quasi-Stationary Symmetric Markov Perfect Equilibrium. Note that this imposes restrictions on value functions over time. For example, if $t \geq T_\theta$ and $s_j = s_r \geq 3$, then $V_1(s_r, s_j, s_-, t) = V_0(s_j, s_r, s_-, t + 1)$. Essentially, the model has a non-stationary phase followed by a stationary phase. Models of technology adoption must somehow contend with the fact that the demand for and costs of adopting a new technology vary over time. One way of dealing with the time-varying nature of demand and costs that appears in the literature is to assume a finite horizon and solve the game played by firms via backward induction; see, e.g., (Igami, 2017). That method raises the issue of assigning continuation values to different industry states in the final time period. In (Igami, 2017), that is done by assuming that the state of the industry doesn't change after the terminal period, and computing the implied discounted stream of profits. Quasi-stationarity instead assumes that firms will keep playing the entry and technology upgrade game forever, so that firms' continuation values are given by the equilibrium value function in the relevant states. The empirical feasibility of this assumption rests on observing a cross-section of markets.

1.3.3 Modeling Flow Profits

It is not uncommon in applications of dynamic games for flow profits to be derived from an estimated demand system paired with an assumption on firms' pricing behavior. Following that route would require data on available plans, their prices, and consumers' choices from the available plans. Unfortunately, such data is not available in my setting. I thus follow a different approach. Suppose that consumer i in market m with demographic characteristics x_i chooses what carrier to subscribe to, what technology to use, and how much to spend on mobile telecommunications services, e_i. Let $\sigma_{fgt}(s, H)$ be the resulting market share of firm-technology pair (f, g) in period t when the industry state is s and the distribution of demographics is H; a model for σ_{fgt} will be specified below. Let M be the size of the market and, as before, let s_f be firm f's state.[22] Finally, denote by $\mathbb{E}[e_i|g]$ the expectation of consumers' expenditures e_i, conditional on a consumer choosing technology g.[23] Firms' profits are then given by:[24]

$$\begin{aligned} \pi_t(s_f, s_{-f}, H) &= M \sum_{g \in s_f} \sigma_{fgt}(s, H) \left(\mathbb{E}_t[e_i|g] + \psi \right) \\ &= M \sum_{g \in s_f} \sigma_{fgt}(s, H) \left(\int \mathbb{E}[e_i|g, x_i] dH_t(x_i|g) + \quad \right) \end{aligned} \tag{1.3}$$

The summation over $g \in s_f$ indicates that we sum over all technologies offered by firm f: $\{g : 0 < g \leq s_f\}$. The parameter ψ captures revenues that the expenditure model may fail to account for and marginal costs of serving customers. In estimation, I will allow ψ to vary by groups of markets; see section 1.4 for details. Note that in equation (1.3), the conditional distribution $H_t(x_i|g)$ is indexed by t. That is because

[22] I set the market size to be twice the population of the municipality. The number of mobile telecommunications subscriptions in Brazil is larger than the population.

[23] Here I condition only on the chosen technology, and not on the firm identity, because firms are assumed throughout to be symmetric.

[24] The expression in the right hand side of 1.3 is an approximation. Firms' profits are equal to $\sum_{g \in s_f} \sum_{i \in fg} e_i$, where the summations are over the technologies offered by firm f and over individuals i subscribing to firm-technology pair (f, g). This approximation holds in the sense that the difference between firms' profits and the right hand side of equation 1.3 is $O_p(\sqrt{M})$, whereas the included term is $O(M)$. This implies that the approximation error becomes negligible relative to the included term for large M. This approximation is analogous to the (implicit) approximation to profit functions used routinely in supply and demand models in empirical industrial organization.

consumers' preferences over technologies are allowed to vary over time (as indicated by the t subscripts in σ_{fgt}), so that the distribution of demographics conditional on technology choice also varies over time.

The main data limitation I face is that I never observe consumer expenditures together with their technology (and carrier) choices. I will therefore make the following assumption:

Assumption 1. $\mathbb{E}[e_i|g, x_i] = \mathbb{E}[e_i|x_i]$.

This assumption says that conditional on individual characteristics x_i, consumer expenditure is mean independent of the technology chosen by that consumer. This is, admittedly, a strong assumption. It would hold, e.g., in a world in which consumers pay per usage (a popular model in Brazil), and technology doesn't affect usage. This assumption would fail if better technologies induce consumers to utilize more data. Assumption 1 would thus be untenable if we were dealing with a population that uses high-bandwith applications. Because we are dealing with small, rural municipalities in Brazil, the assumption is more palatable. Importantly, note that Assumption 1 does not imply that consumers that subscribe to different technologies will spend (on average) the same amount, for individuals with different demographic characteristics are still allowed to sort into different technologies.

Assumption 1 and equation 1.3 imply that

$$\pi_t(s, H) = M \sum_{g \in s_f} \sigma_{fgt}(s, H) \left(\int \mathbb{E}[e_i|x_i] dH_t(x_i|g) + \psi \right) \tag{1.4}$$

I model $\sigma_{fgt}(s, H)$ as arising from a nested logit model. Specifically, consumer i's utility of subscribing to firm-technology pair $j = (f, g)$ in market m and year τ is given by[25]

$$u_{ijm\tau} = \underbrace{\gamma_{r(m),p(\tau)} + \mu_{g(j),p(\tau)} + \beta_{g(j),p(\tau)} y_{m\tau} + \theta_{g(j),p(\tau)} d_{m\tau}}_{v_{g(j)m\tau}} + \xi_{jm\tau} + \zeta_{im\tau}(\lambda) + (1-\lambda)\varepsilon_{ijm\tau} \tag{1.5}$$

where $r(m)$ is the state of municiality m, $p(\tau)$ is the phase (early or late) associated with year τ, $y_{m\tau}$ is GDP per capita, and $d_{m\tau}$ is population density.[26] The term $\xi_{jm\tau}$ is an unobserved product characteristic, $\zeta_{im\tau}(\sigma)$ is a disturbance common to all goods other than the outside good, and $\varepsilon_{ijm\tau}$ is a Type 1 Extreme Value shock. The parameter λ is the nesting parameter, and $\zeta_{im\tau}(\lambda)$ has the unique distribution such that $[\zeta_{im\tau}(\lambda) + (1-\lambda)\varepsilon_{ijm\tau}]$ also has an extreme value distribution (Cardell, 1997).

In equation (1.5), $\gamma_{r(m),p(\tau)}$ is a state-phase fixed effect meant to capture variation in the share of the outside good; $\mu_{g(j),p(\tau)}$ is a technology-phase fixed effect, which captures changes in the popularity of each technology over time; and the effect of income and population density on consumer preferences is also allowed to vary by technology and phase.

The distributional assumptions above imply that market shares are given by

$$\sigma_{jm\tau}(s, v_{m\tau}, \xi_{m\tau}) = \frac{e^{(v_{g(j)m\tau}+\xi_{jm\tau})/(1-\lambda)}}{D} \times \frac{D^{1-\lambda}}{1+D^{1-\lambda}} \tag{1.6}$$

where $v_{m,\tau}$ is a vector collecting the $v_{gm\tau}$, $\xi_{m\tau}$ is a vector similarly defined, and $D := \sum_{j\in s} e^{(v_{g(j)m\tau}+\xi_{jm\tau})/(1-\lambda)}$, where the summation is over the products offered in the market, which are encoded in the industry state s. The predicted quantity of subscribers is $M\sigma_{jm\tau}(s)$.

[25] I specify equation 1.5 at the year level because the demographic variables in it are observed with that frequency. A period in the model, which corresponds to six months, is mapped to its corresponding year and the choice model introduced in the text is used to compute market shares.

[26] Ideally, y_i should be used in equation 1.5. That would add one more integration in the estimation routine. Doing so is work in progress. In the analysis that follows, when calculating $H(x|g)$, I will treat the coefficient on $y_{m\tau}$ as the effect of an individual's income on her utility.

It remains to model $\mathbb{E}[e_i|x_i]$. I assume that individual i's, e_i, is given by

$$\log(e_i) = \alpha_{r(i)u} + \alpha_1 \log(y_i) + \alpha_2 n_i + \eta_i \tag{1.7}$$

In equation (1.7), $r(i)$ indicates i's state of residence; u indicates whether the municipality is classified as urban or rural by the Census; y_i is income; n_i is the number of residents in i's household; and η_i is an error term that is uncorrelated with the included regressors. We now have all the ingredients needed to compute firms' profits in equation 1.4, except for the distribution $H(x_i|g)$. I obtain that distribution using the technology choice model outlined above and Census data on municipality-level demographics; for details, see section 1.4.

The final aspect of the model is an assumption regarding the distribution of $\xi_{jm\tau}$. I introduce this assumption to deal with the fact that the I observe the quantities of subscribers at different levels of geographic granularity over time; see section 1.4 for details.

Assumption 2. *Let $c(m)$ denote the area-code that municipality m belongs to. The unobserved product characteristic $\xi_{jm\tau}$ satisfies*

$$\xi_{jm\tau} = \xi_{jc(m)\tau} + \eta_{jm\tau}$$

where $\eta_{jm\tau} \overset{iid}{\sim} F$.

Assumption 2 says that $\xi_{jm\tau}$ can be decomposed into a random variable that varies only with area-code, on which I place no restrictions, and another RV that varies across municipalities within an area-code, that I assume is *iid* with some unrestricted distribution F.

Under Assumption 2, an argument relying on a large number of municipalities within an area-code implies that

$$\sigma_{jc\tau} = \sum_{m \in c} \omega_m \int \sigma_{jm\tau}(s_{m\tau}, v_{m\tau}, \xi_{c(m),\tau}, \eta_{m\tau}; \theta) dF(\eta_{m\tau}) \tag{1.8}$$

holds approximately. In equation (1.8), ω_m is the fraction of the population in area-code c in municipality m. I will use equation (1.8) in estimation; see section 1.4.

1.4 Identification and Estimation

I start this section by discussing the estimation of the flow profit function in subsection 1.4.1. In subsection 1.4.2 I discuss the estimation of the dynamic parameters of the model, i.e., the entry and upgrade costs and the fine for non-compliance with the regulation.

1.4.1 Estimation of the Flow Profit Function

The flow profit function is given by equation (1.4). Computing profits requires four objects: $\sigma_{fgt}(s, H)$, $\mathbb{E}[e_i|x_i]$, $H_t(x_i|g)$, and ψ. In this subsection, I discuss the estimation of the first three of these objects.

The first task is to estimate the parameters underlying the market share terms, $\sigma_{fgt}(s, H)$. Here I have to deal with the fact the data on mobile subscriptions come at different levels of geographic granularity over time. First, equation (1.6) implies the usual analytical nested logit inversion (Berry, 1994):

$$\log(s_{jmt}) - \log(s_{0mt}) = v_{g(j)mt} + \lambda \log(s_{j|\mathcal{J}_{mt}}) + \xi_{jmt} \tag{1.9}$$

where $\log(s_{j|\mathcal{J}_{mt}})$ is the share of good j in the total number of subscriptions in the market. This equation yields ξ_{jmt} as a function of data and parameters, $\xi_{jmt}(\theta)$. I interact $\xi_{jmt}(\theta)$ with instruments to form moment conditions $\mathbb{E}[\xi_{jmt}(\theta)Z^1_{jmt}] = 0$.

The intuition for the identification of the nesting parameter λ is similar to that in (Berry and Waldfogel, 1999). The nesting parameter determines the extent of business stealing when a new product enters the market. If we can exogenously vary the number of products in the market, we learn the value of λ by observing the effect on the aggregate share of the goods in the market. Following this intuition, I use as instruments for $\log(s_{j|\mathcal{J}_{mt}})$, the logarithm of the area of municipality m, and dummies for whether or not the municipality is one of the regulated ones, interacted with the regulation deadline. The area of a municipality increases the cost of providing service, and thus reduces the number of products in the market. Regulated municipalities with early regulation deadlines will tend to have more products than regulated municipalities with later deadlines. The identifying assumption is that the regulation deadlines are uncorrelated with unobservable product characteristics in 2019. I also use the demographic variables in v_{jmt} as instruments.

The moments discussed above are informative about the nesting parameter and preference parameters in the later period of the data, but not about preference parameters in the earlier period of the data. To construct additional moments to identify those parameters, I leverage assumption 2 and equation (1.8). Equation (1.8), repeated here for convenience, states that market shares at the area-code level are approximately given by

$$\sigma_{jct} = \sum_{m \in c} \omega_m \int \sigma_{jmt}(s_{mt}, v_{mt}, \xi_{c(m),t}, \eta_{mt}; \theta) dF(\eta_{mt}) \tag{1.10}$$

Equating observed market shares at the area-code level with their predicted counterparts, given by the right hand side of equation 1.11, allows one to solve for ξ_{jct} as a function of all the utility parameters. These structural error terms, $\xi_{jct}(\theta)$, could then be interacted with instruments to form moment conditions of the form $\mathbb{E}[\xi_{jct}(\theta)Z^2_{jct}] = 0$. The one hindrance to that approach is the integration with respect to $F(\eta_{jmt})$. Here, again, assumption 2 offers a solution. Given any vector of structural parameters, θ, equation (1.9) gives us $\xi_{jmt}(\theta)$. We can then make use of assumption 2 to recover $\eta_{jmt}(\theta)$, which gives us an empirical distribution of η_{jmt} given θ, $F(\eta;\theta)$. In this way, the integration in equation (1.11) can be performed for any guess of θ by sampling from the implied $F(\eta;\theta)$, and moment conditions can be formed as outlined above.

To summarise the preceding discussion, the steps involved in evaluating the GMM objective function for a given value of θ are as follows. First, use equation (1.9) to obtain $\xi_{jmt}(\theta)$. Second, use assumption 2 to obtain $\eta_{jmt}(\theta)$. Third, solve for $\xi_{jct}(\theta)$ from

$$s_{jct} = \sum_{m \in c} \omega_m \frac{1}{N_s} \sum_{i=1}^{N_s} \sigma_{jmt}(s_{mt}, v_{mt}, \xi_{c(m),t}, \eta_i; \theta) \tag{1.11}$$

where s_{jct} is the observed market share of firm-technology pair j in area-code c and period t, η_i is a vector of $|\mathcal{J}_{mt}|$ independent draws from $F(\eta;\theta)$ and N_s is the number of simulation draws. Fourth, interact ξ_{jmt} with Z^1_{jmt} and ξ_{jct} with Z^2_{jct} and average, to get sample analogs of the moment conditions discussed above; call these sample analogs $\bar{g}^1(\theta)$ and $\bar{g}^2(\theta)$, respectively. For a chosen weight matrix W, the GMM objective is then given by

$$J(\theta) := \begin{pmatrix} \bar{g}^1(\theta)' & \bar{g}^2(\theta)' \end{pmatrix} W \begin{pmatrix} \bar{g}^1(\theta) \\ \bar{g}^2(\theta) \end{pmatrix} \tag{1.12}$$

The GMM estimator is, as usual, $\hat{\theta} := \operatorname{argmin}_\theta J(\theta)$. I have discussed the instruments Z^1_{jct} above. The instruments Z^2_{jct} used in estimation are the population-weighted averages of the demographics included in v_{gmt}. I use the identity matrix as the weighting matrix in estimation.

The term $\mathbb{E}[e_i|x_i]$ in equation (1.4) is calculated from equation (1.7), which is estimated by ordinaty least squares using the Household Budget Survey. From (1.7) it follows that $\mathbb{E}[e_{im}|x_i] = \exp(\alpha_{r(m)u} + \alpha_2 n_i) y_i^{\alpha} \mathbb{E}[\exp(\eta_{im})|x_i]$. I assume that $\exp(\eta_{im})$ is mean independent of x_i and estimate $\mathbb{E}[\exp(\eta_{im})]$ using the residuals from equation (1.7).

The last ingredient needed to use equation (1.4) is the conditional distribution $H(x_i|g)$. By Bayes' rule,

$$h(x_i|g) = \frac{\sigma(g|x_i)h(x_i)}{\int \sigma(g|x_i')h(x_i')dx_i'} \tag{1.13}$$

The term $\sigma(g|x_i)$ is derived from the technology choice model; the unconditional distribution of x_i comes from the Census data. I obtain $h(x_i|g)$ by drawing a uniform random sample from the municipality-level Census data, computing $\sigma(g|x_i)$ for each drawn x_i, and calculating $\sigma(g|x_i)/\sum_j \sigma(g|x_j)$.

The final object in equation (1.4) is the parameter ψ. I will allow the value of ψ to vary across five groups of municipalities. Those groups are determined in the following heuristic way. First, I project the number of firm-technology pairs in municipality m and period t onto municipality and time dummies. Next, I run a linear regression of the estimated municipality fixed effects on averages over time of the municipality characteristics included in the structural model. The residuals from these regressions can be thought of as time-invariant unobserved factors that determine the number of products in a market, and hence are related to profitability in that market. I group municipalities according to the quintiles of the distribution of these residuals and estimate a ψ for each group. These parameters will be estimated together with the dynamic parameters of the model. That is the topic of the next subsection.[27]

1.4.2 Identification and Estimation of Dynamic Parameters

The flow payoffs of the dynamic game introduced in the previous section are linear in the structural parameters. For this class of models (dynamic games with linear flow payoffs), it is possible to show that structural parameters are identified if conditional choice probabilities are identified.[28] The requirement that conditional choice probabilities be identified excludes from this result general models with unobserved state variables However, this result encompasses models where the unobserved state variables possess a group structure and that group structure can be recovered from the data in a first stage, as is assumed here.

The conditional value functions inherit the linearity from the flow payoffs: there exist functions $f_{rt,P^m}(a,s)$ and $g_{rt,P^m}(a,s,z)$ such that

$$\frac{v^m_{r,t}(a,s)}{\sigma} = f_{rt,P^m}(a,s) + g_{rt,P^m}(a,s,z)\sigma^{-1}\Psi$$

where Ψ is a vector collecting all structural parameters (see Appendix 1.D for details). This fact can be used to establish identification.

Since the idiosyncratic errors follow a Type 1 Extreme Value distribution, the

[27]The heuristic procedure discussed in this subsection is related to approaches taken by (Collard-Wexler, 2013) and (Sanches, Silva-Junior, and Srisuma, 2018) to account for unobserved heterogeneity. A recent literature in econometrics has introduced methods to deal with group fixed effects in panel data and structural models. On this, see, e.g., (Bonhomme and Manresa, 2015), (Bonhomme, Lamadon, and Manresa, 2017), and (Cheng, Schorfheide, and Shao, 2019). It is possible that those methods can be adapted to deal with group-level unobserved heterogeneity in dynamic games. A more common approach of dealing with market-level unobserved heterogeneity would be to apply an EM-type algorithm. That approach, however, would require making the strong assumption that unobserved heterogeneity is independent across markets, which seems unlikely in the present case.

[28]This is a known result, see, e.g., (Aguirregabiria and Nevo, 2013). I review the argument here for completeness.

conditional choice probabilities have the logit form:

$$P^m(a|s,r,t) = \frac{\exp(v^m_{r,t}(a,s)/\sigma)}{\sum_{a' \in A(s_f)} \exp(v^m_{r,t}(a',s)/\sigma)}$$

We can apply the usual logit inversion to this equation to obtain:

$$\ln(P^m(a|s,r,t)) - \ln(P^m(s_f|s,r,t)) = \frac{v^m_{r,t}(a,s)}{\sigma} - \frac{v^m_{r,t}(s_f,s)}{\sigma}$$

Using the linear representation of the conditional value functions we can then write

$$\ln(P^m(a|s,r,t)) - \ln(P^m(s_f|s,r,t)) - f_{rt,P^m}(a,s) - f_{rt,P^m}(s_f,s) = \Big[g_{rt,P^m}(a,s,z) - g_{rt,P^m}(s_f,s,z) \Big]' \frac{\psi}{\sigma} \tag{1.14}$$

Equation (1.14) leads to an OLS-like formula for ψ/σ.[29]

The intuition for identification is that the structural parameters are identified by exogenous variation in (π_m, z_m, s, r, t) and the fact that we observe how firms respond to this variation (i.e., we observe conditional choice probabilities). One can, for example, entertain the thought experiment of varying one of the exogenous covariates and observing how the behavior of firms changes. If we vary the distribution of income, for example, flow profits will vary; the extent to which firms respond in their entry and upgrade behavior is informative about the costs of such actions.[30] The fine parameter

[29] This argument has used market-specific CCPs P^m. This is not necessarily inconsistent with the typical assumption that a unique equilibrium is played in the data, as one can simply enlarge the state space to include market-level characteristics and define policy functions on that domain. Either way, those CCPs must be estimable from data, and we therefore require that the equilibria played in the data (or the unique equilibrium defined on an enlarged state space) vary continusouly with the market-level characteristics.

[30] Although useful, this intuition is slightly imprecise. When we vary the distribution of income, the endogenous conditional choice probabilities P^m will also change, thus changing the other terms in $w^m_{rt,P^m}(a,s,z_m)$. This makes clear that functional form assumptions play a role in obtaining identification in dynamic games, which is why all empirical models in this literature are tightly parameterized.

φ is identified by the difference in behavior between regulated and unregulated firms. Time variation provides additional variation to identify φ. Intuitively, for small φ the behavior of regulated firms will change only slightly as the regulation deadline approaches; large φ, on the other hand, should lead to larger changes in behavior.

1.4.3 Estimation

I apply the Nested Pseudo Likelihood (NPL) algorithm of (Aguirregabiria and Mira, 2007) to estimate the dynamic parameters. In light of the results of (Pesendorfer and Schmidt-Dengler, 2010), my choice of estimator requires some justification. A popular alternative is to use a two-step estimator, e.g. (Bajari, Benkard, and Levin, 2007), (Pakes, Ostrovsky, and Berry, 2007) or (Pesendorfer and Schmidt-Dengler, 2008). These estimators all proceed by flexibly estimating policy functions in a first stage and then using those policy functions to construct a second-stage objective function that is then minimized to yield structural estimates. Because my model features substantial non-stationarity, it would be challenging to obtain flexible and accurate first stage estimates of policy functions. For this reason, I opt to use an estimator that makes full use of the already imposed structural assumptions.

As is well known, the computational cost of the maximum likelihood estimator is prohibitive in the case of dynamic games. I thus adopt (Aguirregabiria and Mira, 2007). An alternative that was recently proposed is (Dearing and Blevins, 2019). The estimator proposed by (Dearing and Blevins, 2019) enjoys good theoretical properties. In particular, it is guaranteed to converge, thus overcoming the main issue raised of NPL raised by (Pesendorfer and Schmidt-Dengler, 2010). However, the algorithm in (Dearing and Blevins, 2019) requires solving large systems of linear equations, which renders its application to the empirical setting in this paper substantially more costly than (Aguirregabiria and Mira, 2007).

