Essays in Competition and Externalities

Ilton Gurgel Soares

Submitted in partial fulfillment of the

Requirements for the degree

of Doctor of Philosophy

in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2016

ABSTRACT

Essays in Competition and Externalities

Ilton Gurgel Soares

This dissertation consists of three papers. A common feature of these papers is the interest in how externalities affect consumers and firms' behavior. In the first paper, I study one type of contractual externalities called exclusive dealing, whereby one firm cannot deal with the competitors of the other. More specifically, I propose and estimate an empirical structural model to investigate the effects on prices of upstream mergers in markets with exclusive dealing contracts. The second paper is concerned with markets for a good with network externalities, i.e. a good that generates higher utility the higher the number of consumers purchasing it. The third paper studies externalities of investments on quality improvement. When more than one firm is active, the product improvement externality occurs because as firms chose different quality levels, competition is relaxed and consumers get some consumer surplus from product variety. In the case of winner-take-all markets, the business-stealing externality occurs because as one firm invests in quality upgrade, the competitors become more likely to lose all customers.

The first chapter examines the incentives for price increase in upstream mergers when the supplier has a network of exclusive dealers (ED).[1] The incentives explored in this paper come from changes in the threat point of the bargaining between the supplier and exclusive retailers. The bargaining power of the exclusive dealer comes from local market power of

[1] An ED contract between two firms determines that one cannot deal with the competitors of the other. These contracts are common in a variety of industries such as beer, fuel and automobiles. Franchising relationships usually involve some type of ED agreement since the franchisee typically has to purchase at least some of the goods or services exclusively from the franchisor.

the dealer or due to reputation aspects (when dealers know that the supplier behaves opportunistically after the ED contract is signed, they will be reluctant in becoming exclusive of that supplier or renewing the contract). The change in the threat point post merger is due to the larger network of exclusive retailers, which enables the merged supplier to recapture a larger portion of the consumers that will be diverted from any specific exclusive dealer in case of disagreement on the wholesale price negotiation. The empirical application explored in this paper uses a unique and comprehensive dataset from the Brazilian fuel industry, with information that includes retail and wholesale prices as well as quantities at the station level. Aside from the good quality, this dataset is adequate for the intended analysis because in Brazil fuel stations can either operate independently (in which case they can purchase from any distributor) or sign an ED contract, when they can only purchase from a specific distributor. Moreover, the data spam a period that includes an important merger. I estimate the model using pre-merger data and simulate the effects of combining the networks of exclusive dealers of the merging companies. The simulation shows that the incentives for price increase are sizable, and the mechanism studied in the paper captures a large fraction of the actual price increase observed in the data.

The second chapter, joint with Ilwoo Hwang, studies adoption and pricing when consumers can delay their purchase of a good with network effects. In those cases, price alone does not convey sufficient information for consumers to make their purchase decision and they need to infer about current and future adoption in order to make their decisions. This feature implies that some consumers might find optimal to delay their purchases in order to make their decisions better informed about the success of the network. The multiplicity of equilibria that is typical in the coordination game played by consumers implies that the demand is not well defined for a given price, creating a problem for the firm's pricing decision. We consider a two-period model in which a monopolist sets prices and consumers can delay their purchases to the second period when they will receive information about early adoption. The dynamic coordination problem with endogenous delayed purchases is modeled as a global game, for which we derive conditions for uniqueness of equilibrium. The model is capable of exploring many issues in the economics of network effects such as introductory pricing and early critical mass for platform survival. Our specification nests the pure durables goods and herding models. Numerical results illustrate the amplitude of possible outcomes in the dynamic model with delay. Substantial differences can arise in terms of pricing, adoption and profits when we compare the full specification with multiple benchmarks.

In the third chapter, joint with Michael Riordan, we develop a duopoly model of product quality competition that focuses on how information structure determines equilibrium outcomes. When we introduce private and correlated signals about the fundamental uncertainty about quality differences, each firm can form a more educated guess about what the opponent must be doing, which is the key for uniqueness of equilibria. Equilibrium product improvement decisions are unique if and only if market uncertainty is sufficiently high relative to strategic uncertainty, except in a non-generic special case. A unique equilibrium takes the form of threshold strategies, whereby each firm improves its product upon receiving a sufficiently favorable signal of brand advantage. We show that the unique equilibrium depends on the fundamentals as well as on investment costs and that the probability of miscoordination vanishes as strategic uncertainty decreases. In the type of competition studied here, firms have no incentive to choose the same quality as the competition arising in the marketplace would bring prices to equalize marginal cost. Interestingly, this information structure alleviates substantially the problem of miscoordination observed in the no "information game" and also dominates the complete information game for a large range of parameters in the model.

Contents

List of Tables

List of Figures

Acknowledgements

I am greatly indebted to my advisors Michael Riordan and Katherine Ho for their invaluable support and guidance throughout my PhD. I feel very fortunate to have had the opportunity to work under their supervision. I am also greatly indebted to Chris Conlon, for the numerous conversations, guidance and suggestions, especially on the first chapter of this dissertation. I thank the other members of my committee, Scott Shriver and Fanyin Zheng, for their comments and suggestions.

I also thank my colleagues and professors for their helpful comments: Ildiko Magyiari, Antonio Miscio, Patrick Sun, Joe Hogan, Xing Xia, Samer Shousha, Miguel Morin, Ilwoo Hwang, Naoki Aizawa, Bernard Salanie, Jose Scheinkman, Andrea Pratt, Matt Backus, Yeon-Koo Che, Jushan Bai, Alireza Tahbaz-Salehi, and Francesco Decarolis.

I gratefully acknowledge *Agencia Nacional do Petroleo (ANP)* and *Ministerio do Trabalho e Emprego (MTE)* for providing the data used in the first chapter.

Finally, I will always be grateful to Marcelo Moreira and Flavio Cunha for their help and academic guidance earlier in my PhD.

To Munique and Lucas

Chapter 1

Mergers Under Exclusive Dealing: An Empirical Analysis of the Fuel Industry

Horizontal mergers (that is, mergers between competitors) frequently take place not only among firms selling products directly to consumers, but also among firms in producer markets. These mergers have the potential to affect competition not only in the upstream (input) market, but in the downstream (retail) market as well. Standard merger analysis focuses on the exploitation of market power in order to raise prices. However, many markets are characterized by a vertical structure where upstream firms act as the only supplier of a particular retailer[1]. Under these exclusive dealing (ED) contracts, input suppliers do not

[1] In the beer industry for instance, some brewers sign ED contracts with distributors, granting the distributor exclusivity within a certain territory. See more about the beer industry in Asker (2004) and Sass (2005). In the fuel industry, stations can either operate independently and be free to purchase fuel from any distributor/refinery or they can sign an ED contract with a given distributor/refinery (e.g. Shell or Texaco), carry the logo of the distributor in the station and commit to purchase only from that supplier. For more on the fuel industry see Hastings (2004), Gilbert and Hastings (2005), Manuszak (2010) and Houde (2012). In the automobile industry, car dealers can either operate as multi-brand (independent) stores or under an ED with some specific manufacturer such as Volkswagen or GM for instance. See Nurski and Verboven (2013) for more on ED in the automobile industry. Other examples of industries with ED

compete to sell to their exclusive retailers. In that case, analysis focusing on the input market seems unlikely to predict any effect of horizontal mergers among upstream firms[2].

In this paper I build and estimate a model to study how mergers between suppliers in markets with ED contracts can affect wholesale and retail prices as well as consumer welfare. Upstream mergers[3] in markets with ED have an important difference relative to other types of horizontal mergers: post-merger, not only will the merging firms form a single entity, but the networks of exclusive dealers will also be combined. I show that, with the combined network, the merged supplier is in a better position to negotiate higher wholesale prices with its exclusive retailers, which will ultimately affect the prices paid by consumers. The merged supplier can strike a better deal because, as a result of the consolidation of the networks, it can absorb a larger part of the diverted sales in the case of a hypothetical disagreement with an exclusive dealer. This creates incentives for the merged supplier to increase wholesale prices in its exclusive network. However, the price increase depends on the structure of the downstream competition and need not be uniform.

Figure 1 illustrates how ED between suppliers and retailers can create incentives for price increase after a merger. U_j represents the suppliers (upstream firms) and r_i the retailers (downstream firms). ED contracts are represented by links between upstream and downstream firms. Pre-merger, upstream firm U_1 has an ED contract with retailer r_1 and U_2 has ED contracts with r_2 and r_3. The new firm resulting from the merger of U_1

---

contracts include smartphones (Sinkinson, 2014) and video games (Lee, 2013). Heide, Dutta and Bergen (1998) consider a sample of 147 manufacturers in industrial machinery, electronic and electric equipment industries and point that over 40 percent of those had ED arrangements with distributors.

[2]To some extent, a similar logic was used by the Senior Vice President and Chief Diversity Officer of Comcast, David L. Cohen, who pointed that the merger with Time Warner Cable "will not lead to any reduction in competition or consumer choice in any market. Our companies serve separate and distinct geographic areas. We don't compete for customers anywhere." The main difference between the Comcast-TWC merger and the type of mergers that I study here is that I am interested in mergers between companies that don't serve directly the final consumer.

[3]In what follows I will use the terms "mergers between suppliers", "merger in producer markets" and "upstream mergers" interchangeably.

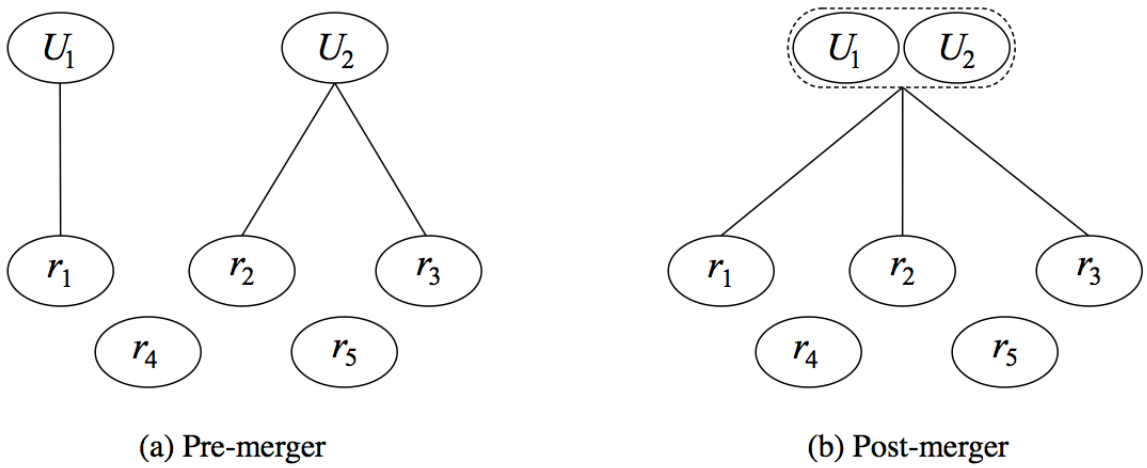


Figure 1.1: Hypothetical market with 5 retailers, 3 of which have exclusive dealing agreements. Retailers r_4 and r_5 can either belong to other exclusivity networks or operate as independent retailers, which are free to purchase from any supplier (some of which are not included in this picture). Figure (a) considers the pre-merger case and Figure (b) depicts the post-merger case, where the networks of exclusive retailers of suppliers U_1 and U_2 are combined.