A Nested Pseudo Likelihood (NPL) fixed point is a pair $(\tilde{\theta}, \{\tilde{P}^m\}_m)$ that satisfies

(i) $\tilde{\theta} = \text{argmax}_\theta \sum_{m,t,f} \ln \Psi(a_{fmt}|s_{mt}, r_{fm}, t, m; \theta, \tilde{P}^m)$

(ii) $\tilde{P}^m = \Psi(\tilde{P}^m; \tilde{\theta})$ for all m

The NPL *estimator* is the NPL fixed point with the maximum value of the pseudo-likelihood. The set of NPL fixed points is known to be non-empty. However, it need not be a singleton. This implies that the researcher must explore the parameter space to ensure that the pseudo-likelihood is being maximized in the set of NPL fixed points.

In practice, one finds NPL fixed points via an iterative algorithm. Starting with a guess for CCPs, $\{\tilde{P}^m\}_m$, the implied pseudo likelihood is maximized (see condition (i) above). One then uses the resulting guess for θ to update firms' CCPs (see condition (ii)) above. These two steps are repeated until the CCPs or the structural parameters converge.

1.5 Estimation Results

Table 1.5 presents the estimates of the static parameters, those in the market shares and expenditure functions. The results show that the market share of 4G is, in both the early and the later periods, the most responsive to income, suggesting that richer individuals have higher demand for high-bandwith uses of mobile communications. The market share of 4G also grows the most with population density in the earlier part of the sample. This is consistent with individuals in more densely populated areas having more social connections and therefore having higher demand for faster connectivity. Surprisingly, this pattern is more muted in the later part of the sample. Estimates of the expenditure model show that richer individuals spend more on mobile telecommunications, as one would expect. Mobile expenses also increase in the number of residents in the household; this is consistent with the notion that individuals in larger families have more reason to communicate and are therefore more active users of mobile telecommunications services.

Table 1.5: Static Parameter Estimates

	Market Shares, Early			Market Shares, Late			Expenditures
	2G	3G	4G	2G	3G	4G	
Intercept	7.718	6.532		3.740	4.001		
Log Income	0.211	0.421	0.819	-0.205	-0.057	0.366	0.356 (0.003)
Pop Dens.	0.269	0.341	0.423	0.160	0.185	0.213	
Residents							0.031 (0.002)
λ		0.628			0.628		
N							71,994
R^2							0.199
State-Phase FEs	Yes	Yes	Yes	Yes	Yes	Yes	No
State-Rural FEs	No	No	No	No	No	No	Yes

The first six columns show estimates of the parameters in equation (1.5), separately by the two phases, early and late. Estimation is based on moment conditions formed using area-code level data for the 2013-2018 period (4,113 observations), and municipality-level data for 2019 (36,290 observations). See section 1.4 for details on estimation. The last column shows OLS estimates of equation 1.7. These estimates are based on survey data on consumers' expenses on mobile telecommunications services and demographic characteristics.

Table 1.6 displays estimates of the dynamic parameters: entry costs, technology upgrade costs, the cost of non-compliance with the regulation, the standard deviation of the cost shocks, and the unobservable profitability parameters, ψ. The costs associated with the introduction of 3G are found to be essentially constant over time. In contrast, the costs of introducing 4G decrease sharply, driven by the coefficient on the municipality's area.[31] Lastly, the fine is found to be very substantial: 6.89 million BRL, which is just over 40% of the median entry cost.

[31] In fact, the cost of upgrading to 4G in the later period is found to be negative.

Table 1.6: Dynamic Parameter Estimates

Parameter	Estimate	Parameter	Estimate
σ	2.461	φ	6.896
$\theta_{e,0}$	19.418	$\theta_{e,Area}$	-0.432
$\theta^E_{3G,0}$	7.363	$\theta^E_{3G,Area}$	0.721
$\theta^L_{3G,0}$	7.634	$\theta^L_{3G,Area}$	0.750
$\theta^E_{4G,0}$	-13.896	$\theta^E_{4G,Area}$	3.555
$\theta^L_{4G,0}$	-11.569	$\theta^L_{4G,Area}$	0.865
ψ_1	-0.199	ψ_2	-0.051
ψ_3	-0.013	ψ_4	0.014
ψ_5	0.067		

σ is the standard deviation of the cost shocks. φ is the cost of failing to comply with the regulation. $\theta_{e,0}$ is the entry cost intercept. $\theta_{e,Area}$ is the coefficient on the logarithm of area in the entry cost function. The remaining parameters are associated with installing 3G and 4G technology. The subscripts 3G and 4G indicate the technology. The subscripts $0, Area$ indicate the intercept and the area term, respectively. The superscripts E, L indicate the two periods, early and late. For example, $\theta^L_{4G,0}$ is the intercept of the cost of introducing 4G technology in the later period. The parameters $\psi_1, \ldots, \psi_5$ are the unobservable profit terms; see equation 1.3 and the discussion therein.

1.6 Counterfactual Analysis

The counterfactual exercises I conduct in this section directly address the questions posed in the beginning of the paper. In subsection 1.6.1, I use the model to analyze the effect of coverage requirements on the time to introduction of 3G technology. I also use the model to quantify the cost that the regulation imposes on firms and to decompose the total effect of coverage requirements into a direct effect on the regulated firm and indirect equilibrium effects. In subsection 1.6.2, I use the model to evaluate alternative regulations. Specifically, I consider policies that subsidize the first firm to introduce 3G technology, as well as an intervention that uses coverage requirements as insurance, in the sense that the regulated firm only incurs noncompliance costs in case no firm introduces 3G technology.

1.6.1 The Effect of Coverage Requirements

To quantify the effect of coverage requirements on the time to introduction of 3G technology and firms' ex-ante expected profits, I use the estimated model to simulate data under two alternative regulatory regimes. First, I solve the game and simulate data under the estimated fine $\hat{\varphi}$. Second, I solve the game and simulate data setting $\varphi = 0$, i.e., with no regulation. I simulate 250 paths of play for each municipality under each of these two regulatory regimes.

First, I compute the fraction of the 250 simulations in which some firm introduced 3G technology by December 2019. Figure 1.4 shows the distribution of those probabilities across municipalities. The figure shows that 66.45% of the municipalities in the sample would have had access to 3G technology with at least 75% probability. For 88.75% of the municipalities, the probability of having 3G access by December 2019 is at least 50%. This suggests that for most municipalities, market forces would most likely than not be sufficient to guarantee provision of 3G service. Figure 1.4 also shows that for 11.25% of municipalities, the probability of having 3G service by December 2019 is less than 50%. In these municipalities, market forces are insufficient to guarantee service provision. All the municipalities that have no service in the beginning of the data are in this group.

The results above may suggest that the regulation has limited effect, given that the probability of having 3G service by December 2019 is high for most municipalities. However, the regulation turns out to have non-negligible effects on the time to introduction of 3G technology. I use the models with and without regulation to simulate data until 2023. For each municipality and regulatory regime, I calculate the average number of years before the introduction of 3G techology or better.[32] Figure 1.5 shows

[32] In those instances in which 3G is not introduced by the end of the simulated data, I set the time to 3G introduction equal to the length of the simulated sample. This implies that the numbers I present on the effect of the regulation are, in some cases, a lower bound.

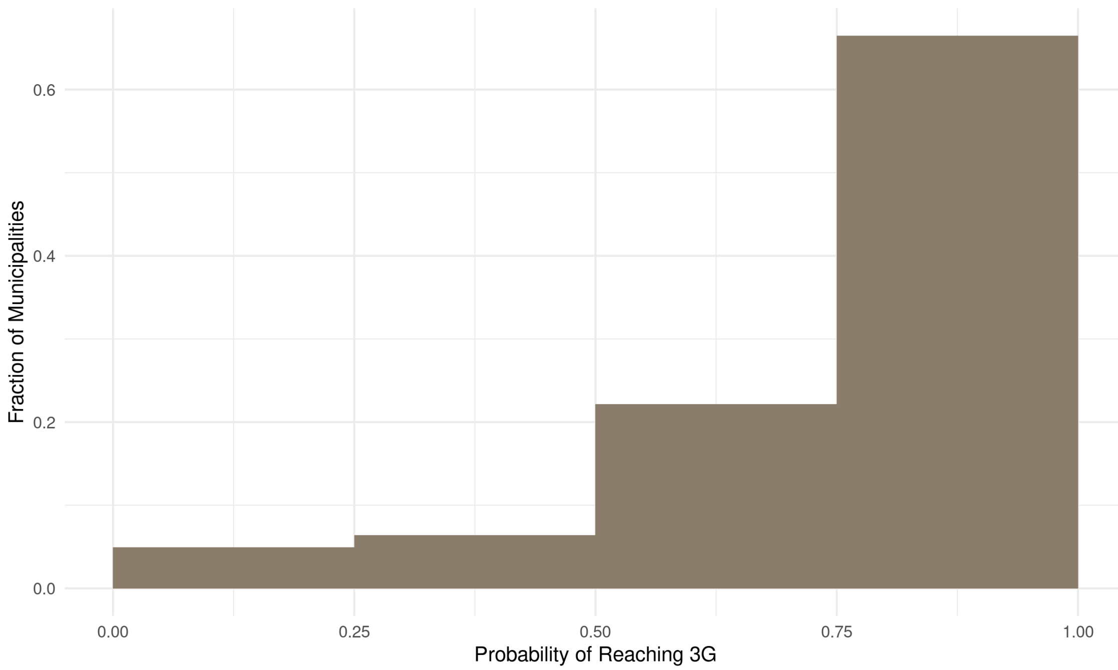

Figure 1.4: Distribution of the Probability of Reaching 3G by December 2019 Without Regulation.

the resulting distributions. In the figure, the label “Status quo” refers to setting the fine to $\hat{\varphi}$. The label “No regulation” corresponds to setting the fine to 0. As can be seen from the figure, the regulation reduces the average time to 3G introduction significantly – by 1.15 years, on average. The regulation also considerably reduces the dispersion in the time to introduction of 3G, mostly by eliminating a long right tail present in the absence of regulation.

Figure 1.6 shows the same information in a different way. For each municipality, I compute the accelleration in the introduction of 3G due to the regulation. Figure 1.6 plots the resulting distribution across municipalities. The effects are concentrated between 0 and 2 years, though there is a long right tail, consisting of the most vulnerable markets. For 4 municipalities, the regulation delays the introduction of 3G, though those effects are quantitatively small. In those cases,the equilibrium

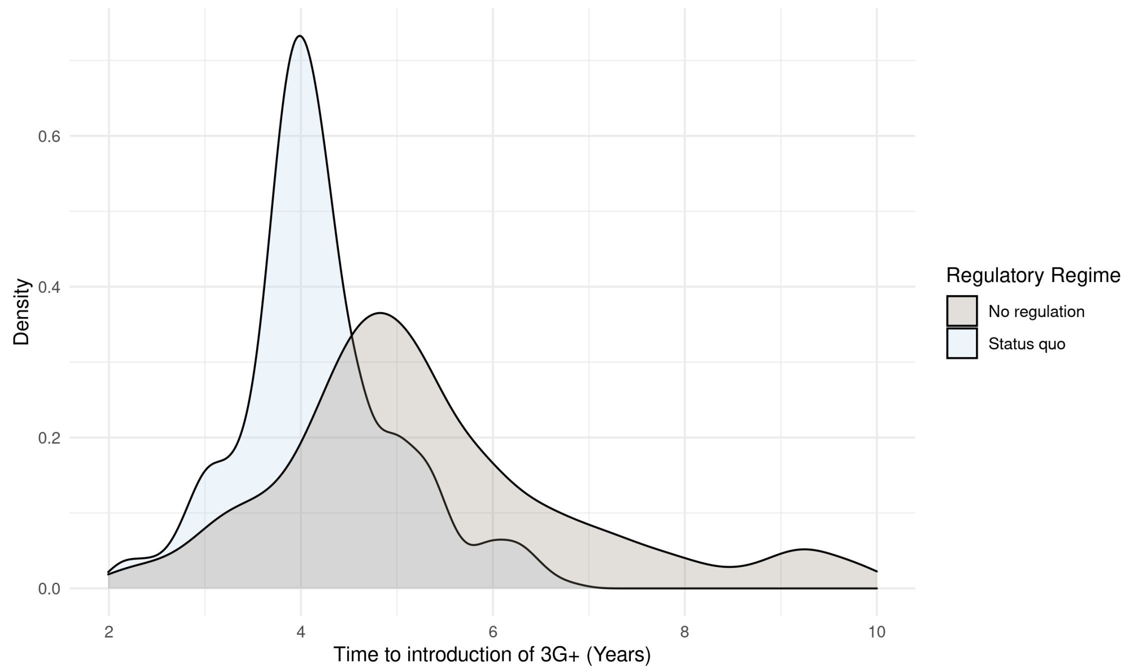

Figure 1.5: Distribution of the time to introduction of 3G technology or better under alternative regulatory regimes.

effects dominate the direct effect of the regulation.

To further understand the determinants of the effects of the regulation, I investigate how the time to 3G introduction in the absence of regulation and the acceleration afforded by coverage requirements relate to observable market characteristics. Specifically, I project the time to introduction of 3G with no regulation and the acceleration induced by coverage requirements onto observable market characteristics and variables that capture the initial market structure. I restrict attention to the municipalities that did not have 3G in the beginning of the sample. Table 1.7 reports the results.

The dependent variable in column 1 of table 1.7 is the time to 3G introduction without regulation, measured in years, and the explanatory variables are a municipality's GDP per capita, population, and area, as well as the number of firms in the

Table 1.7: Explaining Time to Adoption and the Effect of Regulation

	Dependent variable:			
	Time to 3G	Speedup	No. Entrants	Regulation Cost
	(1)	(2)	(3)	(4)
Log GDP	-0.324^{***}	-0.221^{***}	0.031^{***}	
	(0.052)	(0.027)	(0.007)	
Log Population	-0.363^{***}	-0.130^{***}	0.084^{***}	
	(0.070)	(0.037)	(0.010)	
Log Area	0.415^{***}	0.253^{***}	-0.049^{***}	
	(0.025)	(0.013)	(0.004)	
No. Firms t = 0	-1.630^{***}	-0.319^{***}		
	(0.082)	(0.051)		
Regulated Active t = 0		0.293^{***}	-0.884^{***}	
		(0.032)	(0.009)	
Regulated				2.889^{***}
				(0.050)
Active				0.127^{***}
				(0.017)
Regulated * Active				-2.732^{***}
				(0.041)
Constant				0.007^{***}
				(0.001)
Group FEs	Yes	Yes	Yes	No
Observations	689	665	665	3,008
R^2	0.797	0.747	0.948	0.863
Adjusted R^2	0.795	0.744	0.947	0.863
Note:				*p<0.1; **p<0.05; ***p<0.01

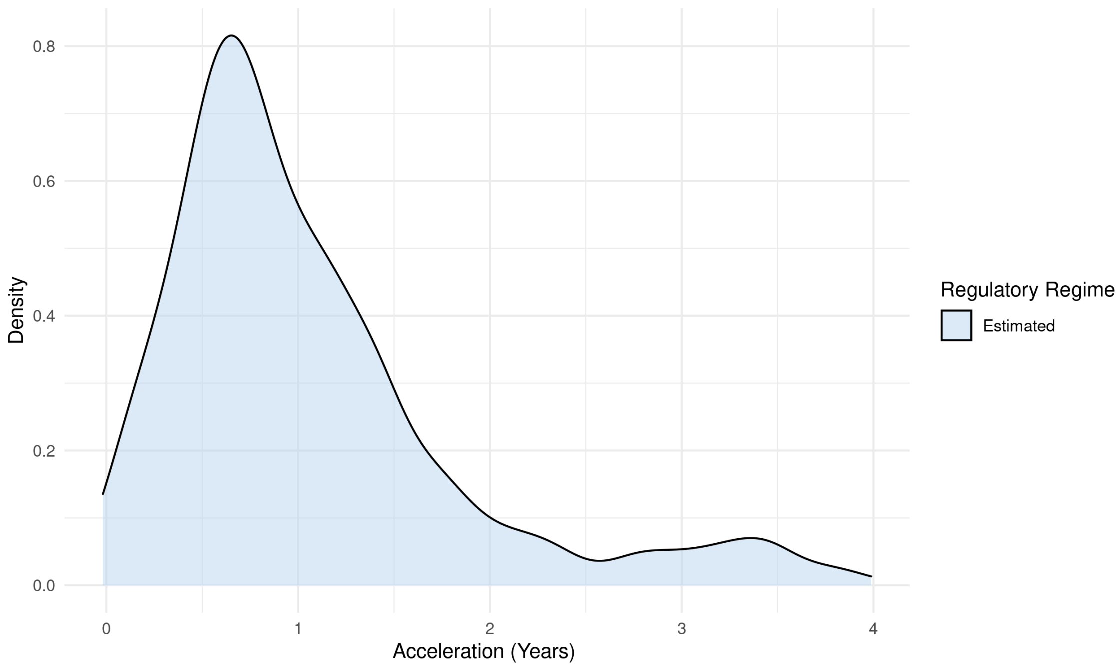

Figure 1.6: How much faster is the introduction of 3G+ under regulation?

beginning of the data.[33] The results show that the time to 3G introduction without regulation is decreasing in a municipality's GDP per capita and in its population, and it is increasing in a municipality's area. Moreover, the time to 3G introduction is decreasing in the number of firms in the market in $t = 0$. These results are all intuitive: firms are more likely to enter and upgrade their technologies in richer, more populous, and smaller markets; since incumbents have a lower cost of introducing 3G than potential entrants, a larger initial number of firms leads to faster 3G introduction.

The second column in table 1.7 models the acceleration in the introduction of 3G generated by coverage requirements, measured in years, as a function of the same variables included in column 1, and additionally a dummy for whether the regulated firm was active in the market in the beginning of the data.[34] The coefficients on the

[33] I take averages over time of these municipality characteristics. The sample is restricted to those markets that do not have 3G service in the beginning of the data.

[34] To aid in the interpretability of the coefficient on the dummy, this regression and the one in the

market characteristics and the number of firms show the same pattern as column 1, i.e., markets where, in the absence of regulation, 3G would be introduced faster also experience a smaller acceleration, as one might expect. Lastly, the estimates imply that regulating an incumbent leads to a larger acceleration than regulating a potential entrant, of just under four months.

The third column in table 1.7 reports a regression of the average number of entrants in a market, computed from simulations at the end of 2022, on market characteristics and whether the regulated firm was active in $t = 0$. Regulating an incumbent instead of a potential entrant reduces the average number of entrants by just under 0.9. The sign is expected, as a regulated potential entrant has to enter the market. The coefficient is less than one in absolute value because regulating an incumbent implies that there is one more unregulated potential entrant, and thus more entry that is not due to the regulation. Moreover, unregulated potential entrants may be more likely to enter the market when an incumbent firm is regulated, because they expect less future competition. I will show below, however, that for most markets the magnitude of this mechanism is not of first order importance.

Next, I use the model to calculate the cost that the regulation imposes on firms.[35] Solving the dynamic game under the estimated fine and under no regulation, I obtain, for each municipality, firms' ex-ante expected profits under those two regimes. The cost of the regulation is the aggregate difference in firms' ex-ante expected in profits in the no-regulation and the status-quo regimes:

$$\text{Regulation Cost} = \sum_{m} \sum_{f} \left(V^{m}_{\varphi=0}(s_{f0}, s_{-f0}, t=0) - V^{m}_{\varphi=\hat{\varphi}}(r_f, s_{f0}, s_{-f0}, t=0) \right)$$

---

third column further restrict attention to those municipalities that had at least one active firm in the beginning of the data

[35] Note that in the real world part of this cost is borne by the government, via reduced revenue in spectrum auctions.

where $V^m_\varphi(\omega)$ is the firm's ex-ante expected profit in municipality m and state ω when the fine is set to φ.[36] For the set of municipalities used in estimation, I calculate that the cost of the regulation amounts to 2.11 billion 2010 BRL, or 1.2 billion 2010 USD.[37] This amounts to 24.14% of firms' aggregate ex-ante expected profits with no regulation.

To understand the sources of these costs, the last column in table 1.7 reports estimates of a regression of the municipality-firm-specific regulation cost, $V^m_{\varphi=0}(s_{f0}, s_{-f0}, t = 0) - V^m_{\varphi=\hat{\varphi}}(r_f, s_{f0}, s_{-f0}, t = 0)$, onto a dummy for whether or not the firm is regulated, a dummy for whether or not the firm was active in the market in $t = 0$, and their interaction. The estimates show that for unregulated potential entrants (i.e., Regulated = 0 and Active = 0), the cost of the regulation is essentially zero. It is slightly positive because the regulation leads to a more competitive market when these potential entrants do enter, thus reducing their profits. As discussed in more detail below, that effect is small, which explains the small cost imposed on these firms. The cost for unregulated active firms is larger, because these firms are directly affected by the extra competition brought about by coverage requirements. On average, these firms lose about 134,000 USD because of the regulation, which is equivalent to 3.63% of their ex-ante expected profit without regulation. That effect depends on the technology of the incumbent firm: unregulated firms with 2G technology lose about 96,000 USD (3.22% of their profits), whereas unregulated firms with 3G technology lose about 346,000 USD (4.51% of their profits).

The bulk of the regulation costs falls on the regulated firms. The cost imposed on regulated firms with 2G technology is, on average, 291,000 USD, or 9.74% of their

[36] Note that in the first term, $V^m_{\varphi=0}(s_{f0}, s_{-f0}, t = 0)$, I do not include the regulation indicator r_f as an argument because there is no regulation in that case; r_f does appear as an argument in the second term.

[37] This conversion uses the average exchange rate in 2010.

profits under no regulation. The cost imposed on regulated firms that are not active in the market is very substantial: it is equal to 2.90 million USD, which is almost 13 times their ex-ante expected profits of (on average) 223,000 USD. This cost comes from the fact that these firms are forced to enter the market when they might have chosen not to do so. Overall, 84.62% of the costs imposed on firms come from those imposed on regulated potential entrants; 9.71% come from regulated incumbents; and the remaining 5.67% come from unregulated firms, i.e., they amount to costs stemming from competition effects of the regulation.