and U_2 will have a network of exclusive dealers consisting of the combination of the two networks. Pre-merger, in case of a hypothetical disagreement between U_1 and r_1, some of r_1's customers are diverted to r_1's competitors. Therefore, from U_1's perspective all of r_1's potential sales are lost. However, post-merger, in the case of the same hypothetical disagreement, the merged supplier would still be able to capture the share of r_1's sales that are diverted to r_2 and r_3, which now belong to U_1's consolidated network. When negotiating wholesale prices, the merged upstream firm will take the value of the diverted sales into consideration as an opportunity cost. Hence, the incentive for the merged supplier to raise wholesale prices originates as a response to its ability (post-merger) to absorb a larger part of the diverted sales of any member of the network in case of a disagreement. [4]

[4]It is key to take ED into consideration in order to capture the incentives for price increase. Without ED, each retailer is free to purchase from any supplier and wholesale prices are determined in a competitive way. For concreteness, consider the case of fuel. When wholesale prices can be different for the various independent retailers, they can be thought as being determined by a mechanism similar to a price quote or

My model combines three components to capture the strategic interactions involving suppliers and retailers in a market characterized by ED. The first component describes the vertical negotiations over wholesale prices between suppliers and exclusive dealers. Following Horn and Wolinsky (1988) and the empirical literature on bargaining[5], I assume that wholesale prices are determined as a solution to the Nash bargaining problem conditional on all other prices. The second component models retail price competition accounting for the importance of geographic differentiation. The third component is the individual consumer's demand, which builds on Berry, Levinsohn and Pakes (1995) to estimate price sensitivity and transportation costs for consumers. The demand model is tailored to the application in the Brazilian fuel industry, which requires accounting for the consumption of both gasoline and ethanol due to the popularity of flex fuel vehicles in that country. I assume that consumers are located in the path defined by their commuting behavior as in Houde (2012). Throughout the analysis, I take the network of exclusive dealers as given.

The difficulty in obtaining data on supply arrangements is perhaps the main reason why the literature on ED is still remarkably limited. In this paper I construct a rich panel data on the Brazilian fuel industry combining different sources. The data contains detailed information about vertical transactions including retail and wholesale prices as well as volumes at the station level. Additional information at the station level that I observe includes location, brand affiliation, number of attendants and ancillary services.

In the estimation, I use data only from a period that precedes an actual merger between two large suppliers, which combined had ED agreements with nearly 20% of the retailers

procurement auction. In that case, a merger between two distributors would reduce the number of bidders, increasing the expected value of the wholesale price (winning bid). When the number of suppliers is large, this effect can be negligible. In that case, the merger effects on wholesale prices will be driven basically by the efficiency gains reduce wholesale prices. This logic illustrates that the ability of the suppliers to charge prices considerably higher than marginal costs in this industry is a consequence of the ED contracts.

[5]E.g. Crawford and Yurukoglu (2012), Ho and Lee (2015), Crawford, Lee, Whinston and Yurukoglu (2015).

in the country. The estimation was conducted in the metropolitan area of Vitoria, which is the capital of the state of Espirito Santo. This state was under suspicion by the antitrust authority for being the one with the largest combined market shares per merger in the country (around 27%).[6] The estimated parameters and model are then used to simulate the effects of the merger. The combination of the networks of exclusive dealers will induce a new set of equations characterizing the equilibrium wholesale prices. These new wholesale prices are then used in the equilibrium condition for the retail pricing to obtain predicted consumers' prices for both types of fuel. The simulated wholesale and retail prices are the fixed points of this interaction between the equilibrium conditions from the retail pricing and vertical negotiation. The data span the periods pre and post merger, providing the opportunity to observe the realized prices at the time of the merger and conduct a retrospective analysis.

I find that the bargaining weight of the major distributors varies between 0.52 and 0.60, significantly smaller than unity, which is the "take it or leave it" value. On average, the model predicts a wholesale price increase of 4.8 cents per liter (cpl) for the merged distributors and 1.3 cpl for the non-merged. These changes correspond to an increase of 30% in the margins of the merging distributors compared to the average margin pre merger. At the retail level, the model predicts 4.3 cpl price increase for the exclusive stations of the merged distributors and 2.7 cpl for the remaining exclusive stations. In addition, I find that the average markup of the retailers is approximately 6%.

Since the data span a period which includes the merger studied, I am able to conduct an ex-post evaluation of the model simulation. Actual data confirms the predicted increase in the wholesale margins, as well as the difference in the wholesale price increase between the merged and non merged firms. The observed increase in the wholesale margins was

[6] See page 6 of the Concentration Act for the merger (08012.001656/2010-01) available at www.cade.gov.br.

even larger than what was predicted in the simulation: the model predicts nearly 60% of the actual average increase in the wholesale margins.

Strategic complementarity at the downstream level implies that the unbranded retailers will eventually increase their retail prices in response to a price increase of the branded retailers. While the model correctly captures this response for the unbranded stations, it does not predict an increase of wholesale prices for independent retailers. The reason is that independent retailers can purchase from any distributor and in that case wholesale prices are determined in a competitive way, depending on the cost structure of the distributors and not on the price charged by the independent retailer.

In terms of the demand estimates I find that the demand for fuel at any given station is very elastic: average price elasticity of 20%. This value is higher than the one predicted by Houde (2012) for the Canadian market (between 10% and 15%) and similar to what is found in Manuszak (2010) using Hawaiian data. One additional reason why station-level price elasticity in the Brazilian market is expected to be high is the coexistence of two types of fuel that are substitutes for a sizable fraction of the consumers (flex fuel vehicle owners).

Another relevant finding in the demand estimation was that consumers value brands. On average, an unbranded (independent) station has to give a discount of approximately 1.5% in order to make consumers indifferent relative to purchasing the fuel in a branded station. The money value estimated disutility of driving is twice as big as the average wage in the country, suggesting that consumers tend not to deviate too much from their paths for buying fuel.

The remainder of this paper proceeds as follows. In Section 2 I introduce the data and industry background. Section 3 presents the empirical model. In Section 4 I discuss identification and estimation. Section 5 presents the results and Section 6 the merger analysis. Section 7 concludes.

1.1 Related Literature and Contributions

This is the first empirical paper to study the effects of upstream mergers in markets with ED. It builds on and contributes to three related literatures. The first is a small but growing empirical literature on markets with ED agreements. The second is the large literature on horizontal mergers[7]. The third is the literature on vertical and bilateral negotiations, more specifically on structural models of bargaining.

The theoretical literature on ED was motivated in large part by the Chicago school argument that in order for an ED agreement to be mutually beneficial it must be associated to efficiency gains. Aghion and Bolton (1987) show that an incumbent supplier and a retailer can exclude an efficient entrant if the contract includes liquidated damages. Chen and Riordan (2007) show how a vertically integrated firm can use exclusive contracts to exclude an equally or more efficient firm that is already in the market. Segal and Whinston (2000) demonstrate that if the manufacturer offering ED contracts cannot discriminate among retailers, both exclusionary and non-exclusionary equilibria exist. All retailers are worse off, so exclusion will only succeed if retailers cannot coordinate their actions to jointly refuse an exclusionary contract. Fumagalli and Motta (2006) show that when ED is between suppliers and retailers instead of suppliers and final consumers, the coordination problem may not occur. In that case, one single deviant retailer may be able to serve the whole market by buying at a lower price from the entrant, enabling the entrant to cover its fixed costs. Johnson (2014) presents a theory in which ED does not serve to exclude or disadvantage rivals. Instead, ED gives each supplier the ability to internalize competition amongst the retailers in the network. Relative to the case without ED, he shows that in equilibrium retail prices increase, benefiting suppliers and retailers but harming consumers.

The papers in the theoretical literature that are more closely related to mine are Milliou

[7]See chapters 3 and 4 of Whinston (2006) for an excellent survey of the history and recent advances in the horizontal merger and exclusive dealing literatures.

and Petrakis (2007) and Fumagalli, Motta and Persson (2009). Milliou and Petrakis (2007) study horizontal mergers in the upstream sectors with ED when bargaining is present. Their model assumes Cournot competition at the downstream level and two part tariffs, which leads to wholesale prices lower than marginal costs, and even more so post merger. This generates a reduction on the retail prices following the merger, in the absence of any efficiency gains. There are two important differences between my model and Milliou and Petrakis (2007). First, I consider that the bargaining is only over the wholesale prices. The second difference is that in my model downstream competition is assumed to be Bertrand with differentiated products. Fumagalli, Motta and Persson (2009) consider the case of a merger between an incumbent supplier and a potential entrant in a market with ED. They show that the incumbent can use ED contracts to improve its bargaining position in the merger negotiation with the entrant. Instead, in this paper I am interested in mergers between two incumbent firms. I those cases, I show that and the improvement in the bargaining position comes from the merger itself and with respect to the exclusive dealers.

The empirical literature on ED is remarkably limited, in large part due to the difficulty in obtaining adequate data, specially on wholesale prices[8]. Moreover, the existing empirical work on ED has primarily focused on foreclosure. For instance, Asker (2005) tests for foreclosure due to exclusive dealing relationships in beer distribution and finds no significant evidence that exclusive dealing increases market power. Hortacsu and Syverson (2007) find that vertical mergers between cement and concrete producers were, on average, efficiency enhancing, leading to lower intermediate and final good prices and larger quantities. Lee (2013) measures the impact of ED on industry structure and welfare in the video game industry and finds that ED favored the entrant platforms. In contrast, in this paper I focus on how upstream mergers can create incentives for changes on prices under ED and how

[8]In order to circumvent this data limitation, empirical research modeling vertical negotiations has relied on theoretical assumptions to infer wholesaler behavior (e.g. Villas-Boas (2007), Mortimer (2008), Hellerstein (2008), Manuszak (2010)).

this affects competitors, competition and consumers.

This paper is also related to a large literature on horizontal mergers, more specifically to recent contributions on the predictions of merger effects, measurement of the effects of actual mergers and mergers in producer markets. In large part, the literature on horizontal mergers has considered one-tier industries. For instance, Nevo (2000), Pesendorfer (2003) and Houde (2012) consider the case in which merging firms directly set consumer prices[9]. To the best of my knowledge, the only empirical papers on mergers in producers' markets that explicitly consider downstream pricing are Villas-Boas (2007) and Manuszak (2010). There are important differences between this paper and two just mentioned. First, I focus on the case of upstream mergers under ED. Second, none these papers have information about wholesale prices, which imposes some restrictions on the type of behavior that they can allow the upstream firms to have. In practice, both papers assume that the upstream firm charges a single price for all retailers. I observe wholesale prices and can take into consideration the variation in prices and asymmetric incentives to raise prices within the network of exclusive retailers. Finally, I observe prices both before and after the merger which allow me to perform a retrospective analysis of the merger.

The literature on ex-post evaluation of merger simulation is very recent. The motivation for comparing predicted changes from merger simulation with observed prices are to evaluate the accuracy of these forecasts, which can also serve as a test of the assumptions imposed in the underlying model. Peters (2006) uses merger simulation to predict price effects of five airline mergers from the 1980s and compares the predicted prices with observed post-merger prices. Weinberg (2011) studies the effects of mergers on the prices of the merged firms and competitors. Houde (2012) studies spatial competition with an application to a real vertical merger, comparing diff-in-diff and counterfactual simulation

[9]The last Horizontal Merger Guidelines, issued in 2010, do not address any special aspect related to upstream mergers in markets with ED contracts.

methods. Bjornerstedt and Verboven (2013) compare the predictions from a merger simulation in the Swedish market for analgesics with the actual merger effects. One important difference of what I do and the cited papers is that I account for the divisions between downstream and upstream firms and the vertical negotiations between them. All the above mentioned papers assume that the merging firms directly set consumers prices. Moreover, none of those papers is related to ED.

The vertical Gross Upward Pricing Pressure (vGUPPI) proposed by Moresi and Salop (2013) explains how a vertical merger can create unilateral incentives to raise prices. They consider the case in which upstream firms are able to charge different prices from the downstream ones. The vGUPPI is very similar to the GUPPI proposed in the Horizontal Merger Guidelines, with the difference that horizontal diversion ratios between two competitors are replaced by the diversion ratio from the upstream merging firm to the downstream merging partner. In my setup, a horizontal merger with ED has a vertical element because the acquirer is also gaining an ED network from the merger. In this paper I show that the incentives for price increase from mergers under ED also depend on the diversion ratios, but only with respect to the new merged network of exclusive dealers.

Finally, related to the literature on structural bargaining models, this paper closely follows Crawford and Yurukoglu (2012), Crawford, Lee, Whinston and Yurukoglu (2012) and Ho and Lee (2015). The major difference with respect to these papers is that I am interested on the case of a market with ED, which changes the structure of the bargaining. Another closely related paper is Gowrisankaran, Nevo, Town (2014), which studies how hospital mergers can affect vertical negotiations between hospitals and MCOs. The vertical negotiations studied in that paper do not involve exclusivity. Another difference relative to Gowrisankaran, Nevo, Town (2014) is that I account for downstream competition (as in Crawford and Yurukoglu (2012) and Ho and Lee (2015)).

1.2 Industry Background and Data

Demand for gasoline and other fuels are an important component of households' budget and changes in their prices can have substantial effects on consumers' welfare. In the U.S. for example, gasoline spending occupies between 4.5% and 12.4% of households' disposable income (Houde, 2010). Based on data from the Personal Consumption Expenditures by type of product from the Bureau of Economic Analysis, Langer and McRae (2013) note that gasoline is the largest non-durable item for most households. They also point to a Gallup poll on June of 2008, a period of high gasoline prices, where one quarter of the U.S. households reported that these prices were the single most important problem facing the country.

Even though gasoline is a fairly homogeneous product, gasoline at the pump is a differentiated product because of some aspects like location and ancillary services. An additional reason why the fuel industry is not perfectly competitive is due to the dominance of major oil companies, which represents a concern to antitrust authorities.

The fuel industry includes the processes of production, distribution and retailing. Gasoline is produced in the refineries and ethanol is produced in the distilleries. These products[10] are then sent to the distributors, who are in charge of mixing fuels and additives as well as storing, selling and transporting these to the jobbers and retailers. The retailer, which is the only party authorized to sell to individual consumes, can operate under an ED contract with one distributor or independently, when it is free to purchase fuel from any distributor. Petrobras is the main Brazilian oil refinery, producing more than 90% of the total volume of gasoline consumed in the country. Moreover, the refinery price of gasoline is insensitive

[10]Gasoline can be of types A and C. Gasoline A is pure gasoline, produced in the refineries, petrochemistries or imported. Gasoline C can be standard, with additive or premium. The gasoline C standard is a mixture of gasoline type A and ethanol. This mixture is realized by the distributors. Gasoline C with additive is a mixture of Gasoline C standard and additives. These additives contain detergents that help cleaning the engine.

to supply and demand because it is regulated by a public sector entity. All other prices are freely determined in the market, including the the producer prices of ethanol, wholesale price and retail prices.

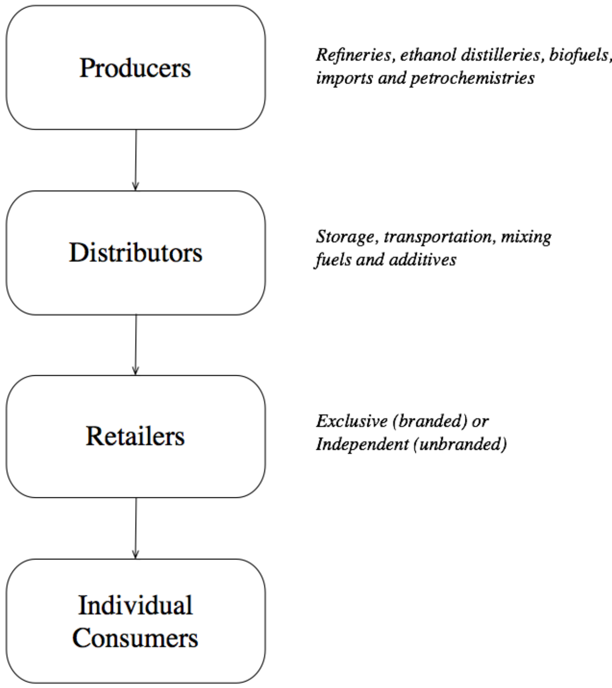


Figure 1.2: Structure of the fuel industry.

Virtually every fuel station in Brazil sells both gasoline and ethanol. There are some features that differentiate gasoline and ethanol and that might affect consumers' choices. First, the calorific value of ethanol is equivalent to around 0.7 of that for gasoline, which implies that for a consumer indifferent between the two types of fuel there is a threshold values of the ratio of prices that would lead to consume one or another. The second difference is that a car running on ethanol is less hazardous to the environment. Third, ethanol has a higher octane rating (110 vs 87-93 on gasoline). Finally, gasoline engine demands less fuel, thus requiring less frequent refueling.

Gasoline sold at the station is a mixture of anhydrous ethanol and gasoline in a proportion that is defined by the regulator and varies between 15% and 25%. Both conventional and flex-fuel cars can use this type of fuel, but the latter category, which has become the dominating passenger car type, can use any blend ratio up to 100 percent hydrous ethanol. The fuel taxes applied to gasoline and ethanol, are modified frequently to make the two fuel types competitive.

The length of the contract between distributors and stations varies depending mostly on the size of the financing, if any, that the station used for renovation or to enter the market. The branded stations also get the support from the distributors in many items such as help with the business plan and structure of the gas station, lease of equipment, advertisements, training for the managers and employees, and marketing promotions (e.g. car raffle). At the end of the exclusivity contract, the retailer is free to switch to a different brand or become independent.

The relation between branded stations and distributors is similar to a franchising agreement, which is potentially very different from vertical integration in terms of incentives. An ED contract can, at least in theory, replicate the effect of vertical integration. In practice, because of limitations arising from transaction costs or legal issues, ED and vertical integration are not equivalent. One example is the possibility of opportunistic behavior that can arise in ED relationships. This is not a problem faced by stations when they are vertically integrated with refineries or distributors.

Houde (2012) notes that between 52% and 72% of branded stations were company-owned in 2001 in the Canadian market. Moreover, in the case of branded stations with ED contracts in that country, wholesale prices are set at the station level in a weekly basis and that "lessee station owners also negotiate a price-support clause that ensures them a minimum profit margin".

1.2.1 Data

The data used in this paper comes from several sources. The main piece of information comes from a detailed survey conducted by the National Petroleum Agency (ANP), the Brazilian regulatory agency of oil and natural gas. Every week since July of 2001, ANP collects data on wholesale and retail prices for gasoline, ethanol and diesel at individual fuel stations in over 500 municipalities in Brazil. In general, between 40% and 50% of the fuel stations are surveyed each week. Coverage reaches 100% in the smaller municipalities. In the larger cities, the survey adopts a rotating sample that eventually covers all stations. The survey provides information about location of stations and distributors as well as brand affiliation and shipping mode (CIF or FOB).

I combine the price data with information about storage capacity of the fuel tanks and number of nozzles for each type of fuel in each station and monthly information about volumes purchased from the distributor for the period between January of 2007 and December of 2011, also provided by ANP. Additionally, I collected data on secondary activities of the station such as existence of car wash, oil change and convenience store from the Department of Federal Revenue of Brazil (*Receita Federal*). A summary of the main characteristics of the stations is displayed in Table 1. Independent retailers are in general competing more aggressively on prices and not so much in terms of additional services. In particular, the number of attendants and nozzles, used to measure the service speed (time spent in the station), is considerably lower in independent stations. Moreover, stations attached to major distributors on average offer a larger variety of ancillary services than independent retailers, with the sole exception of tire repair.

Table 1.1: Average characteristics of exclusive and independent retailers (Vitoria metropolitan area)

Variable	**Major Brands**		**Unbranded**	
	mean	**std. dev.**	**mean**	**std. dev.**
Number of attendants	9.49	5.34	6.13	6.02
Number of nozzles (gasoline)	5.22	1.69	3.47	2.04
Convenience store	0.26	0.41	0.14	0.39
Oil change	0.26	0.35	0.09	0.28
Car wash	0.28	0.38	0.15	0.22
Highway located	0.31	0.41	0.27	0.32
Tire repair	0.08	0.28	0.15	0.36

Major brands include include BR, Ipiranga, Shell and Esso.

Data on prices of ethanol at the producer (distillery) level were obtained from ESALQ. These indices are reported weekly, consisting of information of average prices of hydrous and anhydrous ethanol. Information on taxes on both types of fuel was obtained from ANP and SINDICOM.

Figure 3 illustrates the monthly variation on the wholesale prices (FOB shipping[11]) in the Vitoria Metropolitan Area in 2007. The original price data is at the weekly frequency. For the purpose of estimation, I average both retail and wholesale prices at the monthly level in order to have the same frequency as the volume data. Wholesale prices can vary substantially within the same distributor for different exclusive retailers. This variation in prices within the network is not a feature exclusively of the Brazilian market. In the U.S.,

[11]CIF and FOB are types of shipping agreements and differ in who assumes the expenses and responsibility for the goods during transit. In the case of CIF (Cost, Insurance and Freight) - Insurance and transportation are paid by the seller until the goods are received by the buyer. When shipping if FOB (Free on board), the retailer is responsible for the transportation and all costs once the fuel is picked up at the distributor.

the refiner/distributor can also set different prices for stations within its own network. The Brazilian market also allows price discrimination with respect to the unbranded stations. However, when selling to unbranded stations, the U.S. refiners/distributors must post a rack price that is the same for all purchasers at that rack[12]. In the Appendix A.1. I provide the portion of an ED contract of a major distributor where it is specified that the wholesale prices are "freely agreed between the parties".