The fact that the regulation costs imposed on incumbents is substantially smaller than that imposed on potential entrants, combined with the fact that regulating incumbents leads to a faster introduction of 3G (see column 2 of table 1.7), may suggest that coverage requirements should be imposed on active firms. In practice, the extent to which this policy can be pursued, however, is limited by concerns of competitive neutrality. Such a policy would also provide poor incentives to firms, as entering new markets would make a firm more likely to be regulated in the future. Furthermore, this policy has an opportunity cost that is illustrated by column 3 of table 1.7: imposing the coverage requirement on an incumbent leads to less competition in the market than imposing the requirement on a potential entrant.[38] For the imposition of coverage requirements on potential entrants to be better, in aggregate welfare terms, than imposing those requirements on incumbents, the added competition (which is of 0.88 firms on average, according to table 1.7) has to generate an increase in consumer surplus of 8.60 BRL per month. This is equal to 50.65% of the average predicted expenditure for this set of municipalities.[39]

[38] Column 1 of table 1.7 suggests a second cost. Imposing the coverage requirement on a potential entrant may also accelerate the introduction of subsequent technologies. Below I investigate the effects of alternative coverage requirements on the adoption of 4G (which hasn't been directly regulated for this set of municipalities).

[39] This number is obtained by dividing the added cost from imposing the requirement on a potential

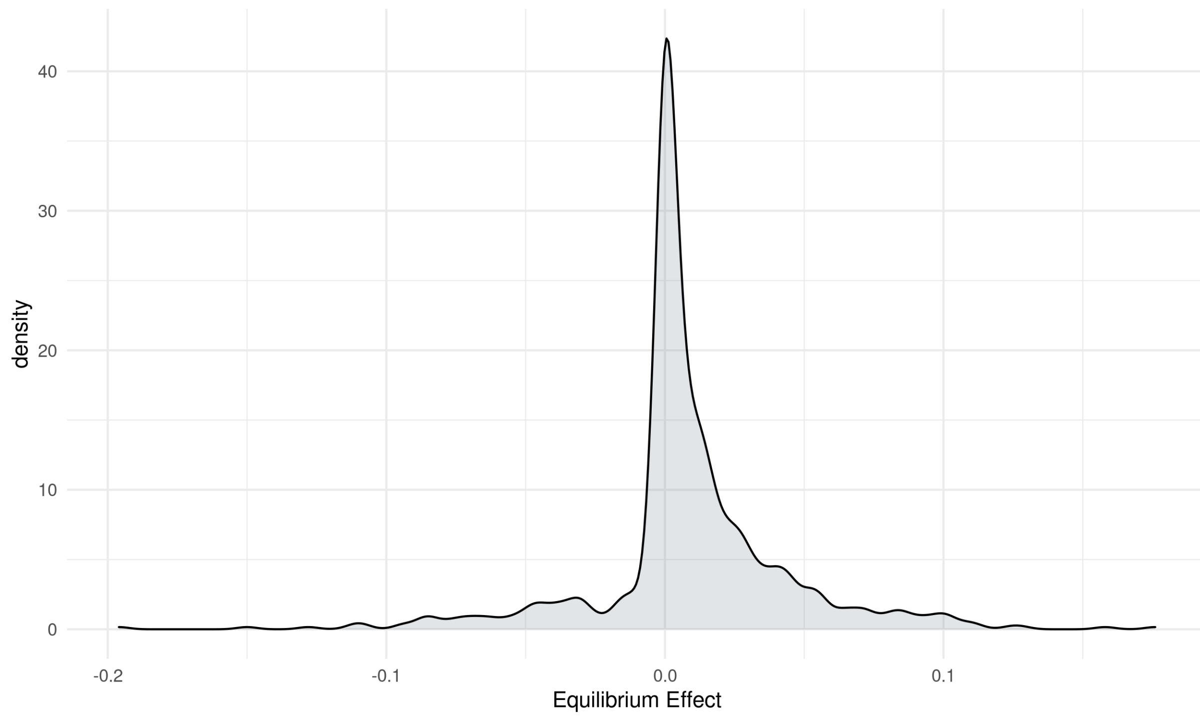

Figure 1.7: Equilibrium Effects

I close this subsection on the effects of coverage requirements by decomposing these effects into direct effects on the regulated firm and indirect equilibrium effects. To do so, I proceed in three steps. First, I solve the game and simulate data in the absence of regulation. I then solve for the regulated firm's optimal policy given the estimated fine and holding the policy functions of the unregulated firms fixed at their equilibrium policies without regulation. Next, I solve for the Markov Perfect Equilibrium under regulation. The difference between the time to adoption under the equilibrium policies with regulation and the time to adoption when only the regulated firm responds to the regulation gives the desired equilibrium effects.

Figure 1.7 shows the distribution, across municipalities, of the equilibrium effects.

entrant (assuming a single active firm in the market and setting the costs on inactive firms to zero) by the average population in the subsample of municipalities that don't have 3G in the beginning of the sample, which is 4,689, using the discount factor used in the model to arrive at a monthly gain in consumer surplus.

Most of the values are positive: the equilibrium adjustment leads to a longer time to introduction of 3G, relative to the case when only the regulated firm adjusts its behavior to the policy. This reflects the reduced incentives to enter and upgrade faced by unregulated firms, resulting from the increased future competition generated by the regulation. Quantitatively, however, the equilibrium effects are very small. The total effects of the policy are therefore almost entirely explained by the direct effects on the regulated firm.

1.6.2 Alternative Regulatory Interventions

The final question posed in the beginning of this paper was whether we can design more effective regulation than coverage requirements. As before, I am mostly concerned with two dimensions of a policy's effect: to what extent it accelerates the introduction of the new technologies, and the cost of adoption of these new technologies. I will also highlight the effect of different policies on market structure. In this subsection, I evaluate two alternative forms of intervention: using coverage requirements as "insurance" against lack of service, and subsidizing the first firm to introduce 3G.

Coverage Requirements as Insurance

The regulation currently in place consists of tasking one firm with introducing 3G technology by a given date. If that firm fails to do so, it incurs a cost of non-compliance. An alternative implementation of coverage requirements would be to impose costs on the regulated firm only if no firm provides 3G by the regulation deadline. This implementation would achieve introduction of 3G by the regulation deadline (assuming sufficiently strong enforcement), and it would also have benefits relative to the current implementation of the regulation. First, it would reduce the cost imposed on the regulated firm, because if some other firm chooses to introduce 3G, the regulated firm would not be subject to the requirement anymore. Second, this implementation of coverage requirements would do away with negative entry deterrence effects. However, given the results above showing that the equilibrium effects of the regulation are quantitatively small, this benefit should also be small.

Results and discussion to be added.

Subsidizing the Introduction of 3G

The large estimated cost of non-compliance and the counterfactual results above show that coverage requirements provide strong incentives for the regulated firm to introduce 3G. These strong incentives ensure service provision. However, they come at the cost of forcing a firm to enter a market or upgrade its technology when it might not have done so in the absence of regulation. The analysis above established that these costs are substantial, especially when the regulated firm is not active in the market.

A policy that treats firms symmetrically, instead of focusing on a single firm, may save on these costs. The intuition is simple. By providing the same incentive to all firms, the firm that will eventually choose to introduce the new technology will tend to be the most cost-efficient one.

Motivated by this reasoning, in this section I evalute a regulation that subsidizes the first firm to introduce 3G technology or better. I denote the subsidy by β. If more than one firm introduces the new technology, those firms split the subsidy equally. Therefore, I add the following term to firms' flow profits for each state of the game

$$\beta \times \underbrace{\mathbf{1}\left\{\left(\max_{f'} s_{f'}\right) < 3 \leq a_f\right\}}_{\text{Subsidy is paid}} \times \underbrace{\sum_{n=0}^{3} \mathbb{P}\left(\left(\sum_{f' \neq f} \mathbf{1}\{a_{f'} \geq 3\}\right) = n\right) \times \frac{1}{1+n}}_{\text{Expected fraction of the subsidy}}$$

where the probabilities in this expression are derived from the ensuing equilibrium behavior.

I experiment with two subsidy designs. I start with a budget given by

$$\text{Budget} = \sum_m \sum_f \left(V^m_{\varphi=0}(s_{f0}, s_{-f0}, t=0) - V^m_{\varphi=\hat{\varphi}}(r_f, s_{f0}, s_{-f0}, t=0) \right) \tag{1.15}$$

This amount is simply the aggregate cost of the regulation. Note that firms would be willing to pay this amount to move from the status quo world to a world with a subsidy. In that sense, the subsidies considered below are self-financed.

I start by simply splitting the budget in equation (1.15) equally across municipalities. Figure 1.8 shows the resulting acceleration in the introduction of 3G technology obtained under coverage requirements (labeled "status quo" in the figure) and the subsidy. The average effect is very similar; the subsidy accelerates the introduction of 3G by 1.07 years on average, relative to 1.15 years under coverage requirements. The subsidy generates larger accelerations for 63.71% of the municipalities. As figure 1.8 shows, relative to coverage requirements, the subsidy eliminates some small effects, but also loses some large ones. The large effects lost by the subsidy come precisely from those municipalities that would experience relatively late introduction of 3G in the absence of regulation. Consider, for example, those municipalities where coverage requirements generate an acceleration in the introduction of 3G of one year and a half or more. The average time to introduction of 3G without regulation in these municipalities is almost three years more than in the remaining municipalities. This is a set of municipalities where the introduction of 3G is relatively unprofitable, and the homogeneous subsidy provides less incentives for 3G introduction in these municipalities than coverage requirements do. For this set of municipalities, the subsidy leads to 3G introduction 1.2 years later than coverage requirements, on average.

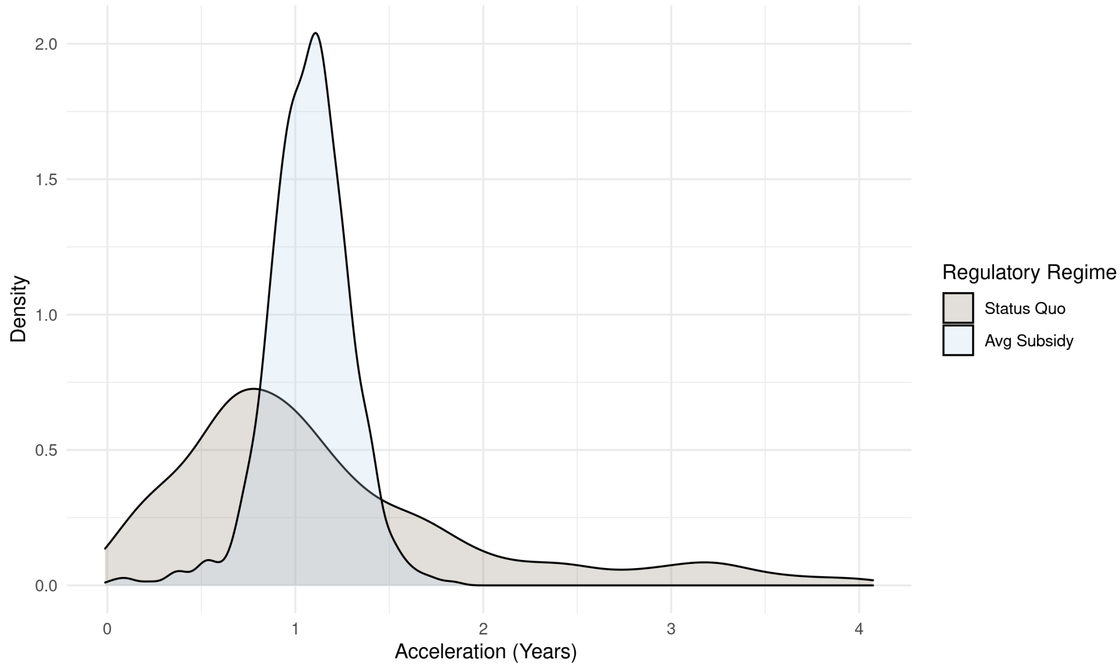

Figure 1.8: Acceleration of 3G Introduction Under Coverage Requirements and Subsidy

The municipalities where coverage requirements generate small accelerations (less than 6 months) are relatively competitive municipalities. The average number of firms in $t = 0$ in those municipalities is 1.54, relative to 0.96 in the remaining municipalities. The introduction of 3G in these municipalities in the absence of regulation is relatively fast: just under 3.5 years, on average, compared to just over under 5.5 years in the other municipalities. In these markets, the effect of the subsidy is very close to the mean effect, so that these markets are moved from the left tail of the "Status Quo" distribution if figure 1.8 to the middle of the subsidy distribution. In summary, relative to coverage requirements, a flat subsidy increases the acceleration of 3G introduction in some localities where there seems to be little need for regulation, and has smaller effects in some municipalities where regulation seems to be particularly important.

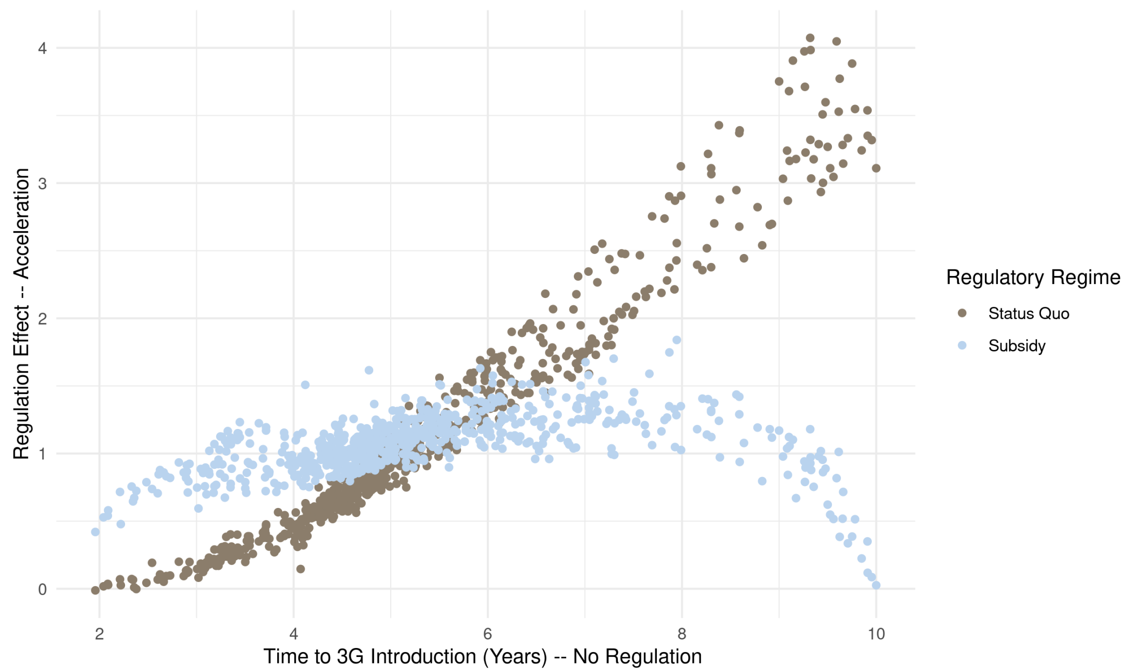

Figure 1.9: Targeting Properties of Coverage Requirements and a Flat Subsidy

This point is shown clearly in figure 1.9. The figure shows a scatterplot of the time to introduction of 3G technology in the absence of regulation against the effects of coverage requirements and the flat subsidy. Each dot in the figure is one municipality. For the case of coverage requirements, we see a positive correlation: the regulation has stronger effects in those markets where, in the absence of intervention, it would take longest for 3G to be introduced. The flat subsidy does not display the same correlation. In fact, its effects are smallest for the most vulnerable municipalities.

In light of these results, I consider a subsidization policy that allocates a larger share of the budget towards the most vulnerable municipalities. Specifically, let τ_m be the time for 3G introduction in municipality m in the absence of regulation and let f be a positive and increasing real function. Allocate to municipality m the fraction $f(\tau_m)/\sum_{m'} f(\tau'_m)$ of the budget specified in equation (1.15). The more convex f

is, the stronger the targeting towards the most vulnerable municipalities. For the results below, I set $f(\tau) = \tau^{3/2}$.[40] Figure 1.10 shows the results. This subsidy leads to a larger acceleration in the roll-out of 3G: 1.27 years relative to 1.15 years under coverage requirements. As shown in the figure, the municipality-specific subsidy restores the desired positive correlation between the effect of the regulation and the time to 3G introduction in the absence of regulation. In fact, this subsidy leads to larger accelerations in the roll-out of 3G in the most vulnerable municipalities than do coverage requirements. This comes at the expense of slightly smaller effects in those municipalities that even in the absence of regulation obtain access to 3G technology relatively quickly. The optimal way to navigate this trade-off (e.g., the optimal choice of exponent in $f(\tau)$) depends on the relative changes in consumer surplus in those two groups of municipalities, which can't be quantified with the limited data available in this study.

Finally, firms substantially benefit from the municipality-specific subsidy relative to coverage requirements.[41] Firms' ex-ante aggregate expected profits grow by 659 million USD, after accounting for their financing of the subsidy (as per equation (1.15)); this amounts to 28% of firms' aggregate profits without regulation. These gains essentially come from reallocating the introduction of the new technology from inactive and regulated firms, who have to pay entry costs to enter the market, to incumbents, who only pay technology installation costs.

This reallocation leads to a more cost-efficient technology roll-out, but at the expense of reduced competition in the market. The subsidy leads to entry of 0.51

[40] This subsidy design relies on τ_m, and one may thus be concerned that its informational requirements are substantial. However, note that the results in table 1.7 show that a substantial portion of the variation in τ_m is explained by observables. Therefore, it might be possible to design a subsidy with similar properties that relies only on data that is available to regulators.

[41] Similar results hold for the flat subsidy.

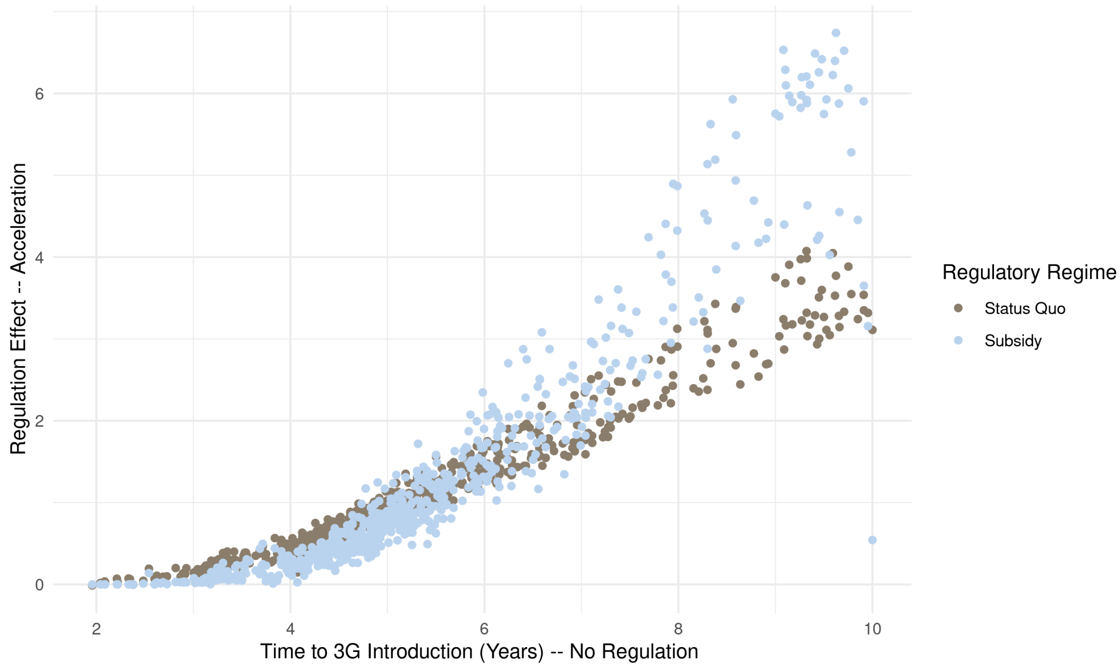

Figure 1.10: Targeting Properties of Coverage Requirements and a Municipality-Specific Subsidy

firms, on average, by the end of 2022. In contrast, coverage requirements lead to entry of 0.84 firms. Moreover, this difference is entirely driven by those markets where the regulated firm is a potential entrant, which are the source of the cost savings discussed above. The average number of entrants in these markets, under coverage requirements, is 1.32, whereas it is equal to 0.57 under the subsidy. In the remaining markets, those where the regulated firm is an incumbent, coverage requirements result in entry of 0.42 firms, on average; the subsidy results in entry of 0.46 firms. In those markets where a potential entrant is subject to the coverage requirement, the coverage requirement would be more desirable than the subsidy if the added competition from one additional firm generated additional consumer surplus of 7.35 BRL; that amounts to 44.69% of consumers' average expenditures in those markets.

1.7 Conclusion

Concerns regarding lack of service provision are widespread and so is regulatory intervention. This paper studies the effect of coverage requirements, a common form of regulation in the mobile telecommunications industry, on the speed of roll-out of new technologies, market structure, and firms' profits. To do so, I use new mobile technology availability data from Brazil to estimate a dynamic model of entry and technology upgrade under regulation.

Counterfactual simulations show that in the absence of regulation, third generation mobile telecommunications technology would have been introduced just over one year later, on average. This faster introduction comes at a high cost: firms' ex-ante expected profits are 24% lower under the existing regulation than they would have been in its absence. I also use the model to evaluate alternative policies. In particular, I find that a policy that subsidizes the first firm to introduce 3G technology, by an amount that the firms themselves would be willing to finance, achieves a slightly larger acceleration of the introduction of 3G and leads to more cost-efficient patterns of roll-out, likely increasing aggregate welfare. These findings have immediate implications for the design of regulation in mobile telecommunications markets, and potentially to other markets where universal service is also a concern.

Some interesting and related questions are not addressed in this paper. First, though my results are informative for the design of regulation, data limitations preclude me from conducting a complete welfare analysis. It would be interesting to combine data such as the one used in this paper with detailed price and quantity data to compare the gains in consumer surplus from having earlier access to new technologies and the regulatory costs imposed on firms. Second, my analysis abstracted away from spatial correlation in firms' costs. It would also be interesting, though challending, to study the introduction of new mobile telecommunications technologies while modeling geographic cost interdependencies. These interesting and challenging topics are left for future research.

1.A Regulation and Delay in the Fudenberg-Tirole Model

1.A.1 The Model

There are two firms. Firm 1 is an incumbent and firm 2 a potential entrant. Time is continuous and the discount rate is r. Firm 1 initially operates as a monopolist with constant marginal cost $\bar{c}$. At any point in time $t \geq 0$, firms can adopt a technology with constant marginal cost $\underline{c}$. Adopting this technology at time t costs $C(t)$, where $C(t) > 0$, $C'(t) < 0$ and $C''(t) > 0$, for all $t \geq 0$.