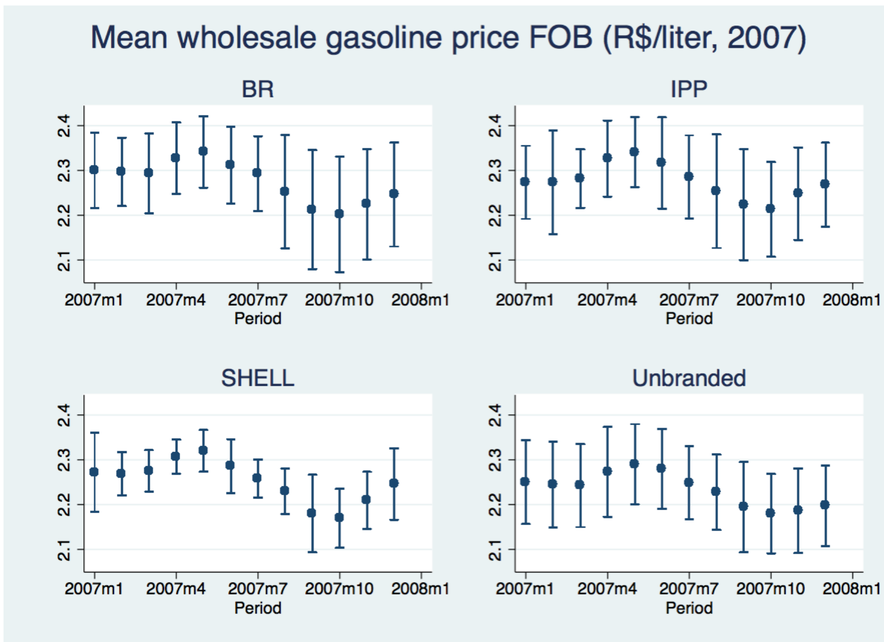


Figure 1.3: Variation of wholesale prices for exclusive retailers of some large distributors and for unbranded retailers

Another source of information that I use is the *Relação Anual de Informações Sociais* (RAIS), a matched employer-employee dataset assembled by the Brazilian Ministry of Labor. The data includes information on the occupation of the worker and I use it to obtain the number of attendants in the stations. The data also includes start and end date on the job, which gives precise information about employment at any point in time.

[12]See Hastings (2010) for more details on the U.S. market.

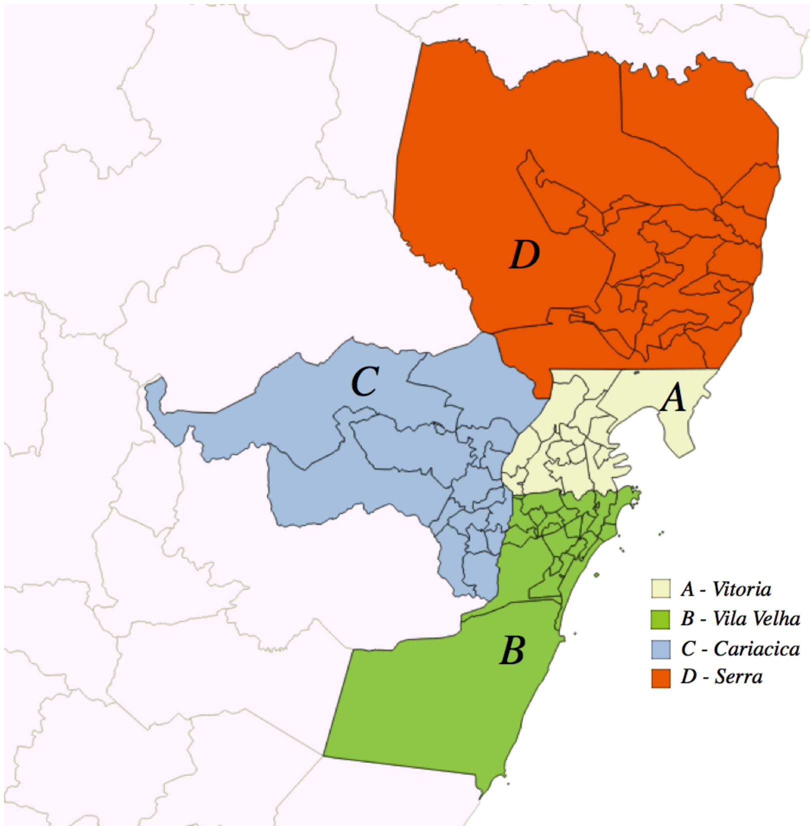


Figure 1.4: Municipalities used in the estimation. The divisions within each municipality are the weighting areas, the smallest level of aggregation available for household location in the Census microdata.

The estimation in this paper is based on data from four municipalities in the state of Espirito Santo: Vitoria (capital), Vila Velha, Cariacica and Serra (See Figure 4). These municipalities are part of the Vitoria Metropolitan Area (VMA) and account for 46.2% of the population in the state of Espirito Santo. The VMA has other three municipalities that were not included in the estimation: Fundao, Viana and Guarapari. The first two because they are not part of the weekly price survey conducted by ANP. The last one because it is substantially different from the other municipalities in terms of consumption of fuel, since this is mostly a vacation destination and fuel consumption is highly seasonal.

I use Census microdata[13] on consumers' home and work locations, as well as commuting time to construct flows within a metropolitan area. To characterize consumer locations, I use the smallest level of aggregation available for household location, the Census weighting area. Each weighting area requires a minimum number of households, contiguity and homogeneity with respect to a certain set of population characteristics and infrastructure. The Census microdata provides information about home location at the level of weighting area. Work location is known only at the level of municipality. In order to construct commuting flows, I combine the information on home and work locations with the commuting time (also included in the Census). The distances between population weighted centroids of weighting areas and retailers were computed in terms of estimated driving time and driving distance using Google maps.

The demand estimation requires a definition of the relevant geographic market for the computation of market shares. This is a complicated task because isolated geographic markets are rare. In order to capture the possible inter relations among the four municipalities, I computed the commuting flows of workers among the four municipalities, displayed in Table 2.

Table 1.2: Commuting flows of workers for the four municipalities considered in the estimation

Origin \ Destin.	***A***	***B***	***C***	***D***	**Total**
***A* (Vitoria)**	78.98%	4.42%	3.45%	13.16%	100%
***B* (Vila Velha)**	34,22%	52,71%	6.13%	6.95%	100%
***C* (Cariacica)**	26.95%	11.86%	50.67%	10.51%	100%
***D* (Serra)**	29.34%	2.32%	1.74%	66.60%	100%

Source: Census microdata, 2010.

[13] I am thankful to Data Zoom, developed by the Department of Economics at PUC-Rio, for providing codes for accessing IBGE microdata.

The diagonal of Table 2 contains information about internal flows, i.e. people that live and work in the same municipality. The off diagonal elements show a substantial flow of workers commuting to a different municipality, remarkably to the capital (Vitoria). For instance, more than 1/3 of the workers living in Vila Velha commute to Vitoria on a daily basis. This is suggestive that stations in Vila Velha are competing with stations in Vitoria, at least for those commuting consumers. Because of the intense flow among the four municipalities, I define the relevant market to be the four municipalities in the Vitoria metropolitan area.

Data on the monthly fleet of vehicles per municipality was obtained from Anfavea. This data was combined with information on the registration of new vehicles by fuel type in order to estimate the fraction of flex vehicles in each municipality. Flex fuel vehicles became commercially available in Brazil in March of 2003, reaching more than 80% of the registration of new vehicles after only three years and near 95% in 2013.

1.3 Empirical Model

Under ED, market power of the upstream firm depends on how it manages competition among the retailers within its network as well as against other retailers. In order to quantify the effects of a merger it is then important to understand how retailers determine price, how consumers choose among the variety of options available and characterize the substitution patterns among stations. The framework described in this Section accounts for these two aspects by formally modeling the individual consumer's demand and retail pricing. These are the building blocks of the model for vertical negotiations between distributors and exclusive retailers.

In an environment without contractual price commitment, I use a bargaining model to characterize how short run wholesale prices are formed in ED relationships. One concern in business-to-business negotiations under ED is the fear of opportunistic behavior (holdup problem), since the supplier can appropriate a large share of the retailer's profits after the exclusivity contract is signed (Williamson, 1985). Since this potential holdup problem can discourage important ex-ante investments from taking place, the supplier might want to commit to a lower bargaining power in the ex-post negotiation over wholesale prices (Grout, 1984). Hence, it becomes important to quantify the bargaining power of each party in those relationships to understand eventual incentives for price increase following a merger[14].

The timing of the model is as follows: in the first stage, exclusive retailers and distributors bargain bilaterally to decide wholesale prices, and retailers simultaneously set retail prices for each type of fuel; in the second stage individual consumers choose which retailer to purchase from and the type of fuel if the consumer has flex fuel vehicle.[15] In the following I provide a detailed description of each component of the model.

1.3.1 Demand

Demand for gasoline (g) and ethanol (e) comes from a population of consumers characterized by a mixture of two groups: group 1 is composed by flex car owners and group 2 by gasoline car owners. The fraction of consumers in group 1 in market m is γ_m. Each consumer i in group 1 can purchase either type of fuel from any of the $r = 1, ..., J$ stations or not at all. A market is considered to be a Census Metropolitan Area in a given month. The size of market m is denoted by M_m and the total number of retailers in market m is denoted by J_m.

[14] See Appendix A.1 for a copy of a contract from a major distributor where it is mentioned that "wholesale prices are freely determined at the time of the purchase through a consensual agreement between the parties".

[15] Any other transfer is considered to be determined by contract and is decided before the bargaining takes place. In this paper I take the networks of exclusive dealers as given.

Location plays an important role in terms of product differentiation in the retail fuel market. The term $d\left(l_i, L_r\right)$ corresponds to the driving time from consumer i to retailer r. Following Houde (2012), the relevant distance considers the mobility of consumers in the product space and defines the location of consumer i as the commuting route between the home and work locations[16]. Given a pair of home and work locations, each consumer is assumed to take the optimal route in terms of travel time. The relevant distance from consumer i to retailer r is defined as the extra time that she takes to go to retailer r on their commuting path:

$$d\left(l_i, L_r\right) = t\left(home_i, L_r\right) + t\left(L_r, work_i\right) - t\left(work_i, home_i\right),$$

where $t\left(a, b\right)$ represents the optimal driving time from a to b.

The indirect utility of consumer i in group j purchasing fuel f from retailer r is

$$u_{irf}^{j} = \delta_{rf} + \lambda d\left(l_i, L_r\right) + \tau_{if} + \varepsilon_{irf}$$

where δ_{rf} is the mean utility of fuel f at station r, i.e. a product pair rf. The mean utility is assumed to be a linear function of observed station characteristics x_{rf}, prices p_{rf} and unobserved characteristics ξ_{rf}:

$$\delta_{rf} = x_{rf}\beta + \overline{\alpha} p_{rf} + \xi_{rf}.$$

The individual deviation from the mean utility is modeled as a function of distances

[16] Another possibility is what is known as single-address approach, adopted by several papers in the literature on retail competition (e.g. Davis (2006), Manuszak (2010), and Thomadsen (2005)). It considers the following distance metric:

$$d\left(l_i, L_r\right) = t\left(home_i, L_r\right) + t\left(L_r, home_i\right).$$

and idiosyncratic taste for each type of fuel, $\lambda d(l_i, L_r) + \tau_{if}$ plus an individual specific unobserved utility ε_{irf}.