Let $p^m(c)$ and $\pi^m(c)$ be, respectively, the monopoly price and profit when marginal cost is c. I focus on the case in which the innovation is *non-drastic*, i.e., $p^m(\underline{c}) \geq \bar{c}$. If both firms are in the market, they compete à la Bertrand. Let $\pi^d(c, c')$ be a firm's profit when its cost is c and its competitor's cost is c'. Under the assumption of a non-drastic innovation and Bertrand competition, π^d satisfies

$$\pi^d(\underline{c}, \bar{c}) = (\bar{c} - \underline{c})D(\bar{c}), \quad \pi^d(\bar{c}, \underline{c}) = 0 \quad \text{and} \quad \pi^d(c, c) = 0 \quad \forall c$$

Firms' strategies specify their decisions to adopt or not the new technology as a function of t and their competitor's technology.[42] Note that due to the Bertrand assumption, a firm will never adopt the new technology after its competitor has

[42] The discussion here is somewhat informal. Fudenberg and Tirole 1985 provide a careful description of appropriate strategies for this game. Their analysis is far from trivial.

adopted, as they would incur the positive adoption cost but their flow profits would stay at zero.

If the incumbent is first to adopt at date t_1, its overall profit is

$$L_1(t_1) = \int_0^{t_1} \pi^m(\bar{c})e^{-rt}dt + \int_{t_1}^{\infty} \pi^m(\underline{c})e^{-rt}dt - C(t_1)e^{-rt_1} \tag{1.16}$$

If the incumbent is preempted at date t_2, its present discounted profit is

$$F_1(t_2) = \int_0^{t_2} \pi^m(\bar{c})e^{-rt}dt \tag{1.17}$$

If the entrant is first to adopt at date t_2, its overall profit is

$$L_2(t_2) = \int_{t_2}^{\infty} \pi^d(\underline{c}, \bar{c})e^{-rt}dt - C(t_2)e^{-rt_2} \tag{1.18}$$

Finally, it the entrant is preempted at time t_1, its profit is given by $F_2(t_1) = 0$.

Figure 1.11 plots the functions L_1, F_1, L_2, F_2.[43] That figure is sufficient to determine the equilibrium outcome of the game.[44] Let t_2^* be defined by $F_2(t_2) = L_2(t_2)$. In Figure 1.11, $t_2^* \approx 5$. Firm 2 will not adopt before t_2^*, as it would prefer to be preempted by firm 1. Knowing this, firm 1 will wait to adopt, as $L_1(t_1)$ is increasing over $t_1 < t_2^*$. Now suppose firm 2 is first to adopt at some $t_2 > t_2^*$. Since $L_1(t_2 > F_1(t_2))$, firm 1 prefers to adopt at $t_2 - \varepsilon$. In equilibrium, firm 1 adopts at $t_1 = t_2^*$, and firm 2 never adopts.

[43] The specification is as follows. $D(p) = 2 - p$, $\bar{c} = 1$, $\underline{c} = 3/4$, $C(t) = \mathbf{1}\{t <= 10\}\left(\frac{t^2}{4} - 5 * t + 25\right) + 0.1$.

[44] But not the equilibrium itself, i.e., the strategy profile.

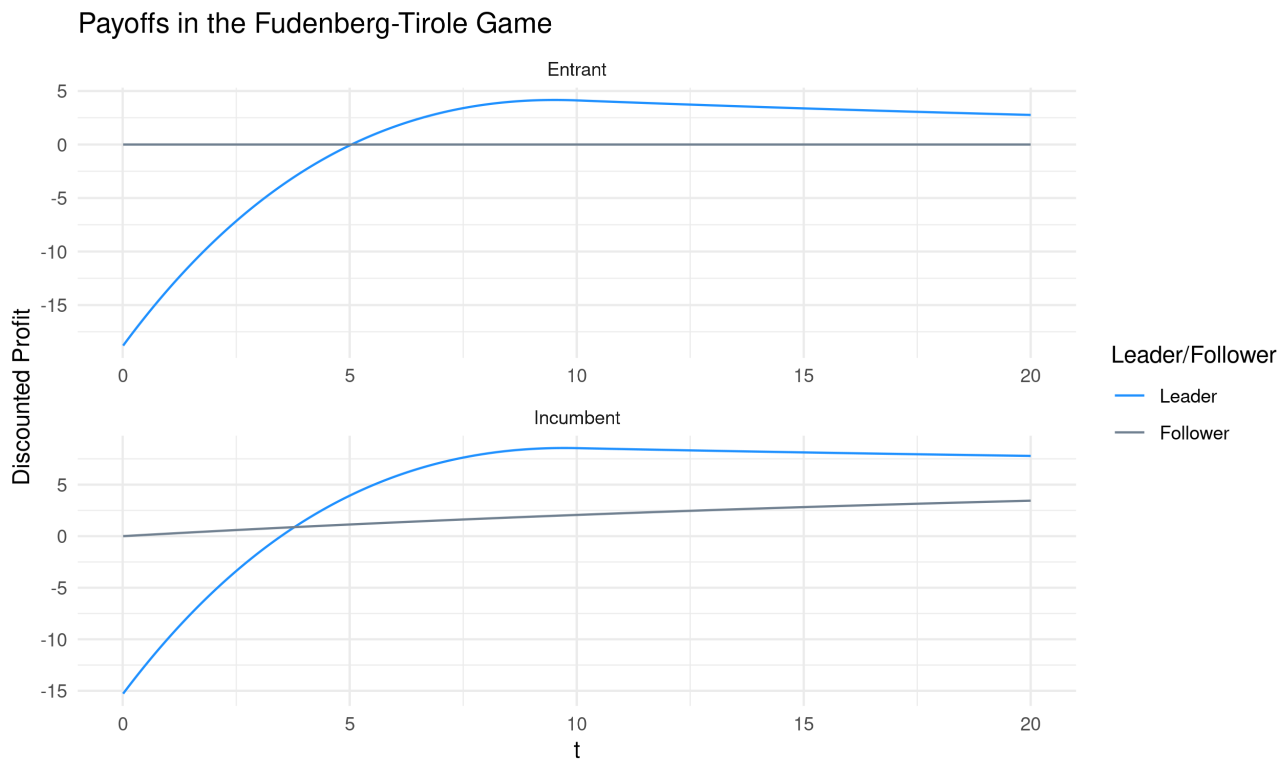

Figure 1.11: Payoffs in the Fudenberg-Tirole Model.

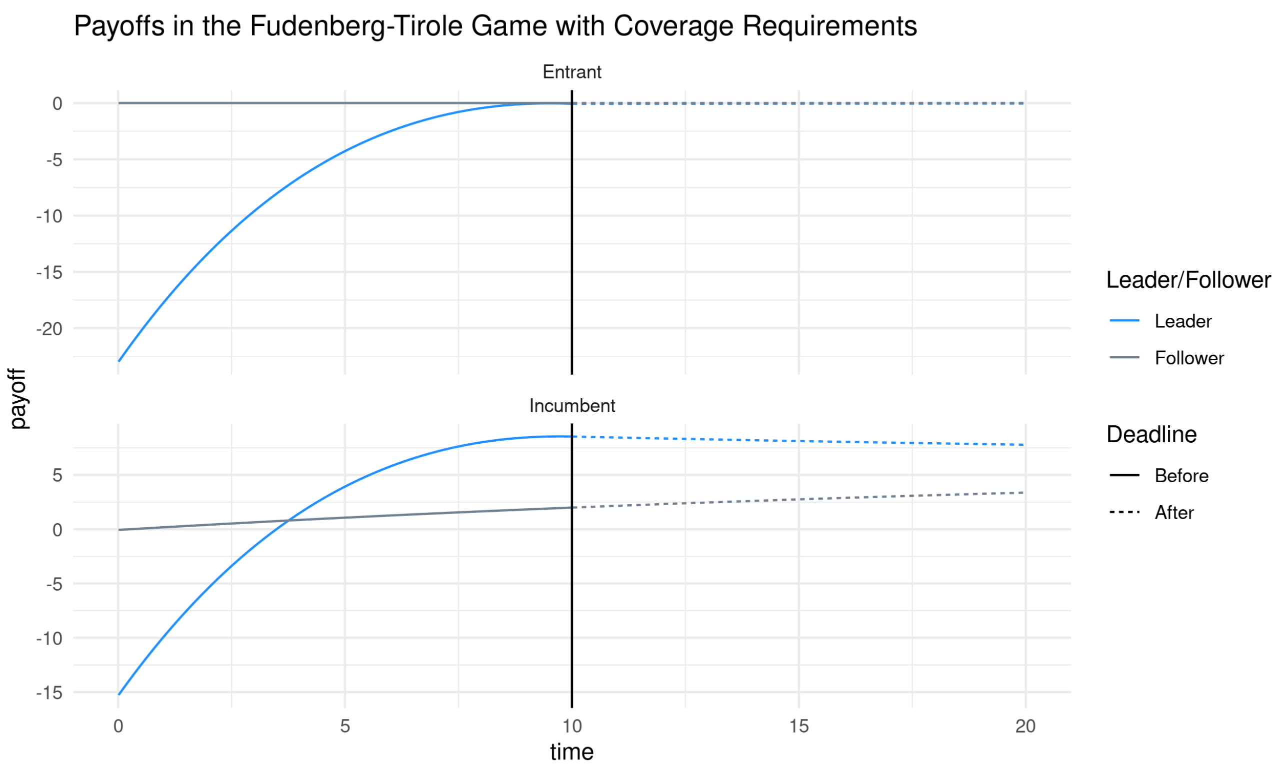

Figure 1.12: Payoffs in the Fudenberg-Tirole Model with Regulation.

1.A.2 Incorporating Regulation

Now suppose that the incumbent is regulated: is must adopt by some exogenously set deadline τ, lest it pay an exorbitant fine. The L_i and F_i functions are now defined (for $t_i \leq \tau$) as follows:

$$\begin{aligned} L_1(t_1) &= \int_0^{t_1} \pi^m(\bar{c})e^{-rt}dt + \int_{t_1}^{\infty} \pi^m(\underline{c})e^{-rt}dt - C(t_1)e^{-rt_1} \\ F_1(t_2) &= \int_0^{t_2} \pi^m(\bar{c})e^{-rt}dt - C(\tau)e^{-r\tau} \\ L_2(t_2) &= \int_{t_2}^{\tau} \pi^d(\underline{c}, \bar{c})e^{-rt}dt - C(t_2)e^{-rt_2} \\ F_2(t_1) &= 0 \end{aligned} \tag{1.19}$$

Figure 1.12 plots these payoffs for the same parametrization underlying Figure

1.11, and $\tau = 10$. As can be seen from the figure, the fact that the incumbent will adopt the technology at time τ, at the latest, eliminates all incentive for the entrant to adopt the new technology. With no need to preempt the entrant, the incumbent is free to delay its own adoption to its most preferred time, which in this example is $t_1^* \approx 9.7$. Therefore, the regulation delays the adoption of the new technology from $t \approx 5$ to $t \approx 9.7$. Of course, if $\tau < 5$, the regulation speeds up the adoption of the new technology.

1.B Descriptive Models – Alternative Specifications

This appendix reports alternative specifications of the descriptive models in table 1.4. In particular, table 1.8 reports models without group fixed effects, and table 1.9 reports models that include characteristics of firms' networks in neighboring states. Specifically, it includes dummies for whether or not the firm provides 2G, 3G, and 4G service in any neighboring municipality. Comparing table 1.8 and table 1.4 shows the importance of the group fixed effects. Without them in table 1.8, the competition coefficients are mostly small in absolute value and sometimes positive. That is in stark contrast with the results in table 1.4, where the competition coefficients are almost all negative and much larger in absolute value. This suggests that the group fixed effects capture important unobserved factors related to how desirable it is to provide service in a given market.

Now let me turn to table 1.9. The first thing to note is that service in neighboring municipalities is important. The estimated coefficients on 3G service and 4G service are sizeable and precisely estimated. Interestingly, the coefficients on 2G service in neighboring municipalities are negative. This is surprising because these coefficients

Table 1.8: Entry/Upgrade Models – Without group fixed effects

	Dependent variable:				
	Out 13-15	Out 16-18	Upgrade 2G 13-15	2G 16-18	3G
	(1)	(2)	(3)	(4)	(5)
Log GDP PC	0.389***	−0.001	0.241***	−0.194***	0.186***
	(0.064)	(0.081)	(0.047)	(0.051)	(0.032)
Log Pop.	0.761***	0.728***	0.851***	0.430***	−0.059*
	(0.066)	(0.086)	(0.051)	(0.058)	(0.035)
Log Area	−0.123***	−0.078**	−0.221***	−0.233***	0.018
	(0.031)	(0.040)	(0.025)	(0.027)	(0.019)
Regulated	1.712***	2.111***	2.312***	0.926***	−0.398***
	(0.110)	(0.126)	(0.076)	(0.104)	(0.040)
Regulated Competitor - Out	−0.662***	−1.167***	0.320**	−0.221	−0.137
	(0.173)	(0.284)	(0.150)	(0.162)	(0.132)
Regulated Competitor - 2G	−0.021	−0.157	−0.304**	−1.202***	−2.345***
	(0.115)	(0.192)	(0.120)	(0.314)	(0.235)
No. Competitors 2G	−0.044	−0.374***	−0.035	0.137***	−0.064**
	(0.069)	(0.097)	(0.036)	(0.049)	(0.027)
No. Competitors 3G	−0.269***	−1.239***	0.047	−0.001	0.190***
	(0.090)	(0.104)	(0.047)	(0.053)	(0.033)
No. Competitors 4G	0.212	−0.466***	−1.343*	−0.307***	0.411***
	(1.031)	(0.107)	(0.719)	(0.056)	(0.034)
Group FE	No	No	No	No	No
Observations	36,230	31,620	24,753	14,002	39,923

Note: *p<0.1; **p<0.05; ***p<0.01

are relative to not having service in the neighboring municipality. The next thing to observe is the effect of the network variables on the competition coefficients. These effects are mostly small, except perhaps for the number of competitors with 4G technology. Albeit small, the effects are always in the direction of increasing (in absolute value) the estimated competition coefficients. This may suggest that there are unobservable factors that are geographically correlated.[45] Finally, and most importantly for the analysis in this paper, note that the effect of the network variables on the regulation variables is very minor, if it exists at all. This suggests that the regulation variables (in particular, whether or not a firm is regulated) are not correlated with the surrounding network infrastructure.

Table 1.10 tests the hypothesis of no correlation between a firm's status as the regulated firm and that firm's infrastructure in neighboring markets. The unit of analysis for the models in table 1.10 is a firm-market pair, and only data from the June 2016 (the first period in the data) is used. The table reports estimation results of a logit model and a linear probability model (included for the sake of interpretability) where the dependent variable is a dummy that takes the value 1 if the firm is regulated, and 0 otherwise. The explanatory variables are a constant and a set of dummies. The variable "2G Service" is equal to 1 if the firm provides 2G service in that market; "3G service" is analogously defined. "2G Service Nb." is equal to 1 if the firm provides 2G service in some neighboring market, and "3G Service Nb." is defined similarly. The results show that, conditional on the technologies offered by a firm in the market, which are included in the structural model, its infrastructure in neighboring municipalities has a small effect on the probability that the firm is regulated. The point estimates are in fact negative. These results suggest that there is no

[45] Variables that are currently omitted and could potentially be included are variables related to the terrain.

Table 1.9: Entry/Upgrade Models – With Neighboring Network Info

	Dependent variable:				
			upgrade		
	Out 13-15	Out 16-18	2G 13-15	2G 16-18	3G
	(1)	(2)	(3)	(4)	(5)
Log GDP PC	1.772***	1.116***	0.693***	0.261***	0.323***
	(0.093)	(0.120)	(0.066)	(0.071)	(0.039)
Log Pop.	2.537***	2.151***	1.337***	1.081***	0.137***
	(0.106)	(0.151)	(0.072)	(0.085)	(0.047)
Log Area	−0.512***	−0.398***	−0.294***	−0.402***	−0.063***
	(0.038)	(0.051)	(0.027)	(0.031)	(0.020)
Regulated	1.716***	2.269***	2.191***	0.887***	−0.275***
	(0.110)	(0.130)	(0.077)	(0.110)	(0.042)
Regulated Competitor - Out	−0.720***	−0.997***	0.131	−0.364**	−0.100
	(0.173)	(0.285)	(0.152)	(0.168)	(0.136)
Regulated Competitor - 2G	0.099	0.023	−0.487***	−1.114***	−2.155***
	(0.114)	(0.195)	(0.121)	(0.319)	(0.236)
No. Competitors 2G	−1.445***	−1.153***	−0.448***	−0.339***	−0.154***
	(0.093)	(0.120)	(0.056)	(0.069)	(0.039)
No. Competitors 3G	−2.072***	−2.265***	−0.667***	−0.741***	−0.015
	(0.122)	(0.146)	(0.082)	(0.088)	(0.050)
No. Competitors 4G	−1.796*	−1.823***	−2.310***	−1.275***	−0.184***
	(1.036)	(0.158)	(0.734)	(0.093)	(0.051)
Nb. Service 2G	−0.398**	−1.135***	−0.174	−0.534	−0.047
	(0.157)	(0.171)	(0.230)	(0.328)	(0.204)
Nb. Service 3G	1.040***	1.523***	0.601***	0.654***	0.490***
	(0.097)	(0.180)	(0.063)	(0.108)	(0.161)
Nb. Service 4G	0.960***	0.495***	0.575***	1.200***	1.640***
	(0.179)	(0.097)	(0.133)	(0.062)	(0.043)
Group FE	Yes	Yes	Yes	Yes	Yes
Observations	36,230	31,620	24,753	14,002	39,923
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 1.10: Testing for Selection on Infrastructure in Neighboring Municipalities

	Dependent variable:	
	Regulated	
	Logit	LPM
	(1)	(2)
2G Service	1.727***	0.237***
	(0.058)	(0.008)
3G Service	0.883***	0.194***
	(0.058)	(0.010)
2G Service Nb.	−0.240*	−0.018
	(0.125)	(0.016)
3G Service Nb.	−0.345***	−0.047***
	(0.052)	(0.008)
Constant	−2.104***	0.116***
	(0.117)	(0.015)
Observations	13,204	13,204
R^2		0.139
Adjusted R^2		0.139
Note:	*p<0.1; **p<0.05; ***p<0.01	

cause for concern that the difference in behavior between regulated and unregulated firms, which identifies the fine parameter φ in the structural model, is driven not by the regulation itself but by omitted differences in firms' neighboring infrastructure. Therefore, despite the importance of neighboring infrastructure shown in table 1.9, I omit these variables from the structural model, as doing so would likely not bias the inference regarding the effects of regulation and would increase the computational burden by several orders of magnitude.

1.C Symmetry Restrictions

The symmetry assumption implies the following restrictions on value functions (and policy functions):

- $V_1(s_1, s_{-1}, t) = V_1(s_1, P(s_{-1}), t)$, for any permutation P.
- $V_0(s_1, s_r, s_-, t) = V_0(s_1, s_r, P(s_-), t)$ for any permutation P.
- If $s_r \geq 3$ and $\exists j \notin \{1, r\}$ s.t. $s_j \geq 3$, then $V_0(s_1, s_r, s_j, s_k, t) = V_0(s_1, s_j, s_r, s_k, t)$
- If $s_1, s_r \geq 3$, then $V_0(s_1, s_r, s_-) = V_1(s_1, P(s_r, s_-))$ for any P.

Add discussion on state space representation.

1.D Conditional Value Functions are Linear in Parameters

In this section I will simplify notation by letting ω denote a generic state of the form $\omega = (t, r, s_f, s_{-f})$. Flow payoffs, net of the idiosyncratic shock, are given by

$$\pi(\omega) - \varphi r \mathbf{1}\{s_f < 2, T < t\} - c(a, s_f)$$

This expression can be written as a linear function of parameters. To see this, first redefine actions a and firm technological states s_f to be vectors indicating the presence of each technology, ordered from 4G to 2G. For example, if a firm offers 3G and 2G, represent s_f as $s_f = (0, 1, 1)$. The deterministic part of costs can then be written as

$$\left[(a' - s'_f) \otimes (\mathbf{1}\{p(t) = E\}, \mathbf{1}\{p(t) = L\}) \otimes z'\right] \underbrace{(\theta'_{4,E}, \theta'_{3,E}, \theta'_{2,E} + \theta_e, \theta'_{4,L}, \theta'_{3,L}, \theta'_{2,L} + \theta_e)'}_{\theta}$$

Define

$$g(a, \omega, z) := \Big(\pi(\omega), \Big[(a - s_f) \otimes (\mathbf{1}\{p(t) = E\}, \mathbf{1}\{p(t) = L\}) \otimes z\Big], r\mathbf{1}\{s_f < 2, T < t\}\Big)$$

and

$$\Psi := (1, \theta', \varphi)'$$

Then we have

$$\pi(\omega) - \varphi r \mathbf{1}\{s_f < 2, T < t\} - c(a, s_f) = g(a, \omega, z)\Psi$$

The value function satisfies the Bellman equation

$$V(\omega, \varepsilon_f) = \max_{a \in A(s_f)} g(a, \omega, z)\Psi + \varepsilon_f(a) + \delta \sum_{\omega'} V(\omega') F_P(\omega' | \omega, a)$$

where F_P denotes the state transitions induced by the equilibrium conditional choice probabilities P and

$$V(\omega') := \int V(\omega, \varepsilon_f) dG(\varepsilon_f)$$

Denote the equilibrium policy by $\sigma^*(s, \varepsilon_f)$. Then (using σ^* as shorthand for $\sigma^*(s, \varepsilon_f)$)

$$V(\omega, \varepsilon_f) = g(\sigma^*, \omega, z)\Psi + \varepsilon_f(\sigma^*) + \delta \sum_{\omega'} V(\omega') F_P(\omega' | \omega, \sigma^*)$$

Integrating both sides of this equation yields

$$\begin{aligned} V(\omega) = & \left(\int g(\sigma^*, \omega, z) dG(\varepsilon_f) \right) \Psi \\ & + \int \varepsilon_f(\sigma^*) dG(\varepsilon_f) + \delta \sum_{\omega'} V(\omega') \int F_P(\omega' | \omega, \sigma^*) dG(\varepsilon_f) \end{aligned}$$

Let $\mathcal{C}(a,\omega)$ be the set of shocks $\varepsilon_f \in \mathbb{R}^{|A(s_f)|}$ such that $a = \sigma^*(\omega, \varepsilon_f)$. Then

$$\begin{aligned}
\int g(\sigma^*, \omega, z) dG(\varepsilon_f) &= \sum_{a \in A(s_f)} \int_{\mathcal{C}(a,\omega)} g(\sigma^*, \omega, z) dG(\varepsilon_f) \\
&= \sum_{a \in A(s_f)} g(a, \omega, z) \int_{\mathcal{C}(a,\omega)} dG(\varepsilon_f) \\
&= \sum_{a \in A(s_f)} g(a, \omega, z) P(a|\omega)
\end{aligned}$$

where here $P(a|\omega)$ are the equilibrium conditional choice probabilities.