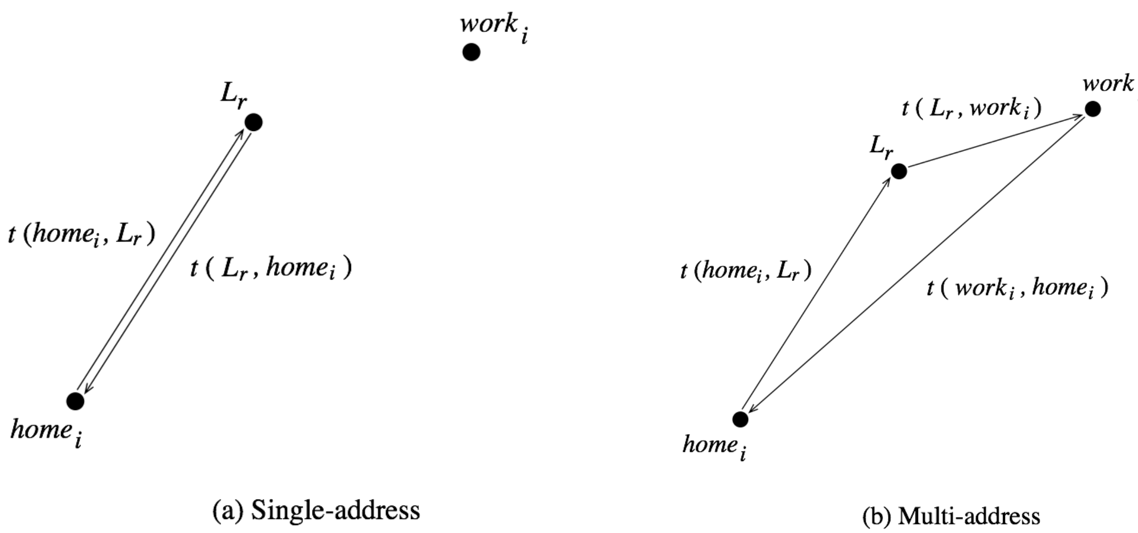


Figure 1.5: Single-address and multi-address distance metrics.

Taste parameter τ_{if} represents consumer i's valuation of fuel type f. Since there is no natural ordering between the two types of fuel, I set $\tau_{ie} = 0$ and assume $\tau_{ig} \sim \mathcal{N}(\mu_\tau, \sigma_\tau^2)$, with $(\mu_\tau, \sigma_\tau^2)$ to be estimated. This structure is consistent with horizontal differentiation between ethanol and gasoline.

The individual specific unobserved utility for each product (ε_{irf}) is assumed to follow a Type 1 Extreme distribution. This assumption implies that the conditional probability that consumer i will buy from station r is

$$P_{r|if} = \frac{exp(\delta_{rf} + \lambda d(l_i, L_r) + \tau_{if})}{1 + \sum_k exp(\delta_{kf} + \lambda d(l_i, L_k) + \tau_{if})}.$$

The expected value of choosing fuel type f for consumer i in group 1 is

$$I_{if} = ln\left(1 + \sum_k exp\left(\delta_{kf} + \lambda d\left(l_i, L_k\right) + \tau_{if}\right)\right).$$

The probability that consumer i from group 1 will choose ethanol is

$$P_{ie} = Pr\left(\tau_{ig} + I_{ig} \leq I_{ie}\right) = \Phi\left(\frac{I_{ie} - I_{ig} - \mu_\tau}{\sigma_\tau}\right)$$

and the probability that consumer i from group 1 will choose gasoline is $P_{ig} = 1 - P_{ie}$.

Retailer r's predicted market share of fuel $f \in \{e, g\}$ considering only consumers belonging to group 1 is:

$$s^1_{rf} = \frac{1}{size\ gr1} \sum_{i \in Group\,1} P_{r|if} P_{if}.$$

Consumers belonging to group 2 are gasoline car owners. In that case $s^2_{re} = 0$ and

$$s^2_{rg} = \frac{1}{size\ gr2} \sum_{i \in Group\,2} P_{r|ig}.$$

Omitting the subscripts for market, retailer r's total market share of gasoline is the average of the market shares in both groups, weighted by the fraction of consumers in each group:

$$s_{rg} = \gamma s^1_{rg} + (1 - \gamma)\, s^2_{rg}.$$

Since only consumers from group 1 can purchase ethanol, retailer r's market share of ethanol is

$$s_{re} = \gamma s^1_{re}.$$

1.3.2 Retail competition

I assume that each of the J multiproduct retailers operates as a single firm[17]. Given brand affiliation and the network of stations in the market, retailers simultaneously choose retail prices for gasoline and ethanol given wholesale prices and other costs. Omitting the market subscript, the problem of retailer r can be written as:

$$\max_{p_{rg},p_{re}} \sum_{f\in\{e,g\}} \left[\left(p_f^r - w_f^r - c_f^r\right) M s_f^r\left(\mathbf{p}\right) - \varphi^r\left(p_f^r - w_f^r\right) M s_f^r\left(\mathbf{p}\right)\right],$$

where φ^r represents the fraction of the gross margin that the retailer pays to the distributor in the form of royalties (or franchise fee) when it has an ED agreement. The parameter φ^r is set to zero if retailer r is independent[18].

Assuming that a pure strategy Bertrand-Nash equilibrium exists, the necessary first order condition can be rearranged to write equilibrium pricing as a function of wholesale prices, retailer's costs and mark-up:

$$\mathbf{p}^r = \mathbf{w}^r + \frac{1}{\left(1-\varphi^r\right)}\mathbf{c}^r - \Delta_r^{-1}\mathbf{s}_r\left(\mathbf{p}\right). \tag{1.3.1}$$

where $\mathbf{p}^r$ and $\mathbf{w}^r$ are the vectors of retail and wholesale prices associated to retailer r, $\mathbf{c}^r$ represents the cost of retailer r in addition to $\mathbf{w}^r$, $\mathbf{s}_r\left(\mathbf{p}\right)$ is the vector of market shares of retailer r and

$$\Delta_r = \begin{pmatrix} \frac{\partial s_g^r(\mathbf{p})}{\partial p_g^r} & \frac{\partial s_e^r(\mathbf{p})}{\partial p_g^r} \\ \\ \frac{\partial s_g^r(\mathbf{p})}{\partial p_e^r} & \frac{\partial s_e^r(\mathbf{p})}{\partial p_e^r} \end{pmatrix}.$$

[17] A coordinated behavior of the retailers can be easily accommodated in this model by assuming that each retailer's objective function is a weighted average of its own profit and the profit of its competitors.

[18] I don't observe individual contracts and in the empirical application I set $\varphi^r = 0.1$, which corresponds to the franchise fee charged by a major distributor in the market. This specific distributor does not differentiate among exclusive dealers in terms of franchise fee.

The FOC is used to simulate the new price equilibrium post merger. The approach consists in finding a fixed point of (1.3.1). The other purpose of the FOC is to uncover $\mathbf{c}^r$, expressing it as a function of observables and terms estimated in the demand model:

$$\mathbf{c}^r = \mathbf{p}^r - \mathbf{w}^r + \Delta_r^{-1}\mathbf{s}_r\left(\mathbf{p}\right).$$

1.3.3 Vertical negotiations between distributors and exclusive retailers

The network of exclusive retailers of distributor D is denoted by $\mathcal{N}^D$. Wholesale price paid by exclusive retailer $r \in \mathcal{N}^D$ to distributor D is determined by bilateral bargaining. In reality, these negotiations can be interdependent in the sense that in case of a disagreement between D and r, all wholesale and retail prices could potentially change. As in all papers in the literature on structural bargaining[19], I follow Horn and Wolinsky (1988) and condition the solution of the bargaining problem on all other prices. Hence, in the hypothetical case of a disagreement between retailer r and distributor D, all other retail and wholesale prices will remain the same. This assumption is made for tractability only and in equilibrium retail and wholesale prices will be optimal with respect to each other. Collard-Wexler et al (2015) show that, under some conditions, this solution is the unique PBE with passive beliefs of a specific simultaneous alternating offers game with multiple parties on both sides.

For all $r \in \mathcal{N}^D$, wholesale prices $w_f^r \in \mathbf{w}$ are negotiated simultaneously, with w_f^r being determined as the maximizer of the Generalized Nash Product (GNP):

$$w_f^r = \underset{w_f^r}{\operatorname{argmax}} \left(\Pi^D - d_{r,f}^D\right)^{b_D} \left(\Pi^r - d_f^r\right)^{1-b_D} \quad \forall r \in \mathcal{N}^D \; and \; f \in \{e, g\} \tag{1.3.2}$$

[19]e.g. Draganska, Klapper and Villas-Boas (2010), Crawford and Yurukoglu (2012), Grennan (2013), Gowrisankaran, Nevo and Town (2014), Ho and Lee (2015), Crawford, Lee, Whinston and Yurukoglu (2015).

where Π^D and Π^r represent the profits of the distributor and retailer, respectively. The terms $d^D_{r,f}$ and d^r_f represent distributor D and retailer r disagreement payoffs when negotiating over the wholesale price of fuel $f \in \{e, g\}$. Parameter $b_D \in [0, 1]$ is the bargaining weight associated to distributor D. Under this structure, each wholesale price w^r_f maximizes the product of distributor D and retailer $r$surpluses from the negotiation taking as given all other prices.

The profit of distributor D with network $\mathcal{N}^D$ is[20]

$$\Pi^D = \sum_{f\in\{e,g\},k\in\mathcal{N}^D} \left(w^k_f - p^{prod}_f - c^D_f\right) M s^k_f (\mathbf{p}).$$

When retailer $r \in \mathcal{N}^D$ and distributor D disagree on the wholesale price of fuel f, w^r_f, distributor D gets profit

$$d^D_{r,f} = \left[\sum_{\widetilde{f}\in\{e,g\},k\in\mathcal{N}^D\setminus\{r\}} \left(w^k_{\widetilde{f}} - p^{prod}_{\widetilde{f}} - c^D_{\widetilde{f}}\right) M \widetilde{s}^{-r,f}_{k,\widetilde{f}} (\mathbf{p})\right] + \left(w^r_{f^c} - p^{prod}_{f^c} - c^D_{f^c}\right) M \widetilde{s}^{-r,f}_{r,f^c} (\mathbf{p}),$$

where $\widetilde{s}^{-r,f}_{k,\widetilde{f}} (\mathbf{p})$ denotes the predicted market share of fuel $\widetilde{f}$ in retailer k if retailer r is not carrying fuel$f$ in that period. Each type of fuel is assumed to be negotiated separately, which means that if D and r disagree on w^r_f, nothing will change in terms of the price purchased by retailer r of the other fuel, denoted by f^c.

From the downstream model we have that the profit of retailer r is

$$\Pi^r = \sum_{f\in\{e,g\}} \left[\left(p^r_f - w^r_f - c^r_f\right) M s_{r,f} (\mathbf{p}) - \varphi^r \left(p^r_f - w^r_f\right) M s_{r,f} (\mathbf{p})\right].$$

In the occurrence of disagreement on w^r_f, the retailer will not sell that type of fuel but will

[20] For the sake of tractability, I assume that the franchise fees don't enter the profit of the distributor. This can be the case for example when these revenues are completely utilized for the purpose of advertisement or other efforts to promote the distributor's brand.

still be able to sell the other fuel, which implies on the following disagreement payoff:

$$d_f^r = \left(p_{f^c}^r - w_{f^c}^r - c_{f^c}^r\right) M\widetilde{s}_{r,f^c}^{-r,f}(\mathbf{p}) - \varphi^r \left(p_{f^c}^r - w_{f^c}^r\right) M\widetilde{s}_{r,f^c}^{-r,f}(\mathbf{p}).$$

The FOC of the maximization problem (1.3.2) in matrix form, stacking the two types of fuel negotiated with retailer r can be written as

$$\mathbf{w}^r = \left(\frac{1}{1-b_D\varphi^r}\right)\left[(1-b_D)\underbrace{\left(\mathbf{p}^{prod} + \mathbf{c}^D + \Omega_r^{-1}\mathbf{S}\mathbf{D}^r\right)}_{costs} + b_D \underbrace{\left(\mathbf{p}^r\left(1-\varphi^r\right) - \mathbf{c}^r\right)}_{value\ added\ by\ retailer\ r}\right] \tag{1.3.3}$$

where $\mathbf{S}$ is a diagonal matrix with shares of each type of fuel sold by retailer r,

$$\mathbf{D}^r = \begin{bmatrix} \sum_{f\in\{e,g\},\,k\in\mathcal{N}^D\setminus r}\left(w_f^k - p_f^{prod} - c_f^D\right)\frac{\Delta s_{k,f}^{-re}}{s_{r,e}} \\ \\ \sum_{f\in\{e,g\},\,k\in\mathcal{N}^D\setminus r}\left(w_f^k - p_f^{prod} - c_f^D\right)\frac{\Delta s_{k,f}^{-rg}}{s_{r,g}} \end{bmatrix} \quad \text{and} \quad \Omega_r = \begin{bmatrix} s_{r,e} & -\Delta s_{r,g}^{-re} \\ \\ -\Delta s_{r,e}^{-rg} & s_{r,g} \end{bmatrix}.$$

Price effects of an upstream merger The decomposition of $\mathbf{w}^r$ tells that the equilibrium wholesale price is a linear combination of the distributor's costs and the value added by the retailer. The vector $\mathbf{D}^r$ contains the value of the diverted sales to retailers belonging to network $\mathcal{N}^D$ in case of a disagreement with retailer r and enters as an opportunity cost for the distributor. This is the channel inducing a wholesale price increase following an upstream merger: the combined network $\widetilde{\mathcal{N}}^D$ implies that the new vector $\widetilde{\mathbf{D}}^r$ will be larger than $\mathbf{D}^r$, with the difference proportional to the diversion values to the new retailers in the network, $k \in \widetilde{\mathcal{N}}^D \setminus \mathcal{N}^D$. In equilibrium, retail prices of the stations in the network will increase in response to the increase in wholesale prices. Strategic complementarity at the downstream level implies that retail prices of the competing stations should also increase.

From equation (1.3.3) again, this leads to an increase in the wholesale price of the other distributors since the value added by their retailers will now be larger. The equilibrium following a merger will be a fixed point of these interactions between retail and wholesale price determination.

1.4 Identification and Estimation

1.4.1 Demand

The demand model is estimated using the nonlinear GMM method proposed by Berry, Levinsohn and Pakes (1995). The set of demand parameters is given by $\theta = \{\theta_1, \theta_2\}$, where $\theta_1 = \{\beta, \alpha\}$ and $\theta_2 = \{\mu_\tau, \sigma_\tau, \lambda\}$ are the vectors of linear and nonlinear parameters, respectively. Linear parameters (β) associated to the mean utility are identified under the assumption that the common characteristics are independent of ξ_{rf}. Identification of α is more complicated because of the correlation between retail prices and ξ_{rf}. The reason for this correlation is that consumers observe the quality index ξ_{rf} when choosing where to purchase fuel, which implies that prices can adjust to variations in this term, which is unobserved by the econometrician.

The unobserved product characteristics can be written as

$$\xi_{rf} = \delta_{rf}\left(\mathbf{s}, \mathbf{x}, \mathbf{p}, \theta_2\right) - x_{rf}\beta - \alpha p_{rf},$$

where $\mathbf{s}$ is the vector of predicted market shares described in Section 4.1. The estimation approach described by BLP is to first obtain δ as a solution to a fixed point problem and then construct the vector ξ to be used in the GMM estimation[21].

[21] The vector δ is obtained by finding the fixed point of the contraction problem

$$\delta^{(t+1)} = \delta^{(t)} + log\left(s^{obs}\right) - log\left(\hat{s}\left(\delta^{(t)}|\theta\right)\right).$$

Since the structural error ξ_{rf} is correlated with retail prices, we need valid instrumental variables. An instrumental variable should be (i) correlated with retail price and (ii) uncorrelated with the unobserved attributes of the station, ξ_{rf}. Retail price can be written as the sum of costs and markup, suggesting that a valid instrument must be some exogenous variable that impacts cost (i.e., a supply side instrument) or something exogenous that impacts mark up (i.e., a demand side instrument).

A natural candidate for instrument is the wholesale price paid by the retailer, which is correlated with the endogenous retail price. However, this variable can be problematic as an IV in the current setup. The reason is that distributors might take into consideration the specific demand drivers in the vertical negotiation, which implies that wholesale price is not a valid instrument.

Instead, I use two cost shifting instruments. The first is the wholesale price of the closest unbranded station[22]. While this variable must be be correlated with the cost of the exclusive retailers, the assumption that wholesale prices for independent stations are determined in a competitive way implies that it should not be correlated with the unobserved attributes of the stations. The second cost based IV is the interaction of producer price (refinery and distillery) with distance between the retailer and distributor. Both terms (producer's price and distance) must be correlated with retail prices through costs. Moreover, the interaction is important to create enough variation at the cross section level.

The advantages of the contraction mapping are the guarantee of a unique fixed point and that this fixed point will be reached for any initial value of δ. The main disadvantage is the slow convergence.The root finding equivalent to the BLP contraction is

$$f(\delta) \equiv log\left(\hat{s}\left(\delta^{(t)}|\theta\right)\right) - log\left(s^{obs}\right) = 0$$

and is solved using Newton's method. The potential problem with Newton's method is that it is not guaranteed to converge for any initial value. In order to provide "good" starting values, I run a few interactions of the contraction mapping from BLP and then switch to the root finding problem. The fast convergence of Newton's method relies on user providing an analytical expression for the Jacobian. If finite differences approximation to the Jacobian is used, Newton's method is very slow to converge.

[22]When the retailer is unbranded, this IV is the wholesale price paid by the station.

In addition to the cost based instruments mentioned above, I also use demand based IVs. The list of demand side instruments includes the exogenous own characteristics as well as average characteristics of the competitors within different distance radius. These instruments are correlated with prices since proximity in characteristic space will impact the stations' markup. Finally, I also added the number of competitors within different distance radius as additional demand based IVs.

For a given set of instruments Z, the vector of estimated demand parameters, $\hat{\theta}$ is characterized by:

$$\hat{\theta} = \text{argmin}_{\theta} \ \xi(\theta)' Z\Phi^{-1}Z'\xi(\theta),$$

where Φ is a consistent estimate of $\mathbb{E}\left[Z'\xi(\theta_0)\xi(\theta_0)' Z\right]$.

Taste parameters μ_τ and σ_τ are associated to consumers' valuation of each type of fuel and ensure that the model is consistent with the horizontal differentiation between the two types of fuel. Identification of these parameters comes mostly from the evolution on the number of flex vehicles, which allows a larger fraction of consumers to decide between the two over time. Identification of λ relies on the panel data dimension of the dataset as discussed in Houde (2012). The argument is that λ can be identified if entry and location choices are correlated with distribution of consumers and independent of ξ_{rf}.

1.4.2 Supply

The FOC for the retail pricing problem can be written as

$$\mathbf{c}^r = \mathbf{p}^r - \mathbf{w}^r + \Delta_r^{-1}\mathbf{s}_r(\mathbf{p}),$$

where $\mathbf{p}^r$ and $\mathbf{w}^r$ are observed in the data and the remaining term comes from the demand estimation. Since all terms in the right hand side are either observed or estimated, we can recover the marginal costs of the retailers using this equilibrium condition. I relate the uncovered marginal costs to some regressors as follows:

$$\mathbf{c}^r = W_r\gamma + \kappa_{b_r} + \gamma_t + \eta_r. \tag{1.4.1}$$

The vector W_r includes characteristics of retailer r such as number of attendants, number of pumps and ancillary services. The term κ_{b_r} captures brand specific intercept and γ_t is the time fixed effect. Equation (1.4.1) is estimated by OLS and the parameter representing the fraction of the gross margins that are used to pay the franchise fee is calibrated to $\varphi^r = 0.1$ when the station has an ED contract with a major distributor and 0 otherwise. This value is consistent with information obtained from industry sources.

I am not imposing the retail pricing equilibrium condition in the demand estimation. The advantage of this approach is that the demand will be consistently estimated in the case of misspecification in the supply side. The disadvantage is the lower precision of the estimates when the assumption on retail competition is valid.

The FOC (1.3.3) is written stacking the two types of fuel negotiated with retailer r. Since matrix $\mathbf{D}^r$ is a function of unobserved c^D, we need to rewrite that expression for the purpose of estimation. The equilibrium wholesale price can be written as the cost plus a margin that is proportional to the distributor's bargaining weight b_D:

$$\mathbf{w}^D = \mathbf{p}^{prod} + \mathbf{c}^D + \frac{b_D}{1-b_D} \underbrace{\left[\mathbf{S}^D\right]^{-1} \Omega \begin{bmatrix} -\Delta_1^{-1}\mathbf{s}_1\left(\mathbf{p}\right) \\ -\Delta_2^{-1}\mathbf{s}_2\left(\mathbf{p}\right) \\ \\ \\ -\Delta_{N_D}^{-1}\mathbf{s}_{N_D}\left(\mathbf{p}\right) \end{bmatrix}}_{B(\mathbf{p})}, \tag{1.4.2}$$

where $\mathbf{w}^D$ is the vector of wholesale prices to all retailers in the network of distributor D. Matrix Ω is block diagonal with blocks

$$\Omega_r = \begin{bmatrix} s_{r,e} & -\Delta s_{r,g}^{-re} \\ \\ -\Delta s_{r,e}^{-rg} & s_{r,g} \end{bmatrix}$$

on the main diagonal and $\mathbf{S}^D$ has market shares on the main diagonal and negative variation in the market shares $\left(-\Delta s_{r,f}^{-r,\tilde{f}}\right)$ in the off diagonal. Given the above expression for the wholesale prices, I assume that the distributor's marginal cost is a linear function of explanatory variables $\mathbf{h}$, $\mathbf{c}^D = \mathbf{h}\Gamma + \eta$ and estimate the following regression:

$$\mathbf{w}^D = \mathbf{p}^{prod} + \mathbf{h}\Gamma + \frac{b_D}{1-b_D} B\left(\mathbf{p}\right) + \eta.$$

The main difference between the equation (1.3.3) and (1.4.2) is that (1.4.2) includes all stations pertaining to network $\mathcal{N}^D$, while (1.3.3) is a matrix representation of the conditions involving each individual station in the network. Endogeneity in this case comes from the fact that $B(\mathbf{p})$ is a function of equilibrium wholesale prices. The cost and bargaining parameters are estimated by GMM under the assumption that $\mathbb{E}[\eta|\mathbf{Z}] = 0$, where $\mathbf{Z}$ is the vector of instruments described in the demand estimation.