Similarly,

$$\int P(\omega'|\omega, \sigma^*) dG(\varepsilon_f) = \underbrace{\sum_{a \in A(s_f)} F_P(\omega'|\omega, a) P(a|\omega)}_{F_P(\omega'|\omega)}$$

The term on the right hand side of this equation is simply the probability that the state moves from ω to ω', induced by the equilibrium conditional choice probabilities. I will denote that term by $F_P(\omega'|\omega)$.

Finally, observe that

$$\int \varepsilon_f(\sigma^*) dG(\varepsilon_f) = \sum_{a \in A(s_f)} \int_{\mathcal{C}(a,\omega)} \varepsilon_f(a) dG(\varepsilon) = \sum_{a \in A(s_f)} P(a|\omega) \mathbb{E}[\varepsilon_f(a) | a = \sigma(\omega, \varepsilon_f)]$$

It is well known that for the Type I Extreme Value distribution, $\mathbb{E}[\varepsilon_f(a)|a = \sigma(\omega, \varepsilon)] = \sigma(\gamma - \ln P(a|\omega))$, where γ is the Euler-Mascheroni constant. Therefore

$$\int \varepsilon_f(\sigma^*) dG(\varepsilon_f) = \sigma \sum_{a \in A(s_f)} P(a|\omega)(\gamma - \ln P(a|\omega))$$

Putting these pieces together, we have

$$V(\omega) = \left(\sum_a g(a, \omega, z) P(a|\omega) \right) \Psi + \sigma \sum_{a \in A(s_f)} P(a|\omega)(\gamma - \ln P(a|\omega))$$
$$+ \delta \sum_{\omega'} V(\omega') F_P(\omega'|\omega)$$

or

$$V(\omega) = \mathbb{E}_P[g(a, \omega, z)]\Psi + \sigma\gamma - \sigma\mathbb{E}_P[\ln P(a|\omega)] + \delta F_P(\omega)V$$

where $\mathbb{E}_P$ denotes an expectation with respect to a using the distribution over a defined by P, $F_P(\omega)$ is a row vector with the transition probabilities in state ω, and V a vector with the value function in each state ω.

We can now stack these equations. Let M_P denote the transition matrix induced by P, $M = [F_P(\omega'|\omega)]_{\omega,\omega'}$. Then[46]

$$V = \mathbb{E}_P[g(a, z)]\Psi + \sigma\gamma - \sigma\mathbb{E}_P[\ln P(a)] + \delta M_P V$$

From this equation we obtain

$$V = (I - \delta M_P)^{-1} \Big\{ \mathbb{E}_P[g(a, z)]\Psi + \sigma\gamma - \sigma\mathbb{E}_P[\ln P(a)] \Big\}$$
$$= \sigma K(P) + (I - \delta M_P)^{-1} \mathbb{E}_P[g(a, z)]\Psi$$

where $K(P) := (I - \delta M_P)^{-1}(\gamma - \mathbb{E}_P[\ln P(a)])$

The conditional value function is, by definition,

$$v(a, \omega) = g(a, \omega, z)\Psi + \delta \sum_{\omega'} V(\omega') F_P(\omega'|\omega, a) = g(a, \omega, z)\Psi + \delta F_P(\omega, a)V$$

[46]In this equation, it is to be understood that the scalar $\sigma\gamma$ is added to all coordinates. The ω-th coordinate of $\mathbb{E}_P[g(a, z)]$ is equal to $\sum_{a \in A(s_f)} g(a, \omega, z) P(a|\omega)$. Similarly for $\mathbb{E}_P[\ln P(a)]$

where $F_P(\omega, a)$ is the distribution over ω' induced by taking action a in state ω. Using the result above for V yields

$$\frac{v(a,\omega)}{\sigma} = \delta F_P(\omega,a)K(P) + \Big\{g(a,\omega,z) + \delta F_P(\omega,a)(I-\delta M_P)^{-1}\mathbb{E}_P[g(a,z)]\Big\}\sigma^{-1}\Psi$$

Chapter 2

Retailers' Product Portfolios and Negotiated Wholesale Prices

Private label products have been on the rise in recent years. It has been hypothesized that the increased popularity of private label products may moderate prices of national brands – see, e.g., (Mills, 1995) and (Morton and Zettelmeyer, 2004). By decreasing retailers' reliance on national brands, the argument goes, private label products increase retailers' bargaining leverage vis-à-vis national brand manufacturers. This increased bargaining leverage leads to lower wholesale and retail prices. The same argument applies to having multiple supply relationships: these relationships strengthen retailers' bargaining positions relative to any one of its suppliers. This should lead to lower wholesale and, possibly, retail prices.

This paper investigates whether these predictions are borne out by the data, and if so, what is the magnitude of those effects. To this end, I propose an empirical model of retail pricing and wholesale price negotiations and use it to measure the effect of private label products and multiple retailer-manufacturer relationships on wholesale prices, retail prices, and consumer welfare. I model retail chains, consisting of a collection of stores, negotiating simultaneously with multiple manufacturers. Stores are local monopolists and set prices to maximize profits, which are aggregated to form the chain's profits. I assume that chains are vertically integrated with respect to their private label products, and thus procure those goods at marginal cost. Branded products are acquired from their manufacturers and resold at the prices chosen by stores. Each chain negotiates simultaneously with its suppliers over linear wholesale prices, and the outcomes of these negotiations are modeled as a Nash equilibrium in Nash bargains – i.e., the wholesale prices charged by a given manufacturer to a given retailer maximize a Nash product for that manufacturer-retailer pair, holding fixed the wholesale prices agreed upon between the retailer and other manufacturers. Due to the local monopoly assumption, price negotiations are independent across retailers. I discuss these assumptions in greater detail in section 2.2.

I estimate the model using scanner data from the IRI Academic dataset (Bronnenberg, Kruger, and Mela, 2008). This dataset includes retail price and quantity information for 30 categories of consumer goods at the week-store-product(UPC) level. The data spans 50 different geographic regions, includes information on product characteristics and all stores are linked to the chains they belong to, which are anonymized. Wholesale prices are not observed. However, they can be inferred from retail price and quantity data paired with the assumed retail pricing model.

My analysis focuses on one category, namely peanut butter. I estimate demand separately for each chain-region pair, which proves to be challenging as standard instruments are not powerful to explain within-chain price variation. I construct a GMM estimator based on a restriction on the correlation between demand and cost unobservables at the level of the store. Using these demand estimates and the assumption of optimal pricing at the store level, I recover marginal costs for each store-week-UPC combination. I then use these marginal costs to obtain estimates of wholesale prices.

The estimated wholesale prices are used as inputs to estimate the bargaining model. I show that the bargaining model admits an inversion: it is possible to solve for manufacturer marginal costs as a function of wholesale prices and bargaining parameters. I use this result to construct a GMM estimator for bargaining and manufacturer marginal cost parameters. In estimation, I use a set of instruments inspired by the bargaining model and whose validity rests on the assumption that unobservable components of upstream marginal costs are uncorrelated across retailers. The estimation methodology proposed in this paper allows for the inclusion of a rich set of covariates in modeling upstream marginal costs, which lends credence to the identification strategy just outlined.

The estimation results show that retailers are indeed able to procure private label products at substantially cheaper prices. I estimate separate bargaining parameters for each of the three largest manufacturers and for the collection of the remaining manufacturers. These estimates are 0, 0.09, 0.45, and 1, which implies that the efficiency of vertical relationships in the grocery channel varies substantially across manufacturers.

Focusing on one retailer, I use the model to evaluate the questions posed above, namely what is the effect of private label products and relationships with multiple manufacturers on wholesale and retail prices.[47] Specifically, I conduct counterfactual exercises in which I remove all products of a given manufacturer from the retailer's product portfolio. I also perform simulations in which I exclude private label products from a retailer's portfolio. I find that wholesale prices do go up, but only marginally – at most by 0.68%. Removing private label products leads to an increase in the

[47]The analysis here focuses on one retailer because of the computational cost of counterfactual exercises. Revisions of this work will perform these exercises on all retailers observed in the data.

wholesale prices of national brands of only 0.10%. These upward pressures on prices are dominated by stores' incentives to reduce prices when their product portfolios shrink: removing the products of a national brand leads to retail price decreases ranging from 0.03% to 4.2%. Removing private label products does lead to increases in the retail prices of national brands, but of only 0.04%.

This paper relates to two intersecting strands of the Empirical Industrial Organization literature. First, it relates to the literature on the structural estimation of empirical bargaining models that employ the solution concept proposed by (Horn and Wolinsky, 1988). Such models were first estimated by (Crawford and Yurukoglu, 2012). Other noteworthy contributions to this literature are (Grennan, 2013), (Gowrisankaran, Nevo, and Town, 2015), and (Ho and Lee, 2017b). The (Horn and Wolinsky, 1988) solution concept mixes ingredients of cooperative and non-cooperative game theory; (Collard-Wexler, Gowrisankaran, and Lee, 2019) provide a fully non-cooperative foundation. This chapter makes methodological contributions to this literature. I use the inversion result mentioned above to derive a GMM estimator of bargaining and manufacturer marginal cost parameters for the multi-product bargaining model with downstream profit maximization. By allowing for optimal downstream pricing, this extends the econometric approach of (Gowrisankaran, Nevo, and Town, 2015). This methodology avoids repeatedly solving for equilibria, and enables the estimation of a flexible specification of manufacturers' marginal costs. This is important for the identification strategy employed in this chapter, which, as noted above, rests on the assumption that unobservable components of upstream marginal costs are uncorrelated across retailers.

This chapter also relates to the literature on inferences on vertical contracting along the grocery channel. Important contributions to this literature are (Villas-Boas, 2007), (Draganska, Klapper, and Villas-Boas, 2010), and, most recently, (Ellickson, Kong, and Lovett, 2018). The paper closest to this chapter is (Ellickson, Kong, and Lovett, 2018), henceforth EKL. EKL seek to quantify the effect of private label products on retailer profitability, and are particularly interested in the bargaining leverage mechanism outlined above. Similarly to this chapter, EKL set up a model of monopolist retailers that simultaneously negotiate with many upstream manufacturers. This chapter differs from EKL in three important ways. First, it has broader focus. As noted above, the bargaining leverage mechanism is not exclusive to private label products. This leads me to also consider the effects of multiple supply relationships on wholesale and retail prices. Second, I model multi-product negotiations, whereas EKL model single-product negotiations. The multi-product model is explicit about the extent that a retailer relies on a given manufacturer, and is thus more suitable for the study of the relevance of multiple supply relationships. Finally, my modeling and econometric approach allow me to model upstream marginal costs very flexibly, but my specification of bargaining parameters is parsimonious. Conversely, EKL's approach allows them to be flexible with respect to bargaining parameters but requires a parsimonious specification of marginal costs.

2.1 Data

The data used in this paper comes from the IRI Academic Database. For a description of the original release of this dataset, see (Bronnenberg, Kruger, and Mela, 2008). IRI provides information on prices and quantities at the store-week-UPC level, for 30 product categories and over 12 years. The data also provides information on the coarse geographic location of each store and to which (anonymized) chain each store belongs to.[48]

[48]There are 50 values for this location variable and these values are not always at the same geographic level. Most locations are cities but (i) there are cities of varying size, (ii) Two of the

My empirical analysis uses data on peanut butter sales during the year 2004. I focus on stores that appear in the data in at least 26 weeks. Conditional on this criterion, I restrict the sample to chains for which I observe at least 5 stores. Finally, conditional on these two criteria, for each store I keep only products that account for at least 5% of total revenues from peanut butter in that store in some week of the year. In the model introduced below, stores are assumed to be local monopolists and every pair (store, week) is treated as a market.[49] The size of the market is assumed to be 1.5 times the maximum total number of units of peanut butter sold in a given store, where the maximum is taken over the weeks in the data.

The selection criteria above yield me with 1,330,205 observations at the store-week-UPC level. There are 22 manufacturers in the data and 199 different products, of which 60 are private label products. The observations are distributed across all the 50 geographic locations in the data and across 86 chains. Figures 2.1 through 2.3 plot the distributions of, respectively, the number of stores, the number of suppliers, and the market share of private label products across chains, which are key determinants of retailers' bargaining leverage vis-à-vis manufacturers.[50] These figures show that there is substantial variation across chains in these three dimensions. This variation will enable the identification of bargaining parameters in the analysis that follows.

---

locations are regions (New England and West Texas/New Mexico) and (iii) two locations are states (Mississipi and South Carolina).

[49] For further discussion of this assumption, see section 2.2.

[50] Figure 2.1 plots the number of stores in the data. The data need not be exhaustive, but assuming IRI's sampling of stores across chains is similar, this suggests that these retail chains do differ in size. In the model and empirical implementation, the number of stores in the data will be assumed to reflect reality: chains' profits are defined to be the sum of stores' profits.

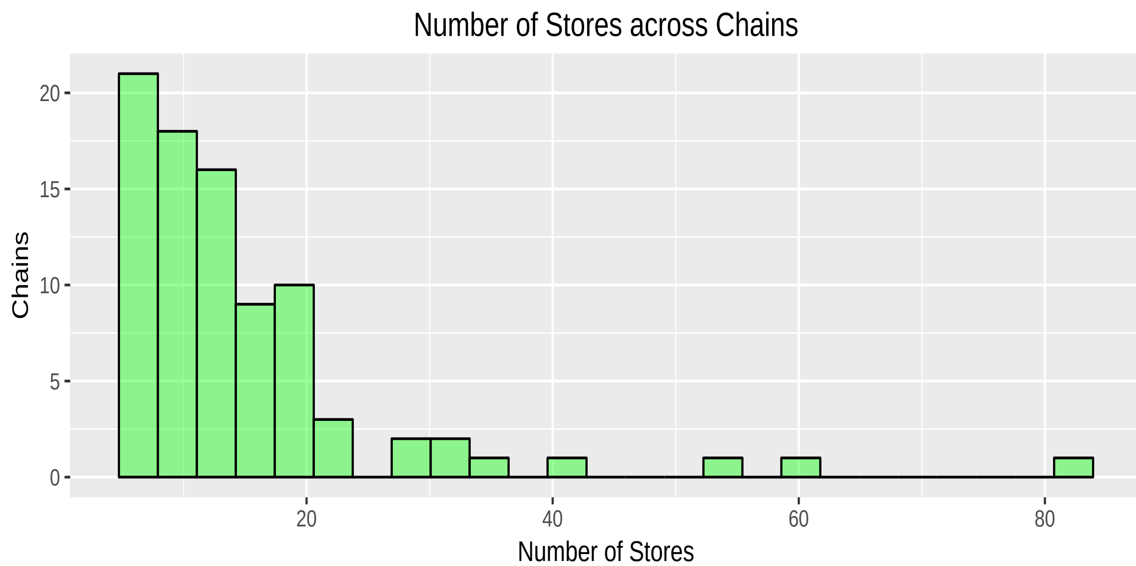

Figure 2.1: Distribution of the number of stores across chains

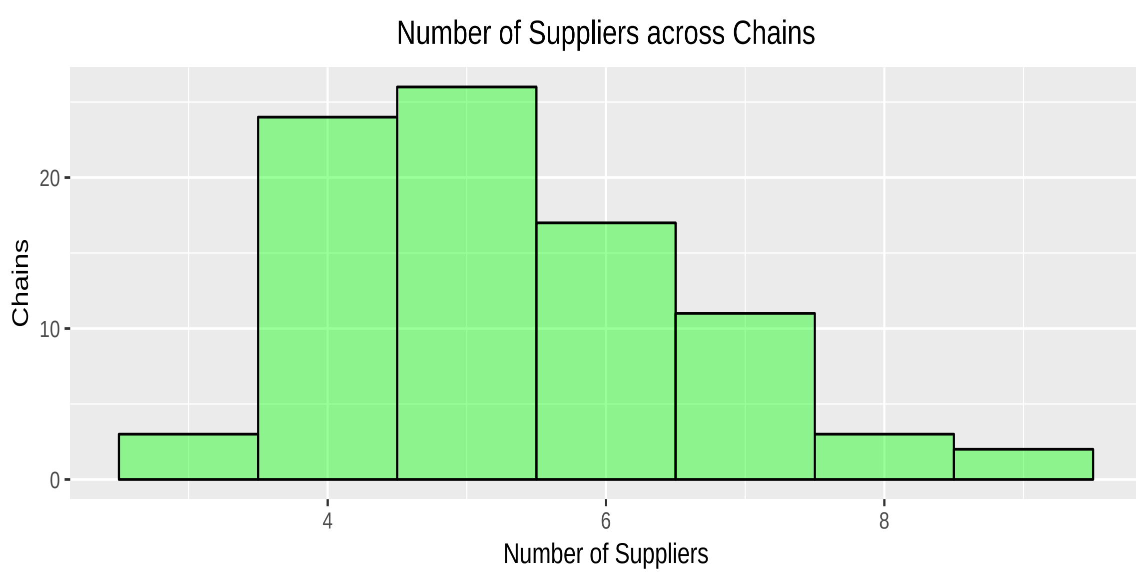

Figure 2.2: Distribution of the number of suppliers across chains

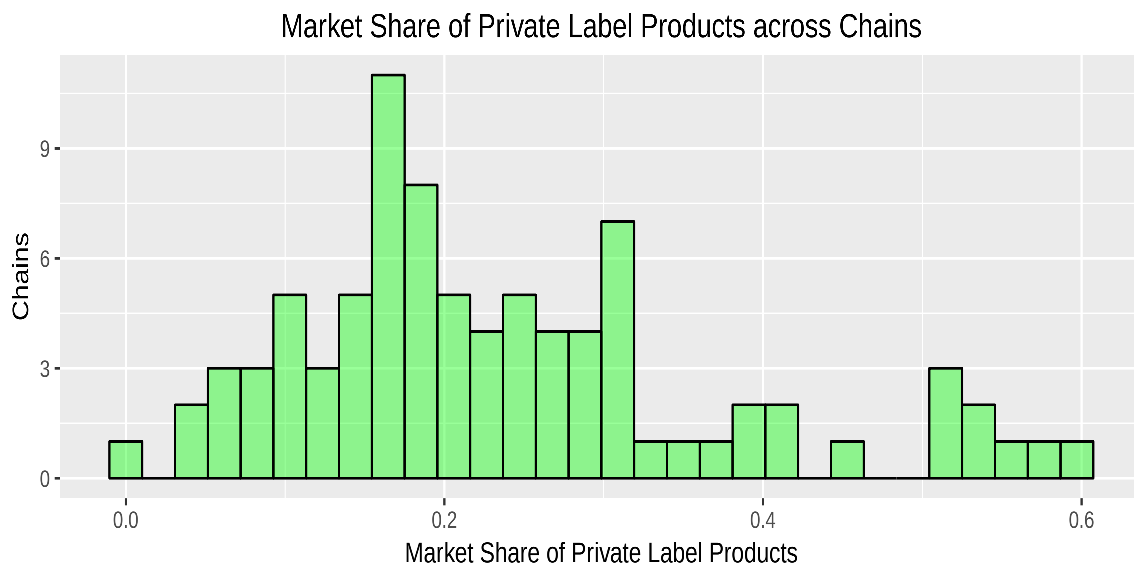

Figure 2.3: Distribution of the share of private label products across chains

2.2 Model

In section 2.2.1, I introduce the main components of the model, leaving demand functions unspecified. In section 2.2.2, I complete the description of the model by specifying consumers' decision problems, from which I derive stores' demand functions.

2.2.1 Model: Retailer Profit Maximization and Negotiated Wholesale Prices

Retail chains are indexed by h. Chain h owns stores $s = 1, \ldots, S_h$. The industry is endowed with a set of products, denoted by $\mathcal{J}$.[51] Products are indexed by $j = 1, \ldots, |\mathcal{J}|$. A store s carries an exogenously given portfolio of products $\mathcal{J}^s \subseteq \mathcal{J}$ and faces a demand function $D^s : \mathbb{R}^{|\mathcal{J}^s|} \to \mathbb{R}^{|\mathcal{J}^s|}$, mapping that store's prices into quantities demanded by consumers. Given marginal costs c^s, stores set prices to

[51] Empirically, $\mathcal{J}$ corresponds to the set of all products in the data.

solve

$$\max_p \sum_{j \in \mathcal{J}^s} (p_j - c_j^s) D_j^s(p)$$

Let $p^s(c^s)$ denote the solution to this problem. In section 2.2.3 I show that the solution to this problem is indeed unique given the demand functions I specify.

There are two strong assumptions in the model of retail prices just introduced. First, I have specified store s's demand as a function of its prices only, and have thus ruled out substitution across stores. Relaxing this assumption would require a model of store choice, which is beyond the scope of the present analysis. Allowing for substitution across stores would also make negotiations over wholesale prices interdependent across retailers, which would substantially complicate the bargaining model proposed below. The second strong assumption is that profit maximization occurs at the level of a store. This conflicts with recent empirical evidence showing that retailers set prices at a coarser level – see, e.g., (DellaVigna and Gentzkow, 2019) and (Adams and Williams, 2019). It should be noted, however, that the assumption of profit maximization at the store level is not essential for the analysis that follows. It could alternatively be assumed, e.g., that retailers make pricing decisions at the level of different geographic regions in the data.

Let $\mathcal{J}_h := \cup_{s=1}^{S_h} \mathcal{J}^s$ be the set of products sold by chain h. Each such product is either a private label product or a branded product, i.e., a product produced by some manufacturer m. I assume that chains are perfectly vertically integrated with respect to their private label products.[52] Let $\mathcal{J}_{h,m}$ be the set of products sold by chain h and manufactured by m. Chain h bargains with manufacturer m over the wholesale prices of these products, $w_{h,m} = (w_{h,j})_{j \in \mathcal{J}_{h,m}}$. Denote the vector of wholesale prices of all the branded products sold by chain h by w_h. I assume that, given w_h, the marginal

[52] An alternative assumption that is also empirically feasible is that all store brand products are produced by a single manufacturer.

costs of the stores owned by chain h are given by

$$c_j^s = \begin{cases} k_{h,j} + \tau^s + \eta_j^s, & \text{if } j \in \mathcal{J}_{PL}^s \\ w_{h,j} + \tau^s + \eta_j^s, & \text{if } j \notin \mathcal{J}_{PL}^s \end{cases} \tag{2.1}$$

where $\mathcal{J}_{PL}^s$ denotes the set of private label products sold by store s, $k_{h,j}$ is the chain's marginal cost of producing good j, and η_j^s are marginal cost shocks.[53]

Chain and manufacturers bargain over wholesale prices $w_{h,m}$ before the shocks to the stores' marginal costs are observed. After wholesale prices are agreed upon, cost shocks to the stores realize and the stores set prices $p^s(c^s)$. The chain then purchases $D_j^s(p^s(c^s))$ at $w_{h,j}$ and sells $D_j^s(p^s(c^s))$ at price $p_j^s(c^s)$ - for each store s. Therefore, the value for the chain of reaching an agreement with manufacturer m at wholesale prices $\hat{w}_m$, holding fixed the wholesale prices agreed upon with other manufacturers, w_{-m}, is given by

$$V_h(\hat{w}_m, w_{-m}; \mathcal{J}_h) = \mathbb{E}_\eta \left[\sum_{s=1}^{S_h} \sum_{j \in \mathcal{J}^s} (p_j^s(\tilde{c}^s(\hat{w}_m, w_{-m})) - \tilde{c}_j^s(\hat{w}_m, w_{-m})) \times \right.$$
$$\left. \times \; D_j^s(p^s(\tilde{c}^s(\hat{w}_m, w_{-m}))) \right]$$

If chain h and manufacturer m do not reach an agreement, the chain ceases to carry all of that manufacturers products.[54] Each store then faces an alternative demand

[53] As figure 2.4 illustrates, most of the price variation within a store is due to sales. Through the lens of the retail pricing model introduced in this chapter, sales are driven by cost shocks. That is far from a satisfactory economic theory of sales. That is not, however, the focus of the analysis here. Moreover, the large variation in stores' costs implied by the data and the retail pricing model is probably inconsequential for the subsequent analysis. That is because wholesale prices will be recovered by essentially averaging retailers' costs across weeks and stores – see, section 2.3.2 – and will therefore depend on the average price level rather than on week to week changes in price.