Identification of the bargaining weights and cost parameters Γ relies on two sources of variations in negotiated wholesale prices between distributors and exclusive retailers: the within network variation and the variation across distributors. The derivation of the bargaining regression uses $\mathbf{p}^r - \mathbf{w}^r - \frac{1}{(1-\varphi^r)}\mathbf{c}^r = -\Delta_r^{-1}\mathbf{s}_r(\mathbf{p})$ from the equilibrium condition in the retail pricing. Hence, $B(\mathbf{p})$ is determined from the substitution patterns obtained in the demand model and the assumption on retail pricing. This implies that identification of the bargaining weights relies on information from marginal costs of the exclusive retailers and hence is conditional on the assumption about retail price competition and consistency of the demand estimation.

The franchise fee does not create a problem for the identification of the bargaining weights because, given φ^r and the retail pricing model, the value of the franchise fee is determined solely by wholesale prices. This is important because the bargaining weights cannot be identified if the bargaining impacts fixed transfers. I assume that all fixed transfers are negotiated at the time when the exclusivity contract is signed, which happens before the bargaining on wholesale price takes place.

As pointed in Gowrisankaran, Nevo and Town (2015), it is empirically difficult to identify bargaining weights and cost shifters at the same level. For this reason, I also do not include the distributors' fixed effects when estimating the bargaining weights for the different distributors.

1.5 Results

1.5.1 Demand

Table 3 presents the parameter estimates of the demand model. The model is estimated using monthly data over the period from January of 2007 to April of 2011, which is the month preceding the merger studied in the next Section.

The price coefficient is precisely estimated and indicates that consumers are highly price sensitive. To have an idea of the magnitude of this estimate, it implies an average price elasticity at the station level of -20.4. This estimate is of the same order of magnitude of studies in other markets such as Houde (2012), which considers the Canadian market and finds price elasticity of demand as high as -15 and Manuszak (2010), which finds for the Hawaiian market elasticities as high as -25.7. Price elasticity in the Brazilian market tends to be higher due to the possibility of freely substituting between ethanol and gasoline for the consumers who own flex vehicles.

The coefficient of the dummy variable for unbranded stations is also precisely estimated. The negative sign implies that consumers are willing to pay extra when purchasing from branded stations. The estimated value implies that, on average, an unbranded station has to give a discount of around 1.5% in order to make consumers indifferent relative to purchasing from a branded station, controlling for all other characteristics. Consumer surveys indicate that consumers normally associate branded stations to higher credibility or higher quality, which is not necessarily true since any brand of fuel of a given octane rating will run an automobile in the same way.[23] Hosken, McMillan and Taylor (2008) find that the only station characteristic that is a good predictor of the retail price heterogeneity is the station's brand affiliation.

[23] Relatedly, Bronnenberg, Dube, Gentzkow and Shapiro (2015) discuss the brand premium for health products and suggest that a sizable share of it can be explained by misinformation and consumer mistakes.

Table 1.3: Demand Estimates

Linear parameters	**Estimates**	**Std. Err.**
Price	-0.325***	0.019
Unbranded	-0.399***	0.007
Attendants	0.040***	0.001
Nozzles	0.026***	0.005
Convenience store	0.092***	0.005
Car wash	0.107***	0.006
Oil change	0.103***	0.010
Highway	-0.327***	0.018
Tire repair	-0.421***	0.007
Highway x tire repair	0.854***	0.054
Nonlinear parameters	**Estimates**	**Std. Err.**
μ_τ (avg. taste for gasoline)	0.92***	0.31
σ_τ (variation in taste)	8.82***	3.41
λ (distance coefficient)	-5.71***	0.12
Time FE	Yes	
Municipality FE	Yes	
Observations	15,135	
First stage F-statistic	23.07	
Average own-price elasticity	-20.38	

*** denotes significance at 1% level, ** at 5% level and * at 10% level.

The estimates of the linear parameters also indicate that consumers value having more attendants and many fueling positions (more nozzles) in the station, both of which are associated to less time spent to refuel. Moreover, convenience store, car wash and oil change significantly increase demand. The negative estimate for availability of tire repair is perhaps capturing the fact that some stations that offer this service are older and not well maintained. The negative coefficient on highway dummy variable suggests that the average consumer dislikes stopping at a highway station, which implies that those stations need to offer a lower price relative to the stations located in the city in order to attract the average consumer. The interaction of highway location and tire repair produces a positive coefficient with a magnitude higher than the sum of both coefficients on each variable separately, indicating that tire repair service significantly increases demand in stations located in highways.

Turning to the nonlinear parameters, the distance coefficient is sizable and precisely estimated. Considering a purchase of 25 liters, the estimated cost of driving an hour is R$17.58 ($= \frac{\widehat{\lambda}}{\widehat{\alpha}}$). This value is twice as big as the average industry wage in the country, which in 2010 was estimated[24] to be R$ 9,48. This result suggests that consumers tend not to deviate too much from their paths for purchasing fuel.

The remaining two nonlinear parameters characterize the individuals' tastes for each fuel. Both coefficients are significant at the 1% level, but not as precisely estimated as the distance coefficient. The positive sign of estimated μ_τ implies that the average consumer has a preference for gasoline compared to ethanol. To get a sense of how large the preference for gasoline is, when prices are equal to the ratio of calorific power[25], 0.7, the consumption of gasoline by the owners of flex fuel vehicles is estimated to be nearly 10% bigger than that of ethanol. The estimates indicate substantial variation in taste, captured by the

[24] From www.bls.gov/data.

[25] One liter of ethanol corresponds to approximately 0.7 liters of gasoline in terms of calorific power.

high value of σ_τ. This result is in line with the findings from Salvo and Huse (2011) and Anderson (2010), who document preference heterogeneity for each type of fuel, with a significant share of flex drivers choosing the most expensive fuel even when ethanol and gasoline energy equivalent prices differ by 20%.

1.5.2 Retail Pricing

Table 4 displays the estimates of the retailers' marginal costs from the retail pricing model. The coefficients are estimated by OLS with time and municipality fixed effects. The high price elasticity of the demand discussed in the last section implies that the market power of retailers is limited. The average gross margin ($\frac{p^r - w^r}{w^r}$) of the retailers is 12.4% and the estimated average markup ($\frac{p^r - c^r - w^r}{w^r}$) is 5.9%.

The number of attendants in the station is a measure of quality since it can proxy for the time spent in the station. The estimated coefficient on the number of attendants implies that the cost per liter of an attendant is 1.2 cents. Considering a station that sells 150k liters per month, this estimate implies a monthly cost of \$1,800, a value compatible with the costs of an attendant during the period studied, including taxes and salary paid by the station.

Table 1.4: Estimates of the marginal costs of retailers

Dep. Variable: Estimated Marginal Cost (in cents)		
Explanatory variable	**Estimates**	**Std. Err.**
Independent retailer	2.779***	0.180
Number of attendants	1.236***	0.144
Storage capacity	-1.024*	0.603
Number of nozzles	0.857**	0.398
Convenience store	1.110***	0.155
Oil change	1.047***	0.166
Car wash	-0.531***	0.152
Highway location	-0.366**	0.159
Time FE	Yes	
Municipality FE	Yes	
R-square	0.615	
Average gross margin	12.4%	
Average markup	5.9%	

*** denotes significance at 1% level, ** at 5% level and * at 10% level. The average gross margin is defined by $(\frac{p^r - w^r}{w^r})$ and the estimated average markup by $(\frac{p^r - c^r - w^r}{w^r})$.

Estimated marginal costs of independent retailers are on average larger than those of exclusive retailers. This can be related to the fact that exclusive retailers receive support from their distributors on things such as business plan and structure of the station as well as training for the managers and employees. Most of the remaining station characteristics are estimated to affect the marginal cost function of stations as expected. Larger stations (proxied by storage capacity) have on average lower marginal costs, but the estimated coefficient is significant only at the level of 10%. Stations located on a highway also have significantly lower marginal cost[26]. One interpretation for the negative coefficient of car wash is that this service is complementary to fuel sales, then reducing the effective cost of serving an additional consumer[27].

1.5.3 Bargaining

The estimates of the bargaining model using pre-merger data are presented in Table 5. Gowrisankaran, Nevo and Town (2015) argue that it is empirically difficult to identify bargaining weights and cost shifters at the same level. For this reason I follow their approach and estimate two specifications. The specification in column (a) allows the bargaining parameters to vary across distributors. In that case, I do not include the distributors' fixed effects. The specification in column (b) assumes equal bargaining weights for distributors and retailers, i.e., $b_D = 0.5$ for every D.

[26]There is a possible interference on the estimates of the coefficients of highway located and size because highway stations tend to be larger. The signs are preserved when I estimate the model with only highway dummy or storage capacity.

[27]This interpretation is analogous to the one provided by Houde (2012), who finds a similar result for car wash and convenience store and suggests that gasoline is a loss-leader product.

Table 1.5: Estimates of the bargaining model

	(a)		(b)	
Bargaining weight estimates	**Estimates**	**Std. Err.**	**Estimates**	**Std. Err.**
b_{BR}	0.52***	0.09	0.5	-
$b_{Esso,Shell}$	0.59***	0.11	0.5	-
$b_{Ipiranga}$	0.60***	0.11	0.5	-
Marginal cost estimates (cents)	**Estimates**	**Std. Err.**	**Estimates**	**Std. Err.**
Distance (km)	0.023*	0.012	0.025*	0.0136
Lag producer's price	0.027***	0.01	0.019**	0.009
BR	-	-	-0.025***	0.001
Esso/Shell	-	-	0.018***	0.001
Ipiranga	-	-	0.007**	0.003
Time FE	Yes		Yes	
Municipality FE	Yes		Yes	
number of observations	8766		8766	

*** denotes significance at 1% level, ** at 5% level and * at 10% level. In specification (a) I estimate the bargaining weights but do not include the distributors' fixed effects. In (b) I set the bargaining weights at 0.5 and include the distributors' fixed effects in the specification of marginal costs.

I find that the bargaining weights of the major distributors varies between 0.52 and 0.60, significantly smaller than unity, which is the "take it or leave it" value. The bargaining weights (specification (a)) and distributors' fixed effects (specification (b)) of the merging firms (Esso and Shell) are estimated together.[28] Bargaining weight varies across distributors, but not in a significant way. Moreover, none of the estimated bargaining weights is statistically different from 0.5. In the merger simulation I use the specification (b), which includes the distributors fixed effects in the equation for the marginal cost and assumes that bargaining weights of distributors and stations are the same.

The distance term in the marginal cost specification is included to capture both the transportation costs as well as any other managerial costs that vary with distance (e.g. monitoring the quality of the fuel and service provided by the station). According to industry sources, the shipping costs are flat up to a distance of around 250 km. Since the maximum distance to a distributor from stations within the Vitoria metropolitan area is approximately 145 km, the shipping costs are basically captured by the constant term. The lag of producer's price captures the cost of carrying stocks into the next period. It is precisely estimated and the magnitude of the estimated coefficient means that the cost of carrying stock to the next period is around 1.9% of the producers' price.

The distributors' fixed effects capture the average cost deviation with respect to the base group, which is the collection of all distributors other than Petrobras, Ipiranga, Esso and Shell and that have exclusive retailers. Among the major distributors, only Petrobras has estimated marginal cost lower than those in the base group.