[54] An alternative assumption that appears in the literature – see, e.g., (Grennan, 2013) and (Ellickson, Kong, and Lovett, 2018) – is that negotiations occur product by product. In that case, if a negotiation were to fail, the chain would cease to carry only that one product. In that model, the extent to which a chain is dependent on a given manufacturer (beyond a single product) does not

$\bar{D}_j^{m,s} : \mathbb{R}^{|\mathcal{J}^s \setminus \mathcal{J}_{h,m}|} \to \mathbb{R}^{|\mathcal{J}^s \setminus \mathcal{J}_{h,m}|}$, where the superscript m indicates which negotiation failed. In case of disagreement, the chain obtains

$$V_h(w_{-m}; \mathcal{J}_h \setminus \mathcal{J}_{h,m}) = \mathbb{E}_\eta \Bigg[\sum_{s=1}^{S_h} \sum_{j \in \mathcal{J}^s \setminus \mathcal{J}_{h,m}} (p_j^s(\tilde{c}^s(w_{-m})) - \tilde{c}_j^s(w_{-m})) \times \\ \times \bar{D}_j^{m,s}(p^s(\tilde{c}^s(w_{-m}))) \Bigg]$$

If chain h and manufacturer m do reach an agreement, the value of the relationship for the manufacturer is

$$V_{m,h}(\hat{w}_m, w_{-m}) = \mathbb{E}_\eta \left[\sum_{s=1}^{S_h} \sum_{j \in \mathcal{J}^s \cap \mathcal{J}_{h,m}} (\hat{w}_j - c_j^m) D_j^s(p^s(\tilde{c}^s(\hat{w}_m, w_{-m}))) \right]$$

where c_j^m is the manufacturer's constant marginal cost of producing good j. Finally, if an agreement is not reached, the value of the relationship for the manufacturer is zero.[55] Let $\mathcal{J}_{h,B} := \mathcal{J}_h \setminus \cup_s \mathcal{J}_{PL}^s$ denote the set of branded products sold by chain h. I can now define the solution concept for this game.

Definition 2. *A vector $w \in \mathbb{R}^{|\mathcal{J}_{h,B}|}$ is a* ***subgame perfect Nash-in-Nash equilibrium (SPNiN) wholesale price vector*** *if, for every manufacturer m such that $\mathcal{J}_{h,m} \neq \emptyset$, the vector $w_m \in \mathbb{R}^{|\mathcal{J}_{h,m}|}$ solves*

$$\max_{\hat{w}_m} V_{m,h}(\hat{w}_m, w_{-m})^{b_{m,h}} \times \left(V_h(\hat{w}_m, w_{-m}; \mathcal{J}_h) - V_h(w_{-m}; \mathcal{J}_h \setminus \mathcal{J}_{h,m})\right)^{b_{h,m}} \tag{2.2}$$

where $b_{m,h}$ is the manufacturer's bargaining power when bargaining with chain h and

play a direct role in price negotiations. Because that is one of the dimensions I am interested in, the multi-product negotiation model is better suited for my purposes.

[55] The assumption underlying the way the value for the manufacturer in both contigencies (agreement or not) is specified is that the bargaining problems that a manufacturer faces with different chains are entirely independent of one another. This is a common assumption in the empirical bargaining literature. It should be noted that it requires (i) absence of downstream price competition and (ii) constant marginal costs for the manufacturer.

$b_{h,m}$ is the chain's bargaining power when bargaining with manufacturer m.[56]

2.2.2 Model: Demand

Each store faces a mass M^s of consumers that either buy a single product at the store or do not buy anything. If consumer i buys product j in store s,[57] which is owned by chain h and located in the geographic region l, she enjoys conditional indirect utility

$$u_{ijs} = \gamma_{jhl} + \phi_s + \alpha_{h,l} p_{js} + \psi_{h,l} a_{js} + \xi_{js} + \varepsilon_{ijs}$$

where γ_{jhl} is a chain and location specific product fixed effect, ϕ_s is a store fixed effect, p_{js} is the price of good j in store s, a_{js} is an advertisement dummy, ξ_{js} are product characteristics that are unobserved by the econometrician and ε_{ijs} are preference shocks.[58]

I will assume, as is standard, that the shocks ε_{ijs} are iid with a Type 1 Extreme Value distribution. Then the share of good j in store s is given by

$$\sigma_j^s(p) = \frac{\exp(\delta_j(p_j))}{1 + \sum_{k \in \mathcal{J}^s} \exp(\delta_k(p_k))}$$

where $\delta_k := \gamma_{khl} + \phi_s + \alpha_{h,l} p_{ks} + \psi_{h,l} a_{ks} + \xi_{ks}$

Store s thus faces the demand function

$$D_j^s(p) = M^s \sigma_j^s(p)$$

[56]Without loss of generality, I impose $b_{h,m} + b_{m,h} = 1$.

[57]In the model, consumers do not choose which store to patronize.

[58]This simple logit model imposes strong restrictions on substitution patterns, as is well known. Allowing for more flexible substitution patterns is important for the questions posed in this paper. A retailer will have greater bargaining leverage against those manufacturers that sell products most similar to the retailer's private labels or to other branded products. By restricting substitution patterns, the assumption of logit demand might meaningfully affect my results. A revision of this chapter will estimate a more flexible demand specification.

The disagreement demand functions $\bar{D}^{m,s}$ are similarly derived from the underlying discrete choice model.

2.2.3 Theoretical Results

This section establishes two results that are used in the subsequent analysis. First, I provide a full characterization of the solution - which turns out to be unique - to a monopolist's profit maximization problem under logit demand, as in the case of stores in the model introduced above.

Proposition 1. *Suppose a monopolist faces a demand function $D : \mathbb{R}^J \to \mathbb{R}^J$ given by*

$$D_j = M\sigma_j(p) = M\frac{\exp(\delta_j(p_j))}{1 + \sum_{k=1}^J \exp(\delta_k(p_k))}$$

where $\delta_k(p_k) = \gamma_k + \alpha p_k$ and $\alpha < 0$. Then

(i) There exists a unique solution p^* to the monopolist's profit maximization problem.

(ii) The solution p^* exhibits constant mark-ups, i.e., there exists a $\mu^* > 0$ such that

$$p_j^* - c_j = \mu^*, \quad \textit{for all } j = 1, \ldots, J$$

(iii) The optimal mark-up μ^* is given by the unique solution to

$$1 + \alpha\mu\sigma_0(c + \mu) = 0 \tag{2.3}$$

where $\sigma_0(p) = 1/\left(1 + \sum_k \exp(\delta_k(p_k))\right)$ is the share of the outside good, $c = (c_1, \ldots, c_j)'$ is the vector of marginal costs and $c + \mu$ means that μ is added to every coordinate of c.

(iv) p^* *is a continuously differentiable function of* $c \in \mathbb{R}^J$ *and its derivatives are given by*

$$\frac{\partial p_j^*}{\partial c_k} = \begin{cases} 1 + \frac{\partial \mu^*}{\partial c_j}(c) & \text{if } k = j \\ \frac{\partial \mu^*}{\partial c_k}(c) & k \neq j \end{cases}$$

where

$$\frac{\partial \mu^*}{\partial c_k}(c) = \frac{\alpha \mu^*(c) \sigma_k(c + \mu^*(c))}{1 - \alpha \mu^*(c)(1 - \sigma_0(c + \mu^*(c)))}$$

Proof. See appendix 2.A. ♠

Existence of a unique solution and the constant mark-up property are private cases of results in (Nocke and Schutz, 2018), but I provide independent proofs of those facts. The analysis of SPNiN equilibria that follows assumes that retail prices depend smoothly on wholesale prices - see Proposition 2. Part (iv) of Proposition 1 establishes that fact and characterizes the relevant derivatives. This explicit characterization makes computation considerably more efficient: the only step that has to be done numerically is the solution of equation (2.3), which is very well-behaved (see the proof of Proposition 1) and thus easy to solve numerically.

The next result, which characterizes SPNiN wholesale price vectors, is key for the approach introduced in section 2.3.3 for the estimation of bargaining and manufacturer marginal cost parameters.

Proposition 2. *Suppose* $w_h \in \mathbb{R}^{|\mathcal{J}_{h,B}|}$ *is a subgame perfect Nash-in-Nash equilibrium wholesale price vector. Let* $c_h \in \mathbb{R}^{|\mathcal{J}_{h,B}|}$ *be the vector of manufacturers' marginal costs of producing the goods sold by chain* h *and let* $m(j)$ *be the manufacturer of product* j*. Then the vector of wholesale markups,* $w_h - c_h$*, satisfies*

$$\left(\sum_{s=1}^{S_h} \Omega^s(w_h) + \Lambda^s(w_h) \right) (w_h - c_h) = - \sum_{s=1}^{S_h} \mathbb{E}_\eta [D^{s,h}(p^s(\tilde{c}(w_h)))]$$

where $D^{s,h}(p) \in \mathbb{R}^{|\mathcal{J}_{h,B}|}$

$$D_j^{s,h}(p) = \begin{cases} D_j^s(p) & \textit{if } j \in \mathcal{J}^s \\ 0 & \textit{otherwise} \end{cases}$$

the matrices $\Omega^s(w), \Lambda^s(w) \in \mathbb{R}^{|\mathcal{J}_{h,B}| \times |\mathcal{J}_{h,B}|}$ *are given by*

$$\Omega^s(w)_{j,k} = \begin{cases} \mathbb{E}_\eta \left[\nabla D_k^s(p^s(\tilde{c}^s(w)))' \frac{\partial p^s}{\partial c_j}(\tilde{c}^s(w)) \right] & \textit{if } j \in \mathcal{J}^s, k \in \mathcal{J}^s \cap \mathcal{J}_{h,m(j)} \\ 0 & \textit{otherwise} \end{cases}$$

and

$$\Lambda^s(w)_{j,k} = \begin{cases} -\frac{b_{h,m}}{b_{m,h} S_h(w)} \left(\sum_{s=1}^{S_h} \mathbb{E}_\eta [D_j^{s,h}(p^s(\tilde{c}(w)))] \right) \mathbb{E}_\eta [D_k^s(p^s(\tilde{c}^s(w)))] & \textit{if } k \in \mathcal{J}^s \cap \mathcal{J}_{h,m(j)} \\ 0 & \textit{otherwise} \end{cases}$$

and $S_h(w) = V_h(w; \mathcal{J}_h) - V_h(w_{-m}; \mathcal{J}_h \setminus \mathcal{J}_{h,m})$

Proof. See appendix 2.A. ♠

The usefulness of Proposition 2 stems from the fact that it allows me to solve for manufacturers' marginal costs as a function of wholesale prices and bargaining parameters. This, together with a model for manufacturers' marginal costs, allows me to estimate bargaining and manufacturer marginal cost parameters without solving the model. See section 2.3.3 for details.

A few final comments on the model are in order. The model introduced above assumes, as does almost all of the empirical bargaining literature, that the set of relationships between retailers and manufacturers are exogenously given. Recent work has provided ways of endogenizing these relationships – see (Ho and Lee, 2017a). Since the determination of those relationships is not the focus of this chapter, I maintain the assumption that they are exogenous. The model also takes the product portfolio at each store as exogenously given. It is possible to imagine a model in which the store's profit maximization problem is both over which products to offer - choosing a subset of the products procured by the chain - and prices. For the choice of products to be non-trivial, a constraint - arising, for example, from finite physical space - must be imposed, otherwise the solution with respect to the product variety is to offer all available products. Since I do not have data to estimate a model of optimal product variety, I take the product offerings at each store as exogenous.

I also assume, in line with the empirical bargaining literature, that if disagreement between manufacturer m and chain h occurs, w_{-m} is held fixed. A perhaps natural alternative assumption is that the wholesale prices that occur under disagreement are themselves the outcome of a Nash-in-Nash bargaining game. Computing equilibria for such a model would require the calculation of a large number of Nash-in-Nash equilibria, which would be computationally demanding. That being said, Proposition 2 goes through without change as long as the disagreement payoff for the chain is independent of $\hat{w}_m$, which would be true in the alternative model just suggested.

2.3 Econometrics

Subsection 2.3.1 provides details on demand estimation; subsection 2.3.2 explains how I construct wholesale prices; subsection 2.3.3 discusses the estimation of bargaining parameters and manufacturers' cost parameters.

2.3.1 Demand Estimation

The conditional indirect utility that consumer i derives from product j when buying it at store s was assumed to be given by

$$u_{ijs} = \gamma_{jhl} + \phi_s + \alpha_{h,l} p_{js} + \psi_{h,l} a_{js} + \xi_{js} + \varepsilon_{ijs}$$

where the ε_{ijs} follow independent Type 1 Extreme Value distributions.[59] As is well known, this model implies the equation

$$\ln(\sigma_j^s) - \ln(\sigma_0^s) = \gamma_{jhl} + \phi_s + \alpha_{h,l} p_{js} + \psi_{h,l} a_{js} + \xi_{js} \tag{2.4}$$

which can be taken to data.

As usual, endogeneity of p_{js} is a concern: if the retailer or the manufacturer has information on ξ_{js}, then prices will be correlated with ξ_{js}. Because of the nature of the questions studied in this paper, I am interested in estimating demand functions that are store-specific (or at least chain-specific). I want to allow, e.g., for patrons of a retail chain to enjoy a product whereas patrons of a different chain dislike it. This allows those different retailers to be differentially dependent on that product, which affects the wholesale prices they negotiate. For this reason, I will not aggregate the data across chains.

Using the data at the retailer level introduces difficulties in the demand estimation. The reason is that standard instruments - for example, the average price of the product in other markets,[60] see (Nevo, 2001) and (Hausman, 1996) - are not powerful to

[59]See 2.2.2 for the definitions of the other terms.

[60]In my context, this would have to be adapted to the average price of the product in other markets and in other chains, because wholesale prices are common across stores that belong to the same chain.

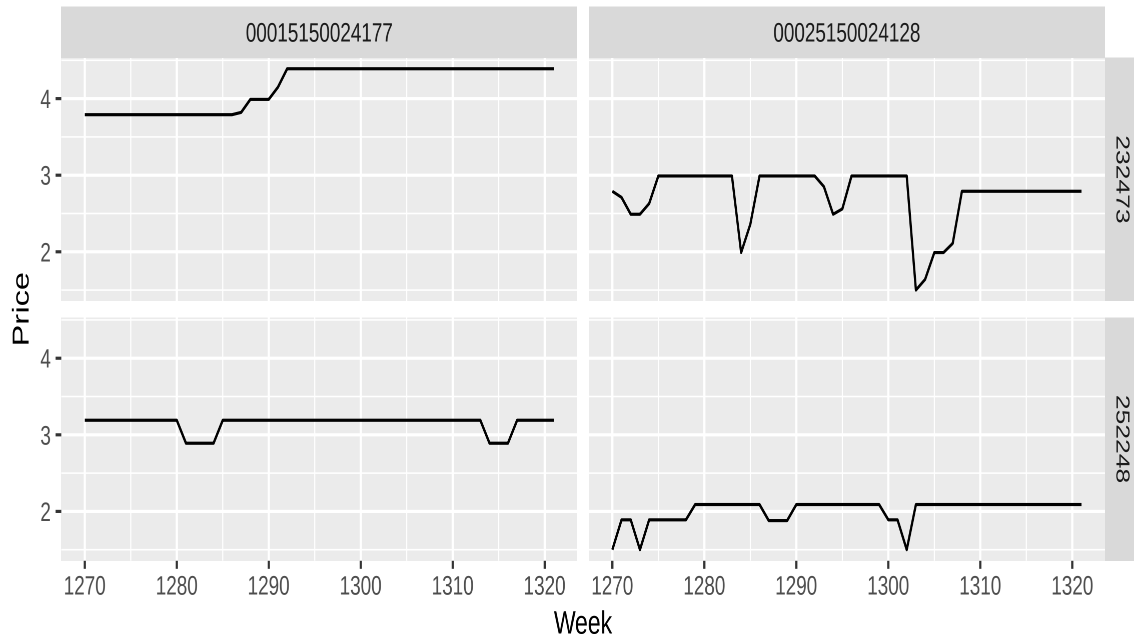

Figure 2.4: Price paths for two products (columns) in two different stores (rows).

explain within-store price variation.[61] The reason is that within-store price variation is driven by sales – as illustrated by figure 2.4 –, for which prices in other markets are not a strong predictor.

A common idea for generating instrumental variables for prices is finding exogenous cost shifters. The model introduced above makes cost shocks to the stores explicit. Cost shocks also enter store prices explicitly: stores set their prices equal to $p^s(\tilde{c}^s(w_h, \eta^s))$. Thus, if stores' cost shocks η^s were observed, they would be an ideal instrument. Based on this intuition, I construct a GMM estimator based on the restriction that stores' cost shocks and unobservable demand factors at the store level are uncorrelated, i.e.,

$$\mathbb{E}[\eta_j^s \xi_{js}] = 0 \tag{2.5}$$

I estimate demand separately for each (chain, location) pair in the data, for a

[61] Aggregating across stores within a chain would not help much, because prices within a chain are highly correlated, as discussed above. See also (**adams2017zone**) and (**dellavigna2017uniform**).

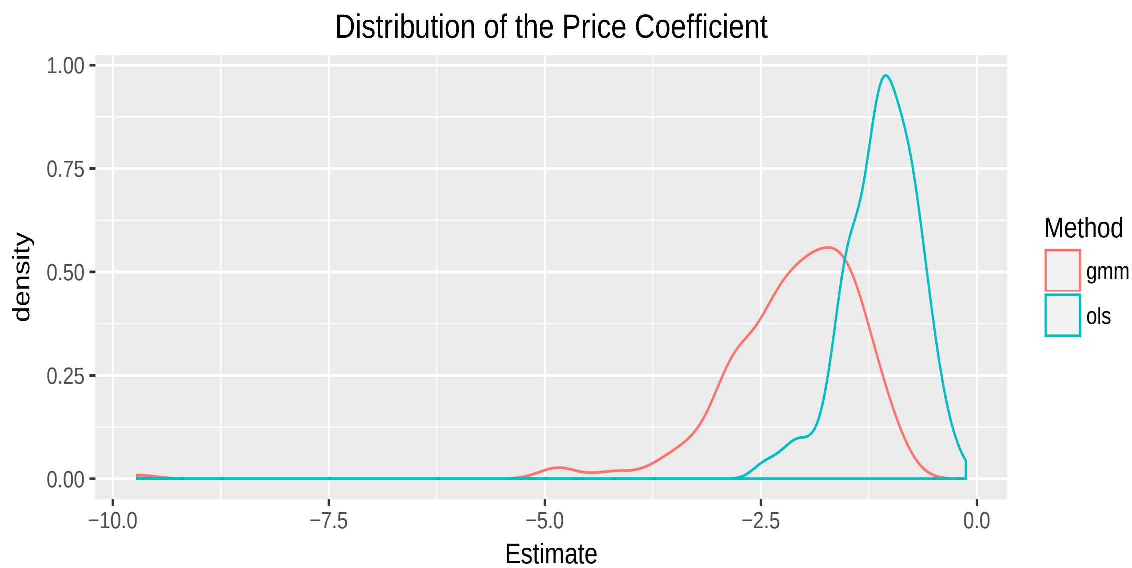

Figure 2.5: Distribution of price coefficient estimates under OLS and GMM.

total of 206 demands, using OLS and the GMM estimator above.[62] Under the OLS estimator, 4.23% of the own price elasticities are less than one in absolute value. For the GMM estimator, this figure is 0.03%. The median of the distribution of own price elasticities is -2.83 for the OLS estimator and -5.02 for the GMM estimator. Because I am modelling the demand for peanut butter within a store, the GMM results seem to be more credible than the OLS results. Under the OLS estimator, 13.7% of the marginal costs inferred from stores' FOCs are negative; that figure is equal to 1.24% for the GMM estimator. Figure 2.5 plots the distribution of the estimates obtained under each method.

2.3.2 Construction of Wholesale Prices

Stores solve

$$\max_{p} \sum_{j \in \mathcal{J}^s} (p_j - c_j^s) D_j^s(p)$$

[62]The GMM estimator is based on (2.5) and exogeneity restrictions for the variables in equation (2.4) other than price.

The first order conditions can be solved for stores' marginal costs:

$$c^s = p^s + (J_{\sigma^s}(p^s)')^{-1}\sigma^s(p^s) \tag{2.6}$$

where σ^s is the share function for store s, given by

$$\sigma_j^s(p) = \frac{\exp(\delta_j^s)}{1 + \sum_k \exp(\delta_j^s)}$$

and $J_{\sigma^s}(p)$ is its Jacobian with respect to prices.

From equation (2.6) I can thus obtain the marginal costs as a function of estimated demand parameters and data.[63] Having recovered stores' marginal costs, I estimate the store marginal cost model (2.1). Specifically, for each chain I run the stores' marginal costs obtained via equation (2.6) on product and store dummies. The coefficient estimates on the product dummies are the estimated wholesale prices. For the purpose of estimating the bargaining part of the model, these wholesale prices are treated as data. I obtain 3250 observations of wholesale prices at the (Chain,UPC) level. Table 2.1 verifies that retailers indeed face lower wholesale costs for private label products.[64]

[63]For the logit model things are even simpler. In that case we have

$$\frac{\partial \sigma_j^s}{\partial p_k} = \begin{cases} \alpha \sigma_j^s(p)(1 - \sigma_j^s(p)) & \text{if } k = j \\ -\alpha \sigma_j^s(p)\sigma_k^s(p) \end{cases}$$

Marginal costs can thus be recovered as a function of α and data - the remaining demand parameters are not needed.

[64]Given that retail prices of private label products are lower than those of branded alternatives, their inferred wholesale costs must also be lower under logit demand and store-level monopoly pricing, due to the constant mark-up property established above.

Table 2.1: Do Stores Face Lower Marginal Costs for Store Brand Products?

	Dependent variable:		
	$\ln(w_{jh})$	$\ln(w_{jh})$	$\ln(w_{jh})$
	(1)	(2)	(3)
Weight			0.031^{***}
			(0.0004)
Private Label Dummy	-0.241^{***}	-0.241^{***}	-0.404^{***}
	(0.026)	(0.026)	(0.014)
Chain FE	No	Yes	Yes
Observations	3,250	3,250	3,250
R^2	0.026	0.100	0.737
Adjusted R^2	0.025	0.076	0.729
Residual Std. Error	0.601 (df = 3248)	0.586 (df = 3163)	0.317 (df = 3162)
Note:			*p<0.1; **p<0.05; ***p<0.01

2.3.3 Estimation of Manufacturer Cost and Bargaining Parameters

The model introduced above allows bargaining parameters to vary both across manufacturers and retailers. In estimation, I will allow them to vary across manufacturers only. Moreover, there are three large manufacturers in the data and many small manufacturers.[65] I will assume that all the small manufacturers have the same bargaining parameter. Therefore, there are 4 bargaining parameters to be estimated.

Proposition 2 says that the vector of SPNiN wholesale prices for chain h, w_h, satisfies

$$\left(\sum_{s=1}^{S_h} \Omega^s(w) + \Lambda^s(w)\right)(w_h - c_h) = -\sum_{s=1}^{S_h} \mathbb{E}_\eta[D^{s,h}(p^s(\tilde{c}(w)))] \tag{2.7}$$

[65] The three large manufacturers are Conagra Foods Inc, Unilever and J M Smucker Co.

Therefore, as long as the matrix on the left hand side of this equation is invertible, I can solve for c_h, the manufacturers' marginal costs of producing the goods sold by chain h as a function of wholesale rices and bargaining parameters.

It also turns out that the matrix on the left hand side of equation (2.7) is block-diagonal, where the blocks correspond to the different manufacturers that negotiate with chain h. This implies that the resulting marginal costs depend only on the bargaining power of that manufacturer. Therefore, I may write

$$c_{jh} = w_{jh} - \mu(w_h, b_{m(j)}), \quad j \in \mathcal{J}_{h,B}, h = 1, \ldots, H \tag{2.8}$$

where μ is the mark-up term obtained by solving the system of linear equations in (2.7), $m(j)$ is the manufacturer that produces good j and, as before, $\mathcal{J}_{h,B}$ is the set of branded products sold by chain h.

Now suppose manufacturers' marginal costs are given by

$$c_{jh} = x'_{jh}\gamma + \nu_{jh} \tag{2.9}$$

where x_{jh} are observable characteristics of the product and the chain and ν_{jh} is an unobservable component of cost. Specifically, I include in x_{jh} manufacturer fixed effects, chain fixed effects, the product's weight, a dummy for reduced sugar products, a dummy for more expensive production processes,[66] a dummy for reduced sodium, a dummy for chunky or crunchy peanut butter and a dummy for flavored products.

Putting equations (2.8) and (2.9) together, I can write

$$\nu_{jh}(b_{m(j)}, \gamma, w_h, x_{jh}) = w_{jh} - \mu(w_h, b_{m(j)}) - x'_{jh}\gamma \tag{2.10}$$

[66] These more expensive production processes include natural, kosher and organic products.

In equation (2.10), the marginal cost shocks ν_{jh} are written as a function of data and structural parameters. However, the expression in the right hand side of that equation depends on w_h, which in turn depends on ν_{jh}. Therefore, to identify the bargaining parameters b_m, instruments generating exogenous variation in w_h are necessary. Given appropriate instrumental variables z_{jh}, estimation of manufacturers' marginal cost parameters (γ) and bargaining parameters (b_m) can be accomplished by GMM, based on the conditional moment restriction

$$\mathbb{E}[\nu_{jh}|z_{jh}] = 0$$

Choice of Instruments and Identification

I use as instruments z_{jh} the following variables:

(i) the cost covariates x_{jh}.

(ii) Product j's expected demand at chain h, under the optimal retail prices implied by the average wholesale prices in other chains, interacted with manufacturer dummies.

(iii) Total expected demand for goods produced by the manufacturer of product j, under the optimal retail prices implied by the average wholesale prices in other chains, interacted with manufacturer dummies.

(iv) Manufacturer mark-up for product j, as obtained from Proposition 2, computed under the average wholesale prices in other chains and assuming $b_m = 1/2$, interacted with manufacturer dummies.

The cost covariates are all assumed to be exogenous with respect to ν_{jh}. Note that the instruments in bullets (ii) through (iv) all use average wholesale prices in other chains, which depend on $\nu_{jh'}$, $h' \neq h$. Therefore, the main assumption underlying validity of these instruments, and hence the main identifying assumption, is that $\mathbb{E}[\nu_{jh}\nu_{jh'}] = 0$, i.e., that manufacturers' marginal cost shocks are uncorrelated across chains. This seems like a reasonable assumption given that (i) negotiations occur only once[67] and (ii) the cost model includes a rich set of product characteristics and manufacturer dummies.[68]

Power comes from the fact that variables (ii)-(iv) influence the value of a relationship to the relevant parties, and thus equilibrium wholesale prices. For example, if the customers of chain h perceive product j to be of high quality, demand for that product will tend to be high and the manufacturer, being aware of that, can charge larger wholesale prices. As another example, the variable in (iv) predicts wholesale mark-ups assuming $b_m = 1/2$. This variable is correlated with actual wholesale mark-ups, which mechanically influence wholesale prices.

The discussion above shows that identification of bargaining and cost parameters comes from variation across chains. To build intuition, think of one manufacturer (labeled m) producing one good and negotiating with two retailers, $h = 1, 2$. The customers of retailer 1 dislike the manufacturer's product and the customers of manufacturer 2 really enjoy the product. The wholesale price at which the good will be sold to retailer 1, call it w_1, will tend to be close to the manufacturer's marginal cost c_m. The wholesale price at which the good will be sold to retailer 2, w_2, on the other hand, will tend to be larger. Exactly how much larger will depend on how much the manufacturer is able to capitalize on the fact that retailer 2's customers enjoy the

[67]In a model with repeated negotiations, $\mathbb{E}[\nu_{jht}\nu_{jh't}] = 0$ would hardly be a compelling assumption, but in that situation one might use the panel structure of the data to generate alternative moment conditions.

[68]Suppose I had ommitted, say, the reduced sugar dummy. These products might use more expensive sweeteners as ingredients, and thus in that model marginal cost shocks might be correlated across chains

product. This is determined by the bargaining parameter b_m.

Estimation and Results

A GMM estimate is the solution to the program

$$\min_{\gamma,b} \bar{g}_n(\gamma,b)' W_n \bar{g}_n(\gamma,b) \tag{2.11}$$

where $\bar{g}_n(\gamma,b) := n^{-1}\sum_{j,h}\nu_{jh}(b_{m(j)},\gamma,w_h,x_{jh})z_{jh}$, z_{jh} is a column vector with the instrumental variables described above and W_n is a weight matrix.[69] I implement the standard two step procedure to obtain an optimal GMM. In the first step, I take as weight matrix $W_n = (Z'Z)^{-1}$, where Z is the matrix of instruments. Let $\hat{\theta}_1$ be the resulting estimate for $\theta = (\gamma, b)$. With that estimate, I construct an estimate of the optimal weight matrix

$$\hat{W}_n^* = \left(n^{-1}\sum_i g_{jh}(\hat{\theta}_1)g_{jh}(\hat{\theta}_1)'\right)^{-1}$$

where $g_{jh}(\theta) = \nu_{jh}(b_{m(j)},\gamma,w_h,x_{jh})z_{jh}$. The second step consists of minimizing (2.11) again, using $\hat{W}_n^*$ as the weight matrix.

It is important to note that, conditional on a value of the bargaining parameters b, estimation is linear on the cost parameters γ - which can be seen from equation (2.10). One evaluation of the GMM objective (2.11), for a fixed value of b, consists of the following steps:

1. Apply Proposition 2 to obtain manufacturers' marginal costs as a function of data and bargaining parameters. Let the resulting vector of marginal costs be

[69] There are 111 instruments in total, 96 of which are cost covariates. The large number comes from the chain dummies.

denoted by $C(b)$.[70]

2. Obtain the implied estimates for γ, given by

$$\hat{\gamma} = (X'ZWZ'X)^{-1}X'ZWZ'C(b)$$

where $C(b)$ is the vector of marginal costs for each (product, chain) pair, obtained in the previous step.

3. Compute $\hat{\nu}_{jh}(b_{m(j)}, \gamma, w_h, x_{jh}) = c_{jh}(w_h, b_{m(j)}) - x'_{jh}\hat{\gamma}$.

4. Compute $\bar{g}_n(\hat{\gamma}, b) = n^{-1}\sum_{j,h}\hat{\nu}_{jh}z_{jh}$ and form the GMM objective in (2.11).

Because γ can be found in closed form for a given value of b, the nonlinear search can be restricted to the bargaining parameters. I minimize the GMM objective using a derivative-free global optimizer with somewhat loose termination parameters.[71] Once the global optimizer has found a solution, that solution is used as the starting point for a derivative-free local optimizer, now with tighter termination conditions.[72]

I compute standard errors using the standard result for the asymptotic distribution of extremum estimators, e.g., (Newey and McFadden, 1994). There are two reasons why these are incorrect. First, they fail to account for the variance coming from the construction of wholesale prices performed before the estimation of the bargaining parameters. Second, the characterization of the asymptotic distribution of extremum estimators is based on the first order conditions for an interior solution for

[70] To compute the terms in equation 2.7, I (i) set ξ_{js} to zero and (ii) compute the expectations with respect to η^s by simulation, sampling from the residuals obtained from running the stores' marginal cost regressions.

[71] Controlled Random Search, see (Kaelo and Ali, 2006). I use the implemention in NLOPT. See http://ab-initio.mit.edu/nlopt.

[72] COBYLA, see (Powell, 1994). I use the implementation in NLOPT. See http://ab-initio.mit.edu/nlopt.

Manufacturer	$\hat{b}_m$	**Std. Dev.**
Conagra	0.0886	0.2393
Unilever	10^{-5}	0.1138
J M Smucker Co.	0.4489	0.0911
Others	1	0.4367

Table 2.2: Bargaining parameter estimates

the estimation program, but the estimate I obtain is not interior. From now on, I will ignore these two caveats.

As mentioned above, bargaining parameters are allowed to vary across manufacturers, but since there are three large manufacturers and many small ones, I assume that all the small manufacturers have the same bargaining parameter.[73] Table 2.2 shows the results.[74] The estimates show that Conagra and Unilever have little bargaining power in the vertical chain. The estimate for Smucker is close to 1/2 and small manufacturers (labeled "Others" in table 2.2) seem to have substantial bargaining power in the vertical chain.

2.4 Counterfactual Exercises

In this section, I use the model to answer the two questions posed in the beginning of this chapter. First, what is the effect of private label products on wholesale and retail prices? Second, what are the effects of multiple supply relationships on wholesale and retail prices? The analysis here focuses on one retail chain.[75] This chain negotiates with all manufacturers and sells private label products. Specifically, the chain I consider sells 34 different products: 12 produced by J M Smucker Co., 11 produced by

[73]Conagra is the smallest of the "large" manufacturers: there are 579 observations at the (UPC, Chain) level where the UPC is manufactured by Conagra. The largest of the "small" manufacturers is Hersheys, with 60 (UPC, Chain) pairs in the data.

[74]10^{-5} was the lower bound imposed on the GMM problem.

[75]These results are preliminary and future revisions will perform the same exercises for all chains.

Conagra Foods Inc., 6 produced by Unilever, 4 private label products and 1 produced by one of the small manufacturers. For the purpose of the counterfactual exercises, I assume there are no shocks to stores' marginal costs, for computational convenience.[76]

The main set of counterfactuals consists of excluding all products of a given manufacturer (including a retailer's own products, i.e., the private labels) from the retailer's product portfolio. I also perform exercises in which I assume that all products of a given manufacturer become private label products and simulate mergers between two manufacturers. I use the model to compute changes in wholesale and retail prices.

A Nash-in-Nash equilibrium wholesale price vector must satisfy the system of first order conditions given by

$$b_m \frac{\partial V_m}{\partial \hat{w}_j}(w) \frac{1}{V_m(w)} + (1 - b_m) \frac{\partial V_h}{\partial \hat{w}_j}(w) \frac{1}{S_h(w)} = 0, \quad \forall j \in \mathcal{J}_{h,B} \tag{2.12}$$

This is a complex system of equations.[77] In particular, it takes into account how stores change their prices once wholesale prices change.[78] To be able to efficiently evaluate the left hand side of equation (2.12), I show in appendix 2.B how the Implicit Function Theorem allows me to further characterize how stores' optimal prices change in response to changes in wholesale prices.[79] For all of the counterfactual scenarios

[76]Computing the expectations in equation 2.7 needs to be done only once for estimation, which is feasible. As explained in this section, solving for SPNiN equilibria involves evaluating those expectations multiple times, which is costly.

[77]Good starting points turn out to be critical to solve this system of as many as 34 equations (the number of equations varies with the counterfactual being considered). For many starting values I do not find solutions. A set of good starting points is given by

$$w_j = c_j(1 + b_{m(j)}\mu)$$

where $\mu \in [0, 1]$. The logic is that under these prices the margins $(w_j - c_j)/c_j$ are proportional to the bargaining parameters.

[78]See the proof of Proposition 2 for the relevant calculations.

[79]Note that Proposition 1 subsumes Appendix 2.B. However, the results shown here were obtained using the analysis in Appendix 2.B. A future revision will use Proposition 1 instead, as it provides a sharper and computationally more efficient characterization of how stores' optimal prices change in response to changes in wholesale prices.

shown below, I find a unique SPNiN.[80]

All tables in this section have the same structure as table 2.3: entry (i, j) shows the results when it is assumed that all the products of manufacturer i are produced and sold by manufacturer j instead; the column "Excluded" shows the results under the assumption that the products of manufacturer i cease to be sold by the retailer. In counterfactual (i, j) it is assumed that the bargaining power of the resulting firm is equal to the bargaining power of manufacturer j. Therefore, the merging firms in counterfactuals (i, j) and (j, i) are the same, but the bargaining power of the resulting firm differs across these two cases.

Table 2.3 shows the average (across products) wholesale price percent change, where the average is computed across products that are branded products in both the benchmark and in the relevant counterfactual. Excluding a manufacturer's products increases the wholesale prices of the remaining products, but not by a lot. Those effects range from 0.003% to 0.68%. In particular, excluding private label products from the retailer's product portfolio leads to only an average 0.096% increase in the wholesale prices of the remaining products. Similarly, assuming that branded products become private label products decreases the wholesale prices of the other branded products, but again the effect is small. Those effects range from no change to an average decrease of 0.053% in the prices of the other branded products. Mergers can have large effects on equilibrium wholesale prices, which can both increase and decrease. These results are driven by differences in firms' bargaining powers.

Table 2.4 shows the average (across products) percent change in retail prices of products that are branded in both the benchmark and in counterfactual (i, j). As shown above, excluding a manufacturer's products leads to a small increase in the

[80] I look extensively for solutions to the system of equations (2.12). As mentioned in the text, I find a unique solution. For each manufacturer, I check that the corresponding Nash product cannot be increased, and find that to be the case.

Table 2.3: Wholesale Price Inflation: Products Negotiated in Both Scenarios

	Excluded	Private Label	Conagra	Smucker	Unilever	Others
Private Label	0.096		0.050	0.197	0.0001	0.092
Conagra	0.677	−0.044		7.051	−1.124	11.234
Smucker	0.138	−0.053	−4.357		−6.180	7.242
Unilever	0.153	0	0.639	3.670		6.056
Others	0.003	−0.004	7.231	−0.400	1.556	

wholesale prices of the remaining products. However, with fewer products, the retailer has an incentive to reduce the price of the remaining products. The reason is that with fewer products it is more likely that marginal consumers will move to the outside option after a price increase. This effect dominates in all cases, except for the exclusion of private labels, when the increase in wholesale price dominates - but the resulting effect on retail prices is small.

Assuming that branded products become private label products leads to increases in the retail prices of the remaining branded products. Technically, the reason is that the retailer's optimal (constant) mark-up increases when the costs of some of the goods decrease.[81] Economically, the lower costs of the new private label products allow the retailer to charge larger mark-ups while still retaining many consumers. The retailer thus optimally increases its mark-up and a fraction of consumers shifts towards the new private label products. Once again, upstream mergers can lead to retail price increases or decreases; the effect depends on firms' bargaining parameters.

Table 2.4: Retail Price Inflation: Negotiated Products

	Excluded	Private Label	Conagra	Smucker	Unilever	Others
Private Label	0.037		−0.040	−0.273	0.0001	−0.485
Conagra	−2.173	0.206		3.388	−0.507	5.911
Smucker	−4.210	2.493	−1.320		−1.855	3.163
Unilever	−0.393	0.00000	0.389	2.270		3.879
Others	−0.027	0.046	5.109	−0.274	1.444	

[81] See Proposition 1.

Finally, table 2.5 shows the percent change in consumer surplus. Excluding products from the retailer's portfolio has a large negative impact on consumers' surplus, because of the diminished variety. The other results are the direct consequence of equilibrium price changes. If the products of a national brand become private labels, consumer surplus increases because of the resulting decreases in price.[82] The welfare effects of mergers are also driven by the associated price effects, which in turn align with the bargaining parameters of the manufacturers involved.

Table 2.5: Changes in Consumer Surplus

	Excluded	Private Label	Conagra	Smucker	Unilever	Others
Private Label	−0.227		−0.575	−3.155	0.0002	−4.266
Conagra	−22.401	1.923		−10.868	1.922	−13.148
Smucker	−33.498	17.903	12.703		17.902	−14.063
Unilever	−3.834	0.00004	−0.407	−2.070		−2.419
Others	−0.226	0.387	−0.225	0.083	−0.650	

2.5 Conclusion

This paper studies vertical relationships between retailers and manufacturers. Specifically, I investigate the effects of private label products and relationships with multiple manufacturers on wholesale prices, retail prices, and consumer welfare.

To address these questions, I develop and estimate a model of wholesale price negotiation between retail chains – which consist of collections of stores – and multiple manufacturers. The model allows for downstream profit maximization and price negotitions over multiple products. I show that a covariance restriction between demand and store-level cost unobservables can be used to estimate demand at the level

[82]Table 2.4 shows price increases for products that are branded under both the baseline and the counterfactual. That does not include the products that become private labels under the counterfactual, whose price decreases.

of a store, yielding credible demand elasticities for the empirical setting considered here. To estimate bargaining parameters and manufacturer marginal costs, I extend the econometric methodology of (Gowrisankaran, Nevo, and Town, 2015) to the case with downstream profit maximization. This method allows for the estimation of flexible specifications for manufacturers' marginal costs. This flexibility underpins the identification of manufacturers' bargaining parameters, which rests on the assumption that unobserved components of relationship specific marginal costs are uncorrelated across chains.

I find that bargaining parameters do vary across manufacturers, suggesting varying levels of vertical efficiency in the grocery channel. Counterfactual exercises show that private label products reduce wholesale and retail prices of national brands, but only slightly. Similarly, negotiations with multiple manufacturers do reduce wholesale prices of competing manufacturers, but those effects are small. Overall, these results suggest that retailer bargaining leverage vis-à-vis manufacturers plays only a minor role in determining wholesale and retail pricing.

The model proposed in this paper makes the strong assumption that stores are local monopolists. Extending the analysis to allow for downstream price competition and bargaining externalities across retailers is a promising area for further research.

2.A Proofs and Auxiliary Results

Proof of proposition 1

Proof. The first order condition of the profit maximization problem with respect to p_j is

$$\sigma_j(p) + \sum_k (p_k - c_k)\frac{\partial \sigma_k}{\partial p_j}(p) = 0$$

Using the logit functional form this becomes (and dropping the argument p)

$$\sigma_j + \alpha\sigma_j(1-\sigma_j)(p_j - c_j) - \alpha\sum_{k\neq j}(p_k - c_k)\sigma_j\sigma_k = 0$$

or, since $\sigma_j(p) > 0$ for all j and all p,

$$1 + \alpha(1-\sigma_j)(p_j - c_j) - \alpha\sum_{k\neq j}(p_k - c_k)\sigma_k = 0$$

Now note that $(1-\sigma_j) = \sigma_0 + \sum_{k\neq j}\sigma_k$, so that the previous equation can be rewritten as

$$\begin{aligned}
0 &= 1 + \alpha\sigma_0(p_j - c_j) + \alpha\left(\sum_{k\neq j}\sigma_k(p_j - c_j) - \sum_{k\neq j}\sigma_k(p_k - c_k)\right) \\
&= 1 + \alpha\sigma_0\mu_j + \alpha\sum_{k\neq j}\sigma_k(\mu_j - \mu_k) \\
&= 1 + \alpha\sigma_0\mu_j + \alpha\sum_{k=1}^{J}\sigma_k(\mu_j - \mu_k) \qquad (2.13)
\end{aligned}$$

where $\mu_k := p_k - c_k$. In a solution to the problem, this equation must hold for all $j = 1, \ldots, J$. Take two arbitrary products j_1, j_2 and subtract the corresponding equations above to obtain

$$\begin{aligned}
0 &= \alpha\sigma_0(\mu_{j_1} - \mu_{j_2}) + \alpha\sum_{k=1}^{J}\sigma_k(\mu_{j_1} - \mu_{j_2}) \\
&= \alpha\sigma_0(\mu_{j_1} - \mu_{j_2}) + \alpha(1-\sigma_0)(\mu_{j_1} - \mu_{j_2}) \\
&= \alpha(\mu_{j_1} - \mu_{j_2})
\end{aligned}$$

Since $\alpha < 0$ and j_1, j_2 are arbitrary, this proves the constant mark-up property (ii).

The problem thus reduces to finding the optimal mark-up μ, i.e.,

$$\max_{\mu} \mu \sum_k \sigma_k(c+\mu) = \mu(1-\sigma_0(c+\mu))$$

where c is the vector of marginal costs and $c+\mu$ means that μ is added to every coordinate of c. The first order condition of this problem is

$$(1-\sigma_0) - \mu\sigma_0'(c+\mu) = 0$$

Noting that $\sigma_0'(c+\mu) = -\alpha\sigma_0 \sum_k \sigma_k = -\alpha\sigma_0(1-\sigma_0)$, this reduces to[83]

$$\phi(\mu) := (1-\sigma_0(c+\mu))(1+\alpha\mu\sigma_0(c+\mu)) = 0$$

Note that $(1-\sigma_0(c+\mu)) > 0$ for all μ. Define $\psi(\mu) := 1+\alpha\sigma_0(c+\mu)\mu$. Note that $\psi(0) = 1$ and that $\lim_{\mu\to\infty}\psi(\mu) = -\infty$, because $\alpha < 0$. By the Intermediate Value Theorem, a solution to $\phi(\mu) = 0$ exists. Moreover,

$$\psi'(\mu) = \alpha[\sigma_0'(c+\mu)\mu + \sigma_0(c+\mu)] = \alpha\sigma_0(c+\mu)[1-\alpha\mu(1-\sigma_0(c+\mu))] < 0$$

for all $\mu \geq 0$. It follows that there's a unique μ^* such that $\phi(\mu^*) = 0$ and this μ^* maximizes the monopolist's profit. This proves parts (i) and (iii).

It remains to prove (iv). So far I have shown that the optimal prices satisfy $p_j^* = c_j + \mu^*(c)$, where I make explicit the dependence of μ on the parameter c. It

[83] Note that imposing the constant mark-up property on equation (2.13) also yields $1+\alpha\mu\sigma_0 = 0$. The argument would then deliver the existence of a unique μ^* satisfying the first order conditions of the original problem, but would not imply that the resulting price vector is indeed a solution for that problem. It would then be necessary to establish the validity of a second order condition. The argument given here, which reduces the profit maximization problem to a unidimensional problem, yields existence and uniqueness.

follows that

$$\frac{\partial p_j^*}{\partial c_k} = \begin{cases} 1 + \frac{\partial \mu}{\partial c_j}(c) & \text{if } k = j \\ \frac{\partial \mu}{\partial c_k}(c) & k \neq j \end{cases} \tag{2.14}$$

As shown above, $\mu^*(c)$ is the unique solution to

$$f(c, \mu) := 1 + \alpha\mu\sigma_0(c + \mu) = 0$$

Since $\frac{\partial f}{\partial \mu}(c, \mu) = \alpha\sigma_0(c+\mu)[1-\alpha\mu(1-\sigma_0(c+\mu))] < 0$, the Implicit Function Theorem implies that $\mu^*(c)$ is C^1 and that

$$\frac{\partial \mu}{\partial c_k}(c) = -\frac{\partial f}{\partial c_k}(c, \mu) \Big/ \frac{\partial f}{\partial \mu}(c, \mu)$$

Noting that $\frac{\partial f}{\partial c_k}(c, \mu) = -\alpha^2\mu\sigma_k(c + \mu)\sigma_0(c + \mu)$, I obtain

$$\frac{\partial \mu}{\partial c_k}(c) = \frac{\alpha\mu\sigma_k(c + \mu)}{1 - \alpha\mu(1 - \sigma_0(c + \mu))}$$

as desired. ♠

I now establish an auxiliary result used in the proof of Proposition 2.

Lemma 1. *Let the value for a store under wholesale prices* w *be given by*

$$V^s(w; \mathcal{J}^s) = \mathbb{E}_\eta\left[\sum_{j\in\mathcal{J}^s}(p_j^s(\tilde{c}^s(w)) - \tilde{c}_j^s(w_j))D_j^s(p^s(\tilde{c}^s(w)))\right]$$

Then

$$\frac{\partial V^s}{\partial w_j}(w) = -\mathbb{E}_\eta\left[D_j^s(p^s(\tilde{c}^s))\right]$$

Proof. For a given marginal cost vector $\tilde{c}^s$, let the value of the store's profit maximization problem be $\pi^s(p^s(\tilde{c}^s); \tilde{c}^s)$ - where $p^s(\tilde{c}^s)$ is the solution to that problem.

Exchanging the order of differentiation and integration yields

$$\frac{\partial V^s}{\partial w_j}(w) = \mathbb{E}_\eta\left[\frac{\partial}{\partial w_j}\sum_{j\in\mathcal{J}^s}(p^s_j(\tilde{c}^s(w)) - \tilde{c}^s_j(w_j))D^s_j(p^s(\tilde{c}^s(w)))\right]$$

The derivative inside the expectation operator is $\frac{\partial}{\partial w_j}\pi^s(p^s(\tilde{c}^s);\tilde{c}^s)$. Because $\tilde{c}^s_j = w_j+\tau^s+\eta^s_j$, $\frac{\partial}{\partial w_j}\pi^s(p^s(\tilde{c}^s);\tilde{c}^s) = \frac{\mathrm{d}}{\mathrm{d}c_j}\pi^s(p^s(\tilde{c}^s);\tilde{c}^s)$, where $\frac{\mathrm{d}}{\mathrm{d}c_j}$ denotes the total derivative with respect to c_j. By the Envelope Theorem,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}c_j}\pi^s(p^s(\tilde{c}^s);\tilde{c}^s) &= \frac{\partial}{\partial c_j}\pi^s(p^s(\tilde{c}^s);\tilde{c}^s)\\ &= -D^s_j(p^s(\tilde{c}^s(w)))\end{aligned}$$

Taking the expectation with respect to η yields the result. ♠

An immediate corollary to Lemma 1 is given below

Corollary 1.

$$\frac{\partial V_h}{\partial w_j}(w) = -\mathbb{E}_\eta\left[\sum_{j=1}^{S^h} D^{s,h}_j(p^s(\tilde{c}^s))\right]$$

where $D^{s,h}_j$ is equal to D^s_j if $j\in\mathcal{J}^s$ and zero otherwise.

Proof. It follows from Lemma 1 and the definition of V_h. ♠

Corollary 1 is used in the proof of Proposition 2 below.

Proof of Proposition 2 .

Proof. Consider the maximization problem in (2.2). After taking the logarithm of the objective function, the first order condition with respect to $\hat{w}_j$ is given by

$$b_{m,h}\frac{\partial V_m}{\partial \hat{w}_j}(w)\frac{1}{V_m(w)}+b_{h,m}\frac{\partial V_h}{\partial \hat{w}_j}(w)\frac{1}{S_h(w)}=0$$

where I write $V_h(w)$ instead of $V_h(w;\mathcal{J}_h)$ to simplify notation and m is the manufacturer that produces product j (denoted $m(j)$ below). Rearranging,

$$\frac{\partial V_m}{\partial \hat{w}_j}(w)=-\frac{b_{h,m}}{b_{m,h}\;S_h(w)}\;\frac{\partial V_h}{\partial \hat{w}_j}(w)\;V_m(w) \tag{2.15}$$

Now note that

$$\begin{aligned}\frac{\partial V_m}{\partial \hat{w}_j}(w)=&\sum_{\{s:j\in\mathcal{J}^s\}}\;\sum_{k\in\mathcal{J}^s\cap\mathcal{J}_{h,m(j)}}(w_k-c_k^m)\mathbb{E}_\eta\left[\nabla D_k^s(p^s(\tilde{c}^s(w)))'\frac{\partial p^s}{\partial c_j}(\tilde{c}^s(w))\right]\\&+\sum_{\{s:j\in\mathcal{J}^s\}}\mathbb{E}_\eta[D_j^s(p^s(\tilde{c}(w)))]\end{aligned}$$

where

$$\nabla D_k^s(p)'=\begin{pmatrix}\frac{\partial D_k^s}{\partial p_1}(p) & \dots & \frac{\partial D_k^s}{\partial p_{J^s}}(p)\end{pmatrix}\;,\;\frac{\partial p^s}{\partial c_j}(c)=\begin{pmatrix}\frac{\partial p_1^s}{\partial c_j}(c) & \dots & \frac{\partial p_{J^s}^s}{\partial c_j}(c)\end{pmatrix}'$$

Plugging this and the definition of $V_m(w)$ into equation (2.15) yields

$$\begin{aligned}\sum_{s=1}^{S_h}\Omega_j^s(w)(w_h-c_h)+\sum_{s=1}^{S_h}\mathbb{E}_\eta[D_j^{s,h}(p^s(\tilde{c}(w)))]=&-\frac{b_{h,m}}{b_{m,h}\;S_h(w)}\;\frac{\partial V_h}{\partial \hat{w}_j}(w)\times\\&\times\sum_{s=1}^{S_h}\mathbb{E}_\eta\left[\bar{D}_j^s(p^s(\tilde{c}^s(\hat{w}_m,w_{-m})))\right]\cdot(w_h-c_h)\end{aligned}$$

where $D_j^{s,h}(p)$ is equal to $D_j^s(p)$ if store s sells product j and it is equal to zero otherwise. Moreover, $\Omega_j^s(w),\bar{D}_j^s(p)\in\mathbb{R}^{|\mathcal{J}_{h,B}|}$ are given by (the extra argument in parenthesis denotes the coordinate)

$$\Omega_j^s(w)(k) = \begin{cases} \mathbb{E}_\eta\left[\nabla D_k^s(p^s(\tilde{c}^s(w)))' \frac{\partial p^s}{\partial c_j}(\tilde{c}^s(w))\right] & \text{if } j \in \mathcal{J}^s, k \in \mathcal{J}^s \cap \mathcal{J}_{h,m(j)} \\ 0 & \text{otherwise} \end{cases}$$

and

$$\bar{D}_j^s(p)(k) = \begin{cases} D_k^s(p) & \text{if } k \in \mathcal{J}^s \cap \mathcal{J}_{h,m(j)} \\ 0 & \text{otherwise} \end{cases}$$

Using corollary 1, the last equation can be rewritten as

$$\sum_{s=1}^{S_h} \Omega_j^s(w)(w_h - c_h) + \sum_{s=1}^{S_h} \mathbb{E}_\eta[D_j^{s,h}(p^s(\tilde{c}(w)))] = \frac{b_{h,m}}{b_{m,h}\ S_h(w)} \sum_{s=1}^{S_h} \mathbb{E}_\eta[D_j^{s,h}(p^s(\tilde{c}^s(w)))] \times$$
$$\times \sum_{s=1}^{S_h} \mathbb{E}_\eta\left[\bar{D}_j^s(p^s(\tilde{c}^s(\hat{w}_m, w_{-m})))\right] \cdot (w_h - c_h)$$

This equation holds for each $j \in \mathcal{J}_{h,B}$. Stacking these equations then yields

$$\sum_{s=1}^{S_h} \Omega^s(w)(w_h - c_h) + \sum_{s=1}^{S_h} \mathbb{E}_\eta[D^{s,h}(p^s(\tilde{c}(w)))] = -\sum_{s=1}^{S_h} \Lambda^s(w)(w_h - c_h)$$

where $\Omega^s(w)$ and $\Lambda^s(w)$ are the matrices defined in the statement of the proposition. Rearranging gives

$$\left(\sum_{s=1}^{S_h} \Omega^s(w) + \Lambda^s(w)\right)(w_h - c_h) = -\sum_{s=1}^{S_h} \mathbb{E}_\eta[D^{s,h}(p^s(\tilde{c}(w)))]$$

as we wanted to show. ♠

2.B Characterizing Retail Price Changes with Respect to Changes in Costs

This section characterizes how stores' optimal prices change in response to changes in marginal costs.[84] The results of this section are used in section 2.4.

Given a product portfolio $\mathcal{J}^s$ and marginal costs $(c_1, \ldots, c_J)$, a store sets prices to solve

$$\max_p \sum_{j=1}^{J} (p_j - c_j) D_j(p)$$

The first order conditions are

$$\sum_k (p_k - c_k) \frac{\partial D_k}{\partial p_j}(p) + D_j(p) = 0, \quad j = 1, \ldots, J$$

Stacking these equations yields

$$J_D'(p)(p - c) + D(p) = 0 \tag{2.16}$$

where J_f is the jacobian of the function f. Equation (2.16) implicitly defines p as a function of c. I'm interested in characterizing $J_p(c)$. Define $H(p, c) := J_D'(p)(p - c) + D(p)$, which is the left hand side of (2.16). By the Implicit Function Theorem,

$$J_p(c) = -H_p(p, c)^{-1} H_c(p, c) \tag{2.17}$$

where H_p denotes the matrix of partial derivatives of H with respect to prices and H_c is similarly defined. From (2.16), $H_c(p, c) = -J_D'(p)$.

[84] Note that Proposition 1 subsumes the results in this section. However, the results reported in this version of the paper use the characterization in this section instead of Proposition 1.

Now note that

$$\frac{\partial H_j}{\partial p_l}(p,c) = \sum_k (p_k - c_k) \frac{\partial^2 D_k}{\partial p_l \partial p_j}(p) + \frac{\partial D_l}{\partial p_j}(p) + \frac{\partial D_j}{\partial p_l}(p)$$

and therefore

$$H_p(p,c) = \sum_k (p_k - c_k) \frac{\partial^2 D_k}{\partial p \partial p'}(p) + J_D(p)' + J_D(p)$$

where $\frac{\partial D_k}{\partial p \partial p'}(p)$ is the Hessian matrix of the demand for good k evaluated at p. This result and $H_c(p,c)$ can now be plugged in equation (2.17) to obtain $J_p(c)$. Computationally, I treat (2.17) as a collection of J systems of linear equations. The solution to the $j-$th system of equations delivers the $j-th$ column of $J_p(c)$.

Bibliography

Adams, Brian and Kevin R Williams (2019). "Zone pricing in retail oligopoly". In: *American Economic Journal: Microeconomics* 11.1, pp. 124–56.

Aguirregabiria, Victor and Pedro Mira (2007). "Sequential estimation of dynamic discrete games". In: *Econometrica* 75.1, pp. 1–53.

Aguirregabiria, Victor and Aviv Nevo (2013). "Recent developments in empirical IO: Dynamic demand and dynamic games". In: *Advances in Economics and Econometrics* 3, pp. 53–122.

Aker, Jenny C and Isaac M Mbiti (2010). "Mobile phones and economic development in Africa". In: *Journal of economic Perspectives* 24.3, pp. 207–32.

Akerman, Anders, Ingvil Gaarder, and Magne Mogstad (2015). "The skill complementarity of broadband internet". In: *The Quarterly Journal of Economics* 130.4, pp. 1781–1824.

Armstrong, Mark (2001). "Access pricing, bypass, and universal service". In: *American Economic Review* 91.2, pp. 297–301.

Bajari, Patrick, C Lanier Benkard, and Jonathan Levin (2007). "Estimating dynamic models of imperfect competition". In: *Econometrica* 75.5, pp. 1331–1370.

Berry, Steven (1994). "Estimating discrete-choice models of product differentiation". In: *The RAND Journal of Economics*, pp. 242–262.

Berry, Steven and Joel Waldfogel (1999). “Free entry and social inefficiency in radio broadcasting”. In: *The RAND Journal of Economics* 30.3, pp. 397–420.

Bertschek, Irene and Thomas Niebel (2016). “Mobile and more productive? Firm-level evidence on the productivity effects of mobile internet use”. In: *Telecommunications Policy* 40.9, pp. 888–898.

Björkegren, Daniel (2019). “The adoption of network goods: Evidence from the spread of mobile phones in Rwanda”. In: *The Review of Economic Studies* 86.3, pp. 1033–1060.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa (2017). “Discretizing unobserved heterogeneity”. In: *University of Chicago, Becker Friedman Institute for Economics Working Paper* 2019-16.

Bonhomme, Stéphane and Elena Manresa (2015). “Grouped patterns of heterogeneity in panel data”. In: *Econometrica* 83.3, pp. 1147–1184.

Bronnenberg, Bart J, Michael W Kruger, and Carl F Mela (2008). “Database paper—The IRI marketing data set”. In: *Marketing science* 27.4, pp. 745–748.

Cardell, N Scott (1997). “Variance components structures for the extreme-value and logistic distributions with application to models of heterogeneity”. In: *Econometric Theory*, pp. 185–213.

Cheng, Xu, Frank Schorfheide, and Peng Shao (2019). *Clustering for Multi-Dimensional Heterogeneity.*

Choné, Philippe, Laurent Flochel, and Anne Perrot (2002). “Allocating and funding universal service obligations in a competitive market”. In: *International Journal of Industrial Organization* 20.9, pp. 1247–1276.

Collard-Wexler, Allan (2013). “Demand fluctuations in the ready-mix concrete industry”. In: *Econometrica* 81.3, pp. 1003–1037.

Collard-Wexler, Allan, Gautam Gowrisankaran, and Robin S Lee (2019). ""Nash-in-Nash" bargaining: A microfoundation for applied work". In: *Journal of Political Economy* 127.1, pp. 163–195.

Crawford, Gregory S and Ali Yurukoglu (2012). "The welfare effects of bundling in multichannel television markets". In: *American Economic Review* 102.2, pp. 643–85.

Czernich, Nina et al. (2011). "Broadband infrastructure and economic growth". In: *The Economic Journal* 121.552, pp. 505–532.

Dearing, Adam and Jason R Blevins (2019). "Efficient and Convergent Sequential Pseudo-Likelihood Estimation of Dynamic Discrete Games". In: *arXiv preprint arXiv:1912.10488*.

DellaVigna, Stefano and Matthew Gentzkow (2019). "Uniform pricing in U.S. retail chains". In: *The Quarterly Journal of Economics* 134.4, pp. 2011–2084.

Doraszelski, Ulrich and Juan F Escobar (2010). "A theory of regular Markov perfect equilibria in dynamic stochastic games: Genericity, stability, and purification". In: *Theoretical Economics* 5.3, pp. 369–402.

Draganska, Michaela, Daniel Klapper, and Sofia Villas-Boas (2010). "A larger slice or a larger pie? An empirical investigation of bargaining power in the distribution channel". In: *Marketing Science* 29.1, pp. 57–74.

Dunne, Timothy et al. (2013). "Entry, exit, and the determinants of market structure". In: *The RAND Journal of Economics* 44.3, pp. 462–487.

Ellickson, Paul B, Pianpian Kong, and Mitchell J Lovett (2018). "Private labels and retailer profitability: Bilateral bargaining in the grocery channel". In: *Available at SSRN 3045372*.

Ericson, Richard and Ariel Pakes (1995). "Markov-perfect industry dynamics: A framework for empirical work". In: *The Review of economic studies* 62.1, pp. 53–82.

Gowrisankaran, Gautam, Claudio Lucarelli, et al. (2011). "Government policy and the dynamics of market structure: Evidence from Critical Access Hospitals". In: *manuscript. University of Arizona*.

Gowrisankaran, Gautam, Aviv Nevo, and Robert Town (2015). "Mergers when prices are negotiated: Evidence from the hospital industry". In: *American Economic Review* 105.1, pp. 172–203.

Grennan, Matthew (2013). "Price discrimination and bargaining: Empirical evidence from medical devices". In: *American Economic Review* 103.1, pp. 145–77.

Hausman, Jerry A (1996). "Valuation of new goods under perfect and imperfect competition". In: *The economics of new goods*. University of Chicago Press, pp. 207–248.

Ho, Kate and Robin S Lee (2017a). "Equilibrium provider networks: bargaining and exclusion in health care markets". In:

— (2017b). "Insurer competition in health care markets". In: *Econometrica* 85.2, pp. 379–417.

Horn, Henrick and Asher Wolinsky (1988). "Bilateral monopolies and incentives for merger". In: *The RAND Journal of Economics*, pp. 408–419.

Igami, Mitsuru (2017). "Estimating the innovator's dilemma: Structural analysis of creative destruction in the hard disk drive industry, 1981–1998". In: *Journal of Political Economy* 125.3, pp. 798–847.

Jack, William and Tavneet Suri (2014). "Risk sharing and transactions costs: Evidence from Kenya's mobile money revolution". In: *American Economic Review* 104.1, pp. 183–223.

Jensen, Robert (2007). “The digital provide: Information (technology), market performance, and welfare in the South Indian fisheries sector”. In: *The Quarterly Journal of Economics* 122.3, pp. 879–924.

Kaelo, P and MM Ali (2006). “Some variants of the controlled random search algorithm for global optimization”. In: *Journal of optimization theory and applications* 130.2, pp. 253–264.

Mills, David E (1995). “Why retailers sell private labels”. In: *Journal of Economics & Management Strategy* 4.3, pp. 509–528.

Morton, Fiona Scott and Florian Zettelmeyer (2004). “The strategic positioning of store brands in retailer–manufacturer negotiations”. In: *Review of industrial organization* 24.2, pp. 161–194.

Nevo, Aviv (2001). “Measuring market power in the ready-to-eat cereal industry”. In: *Econometrica* 69.2, pp. 307–342.

Newey, Whitney K and Daniel McFadden (1994). “Large sample estimation and hypothesis testing”. In: *Handbook of econometrics* 4, pp. 2111–2245.

Nocke, Volker and Nicolas Schutz (2018). “Multiproduct-Firm Oligopoly: An Aggregative Games Approach”. In: *Econometrica* 86.2, pp. 523–557.

Pakes, Ariel, Michael Ostrovsky, and Steven Berry (2007). “Simple estimators for the parameters of discrete dynamic games (with entry/exit examples)”. In: *the RAND Journal of Economics* 38.2, pp. 373–399.

Pesendorfer, Martin and Philipp Schmidt-Dengler (2008). “Asymptotic least squares estimators for dynamic games”. In: *The Review of Economic Studies* 75.3, pp. 901–928.

— (2010). “Sequential estimation of dynamic discrete games: A comment”. In: *Econometrica* 78.2, pp. 833–842.

Powell, Michael JD (1994). "A direct search optimization method that models the objective and constraint functions by linear interpolation". In: *Advances in optimization and numerical analysis*. Springer, pp. 51–67.

Roller, Lars-Hendrik and Leonard Waverman (2001). "Telecommunications infrastructure and economic development: A simultaneous approach". In: *American economic review* 91.4, pp. 909–923.

Ryan, Stephen P (2012). "The costs of environmental regulation in a concentrated industry". In: *Econometrica* 80.3, pp. 1019–1061.

Sanches, Fabio, Daniel Silva-Junior, and Sorawoot Srisuma (2018). "Banking privatization and market structure in Brazil: a dynamic structural analysis". In: *The RAND Journal of Economics* 49.4, pp. 936–963.

Schmidt-Dengler, Philipp (2006). "The timing of new technology adoption: The case of MRI". In: *Manuscript, London School of Economics*.

Valletti, Tommaso M, Steffen Hoernig, and Pedro P Barros (2002). "Universal service and entry: The role of uniform pricing and coverage constraints". In: *Journal of Regulatory Economics* 21.2, pp. 169–190.

Villas-Boas, Sofia (2007). "Vertical relationships between manufacturers and retailers: Inference with limited data". In: *The Review of Economic Studies* 74.2, pp. 625–652.

Wu, Tim (2010). *The master switch: The rise and fall of information empires*. Vintage.