1.6 Analysis of an Upstream Merger

In this Section, I present the results of the merger simulation and ex-post evaluation of the model predictions. I start with a short description of the merger studied. Next, I provide the details of the simulation methodology employed in the analysis, followed by the results,

[28]I also estimated the model allowing for different values of the bargaining weight for Esso and Shell, but they were not statistically different from each other. Making then distinct would create a problem for the merger simulation in terms of which value to use.

where I confront the predictions with the observed prices and shares following the merger.

1.6.1 Brief History of the Merger

Although there are around 200 fuel distributors in the country, the Brazilian distribution sector is very concentrated. As of 2010, the four largest distributors had a joint share of 67.2% in the gasoline market (BR (29.7%), Ipiranga (19.6%), Shell (11.2%) and Esso/Cosan (6.7%%))[29]. The proposed merger between Shell and Esso/Cosan raised a concern in the antitrust authority because of the high participation of the merging firms in some regions. The state of Espirito Santo was the one where Shell and Esso had higher participation, with 27% market share and around 25% of the stations being exclusive of either one of these distributors. After the merger, the exclusive retailers of the merged firm carrying the Esso brand were given the limit of three years to change to the Shell logo.

The exact date at which the two distributors started to operate jointly is not publicly known. The new company (named Raizen) announced that the joint operation should start by the end of the first semester of 2011. Based on ANP data on brand affiliation, the changes to Raizen were observed in the second half of May of 2011 (more precisely, on May/18). I assume that May/2011 was the month when the merged distributors started the joint operation. The vertical line in Figure 6 indicates the merger period. In that month, there was a jump in the average margins of the distributors, both the merging and non merging ones.

[29]Source: Anuario Estatistico ANP, 2011.

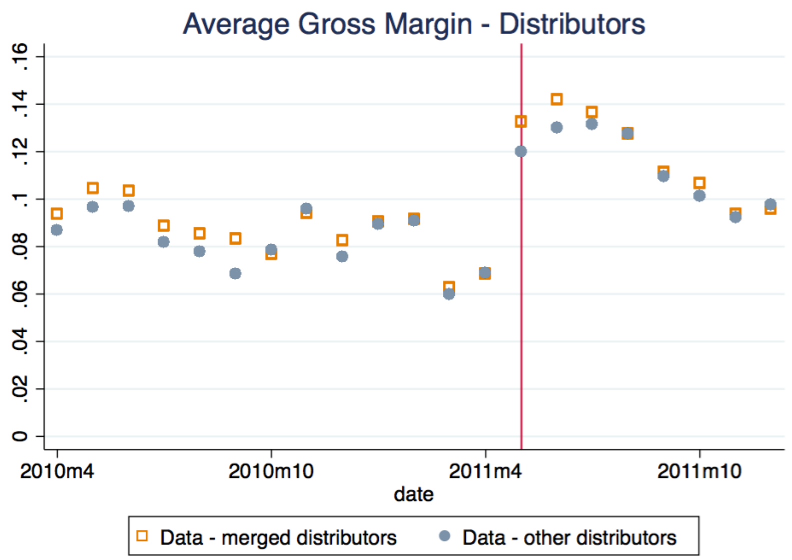


Figure 1.6: This figure illustrates the behavior of the average wholesale margins $\left(\frac{w-p^{prod}}{p^{prod}}\right)$ for the merged firms (orange squares) and non merged firms (blue circles). The vertical line indicates the period of the merger. Data from Vitoria Metropolitan Area (VMA).

1.6.2 Merger Simulation Analysis

To simulate the post merger equilibrium, I combine the networks of exclusive retailers of the merging distributors. The predicted wholesale prices are obtained by

$$\mathbf{w}^{post,r} = \left(\frac{1}{1-b_D\varphi^r}\right)\left[(1-b_D)\left(\mathbf{p}^{prod} + \mathbf{h}\widehat{\Gamma} + \Omega_r^{-1}\mathbf{S}\widetilde{\mathbf{D}}^r\right) + b_D\left(\mathbf{p}^{post,r}\left(1-\varphi^r\right) - \left(W_r\widehat{\gamma} + \widehat{\kappa}_{b_r} + \widehat{\gamma}_t\right)\right)\right],$$

where $\mathbf{h}\widehat{\Gamma}$ and $W_r\widehat{\gamma} + \widehat{\kappa}_{b_r} + \widehat{\gamma}_t$ represent the estimated costs of the distributor and retailer, respectively. Vector $\widetilde{\mathbf{D}}^r$ contains the diversion values considering the combined network.

The predicted post-merger retail price ($\mathbf{p}^{post,r}$) is obtained as the solution to the fixed point problem below:

$$\mathbf{p}^{post,r} = \mathbf{w}^{post,r} + W_r\widehat{\gamma} + \widehat{\kappa}_{b_r} + \widehat{\gamma}_t + \Delta_r^{-1}\left(\mathbf{p}^{post}\right)\mathbf{s}_r\left(\mathbf{p}^{post}\right),$$

which depends on the demand model through $\Delta_r^{-1}(\mathbf{p}^{post})$ and $\mathbf{s}_r(\mathbf{p}^{post})$.

Each component of the model has a clear role in the simulation of the wholesale and retail prices following a merger: (i) the demand is used to predict shares and price sensitivity; (ii) retail pricing is used to predict retailers' costs; and (iii) bargaining model is used to predict distributors' costs. To get $\mathbf{w}^{post}$ and $\mathbf{p}^{post}$ we iterate the bargaining, retail pricing and demand until convergence.

In the analysis that follows I am considering that the bargaining weights of the distributors and retailers are the same (i.e., $b_D = 0.5$). If the bargaining weights of the distributor increase as a result of the merger, it will get a larger share of the total profits, which will likely induce investments in cost savings.

In order to predict the effects of the merger we need to characterize the prices associated to independent retailers, which in the Vitoria Metropolitan Area corresponds to nearly 15% of the stations at the time of the merger. Independent retailer j is assumed to procure upstream from a set of N_j distributors, which compete in a reverse auction with no reserve price to serve retailer j. Each local distributor privately observes its own costs and all distributors are assumed to be risk neutral and behave non cooperatively. The cost considered in the bidding function does not include producer's price and is assumed to be independent across distributors. There are many variations of possible representations of this procurement auction. In the merger simulation I make the simplifying assumption that the wholesale price that each independent retailer is paying varies only to the extent of the variation in the producer's prices. This is equivalent to saying that the realizations of the private costs in the period of the merger are identical to those pre merger.

1.6.3 Results

Table 6 reports the simulated post-merger equilibrium gasoline prices for the Vitoria Metropolitan Area. The results assume zero efficiency gains from the merger.

Table 1.6: Merger simulation and observed data

	1: No merger (simulation)	
	Esso/Shell	**other**
Average wholesale price	2.591	2.587
Average wholesale gross margin	9.80%	9.63%
Average retail price	2.926	2.919
	2: Merger (simulation)	
	Esso/Shell	**other**
Average wholesale price	2.639	2.601
Δ wholesale prices (relative to the no merger simulation)	4.8 cpl (1.8%)	1.3 cpl (0.6%)
Average wholesale gross margin	11.84%	10.23%
Average retail price	2.969	2.946
Δ retail prices (relative to the no merger simulation)	4.3 cpl (1.5%)	2.7 cpl (0.9%)
	3: Observed data	
	Esso/Shell	**other**
Average wholesale price	2.672	2.642
Average wholesale gross margin	13.25%	12.22%
Average retail price	3.017	3.011

Note: Variation in retail and wholesale gasoline prices are measured in cents per liter (cpl). Average wholesale gross margin is computed as $\left(\frac{w-p^{prod}}{p^{prod}}\right)$.

On average, the model predicts a wholesale price increase of 4.8 cents per liter (cpl) for the merged distributors and 1.3 cpl for the non-merged. These changes correspond to an increase of 30% in the margins of the merging distributors compared to the average margin pre merger. At the retail level, the model predicts 4.3 cpl price increase for the exclusive stations of the merged distributors and 2.7 cpl for the remaining exclusive stations. This implies that exclusive retailers of the merging distributors are only partially passing-through the increase in the wholesale prices to retail prices. The average increase in the retail prices of the stations in the competing networks is higher than the increase in their wholesale prices. One possible explanation for this is that part of the price increase comes from the higher wholesale price, but another part comes from the strategic complementarity at the retail level, which incentivizes stations to increase retail prices after a price increase of the competitors.

Since the data span the post merger period, I am able to conduct an ex-post evaluation of the model simulation. The analysis presented here is focusing on the short run effects on prices. Figure 7 displays the simulated wholesale margins in the counterfactual case of no merger. The red diamond represents the distributors associated to the proposed merger and the blue diamond represents the remaining distributors. The simulated margins in the absence of merger do not exhibit any significant difference among the distributors.

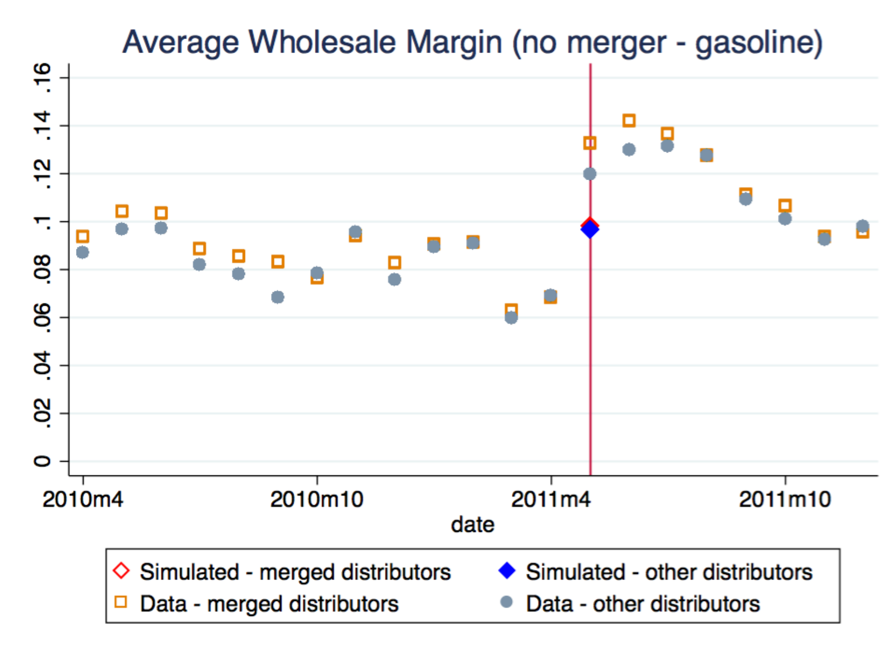


Figure 1.7: Observed and simulated wholesale prices post merger. Data from Vitoria Metropolitan Area (VMA).

The consolidation of exclusive dealing networks following the merger creates incentives for the merged supplier to increase wholesale prices, as we can observe in Figure 8. Importantly, the observed data confirms the model prediction that the increase in the margin of the merging distributors is larger than that of the other distributors. The observed wholesale price increase is of a similar order of magnitude, but larger than what is predicted by the model. The model can capture around 60% of the increase in the wholesale margins of the merging distributors, suggesting that other forces might have played an important role at the time of the merger.[30]

[30]The model also predicts that the price increase will not be uniform. However, the observed variation in the wholesale price for the region studied at the time of the merger was substantially lower than what was observed pre merger and than what was predicted by the model. I am currently estimating the model and merger simulation for another region that did not exhibit the same pattern in terms of reduction in the wholesale price variability.

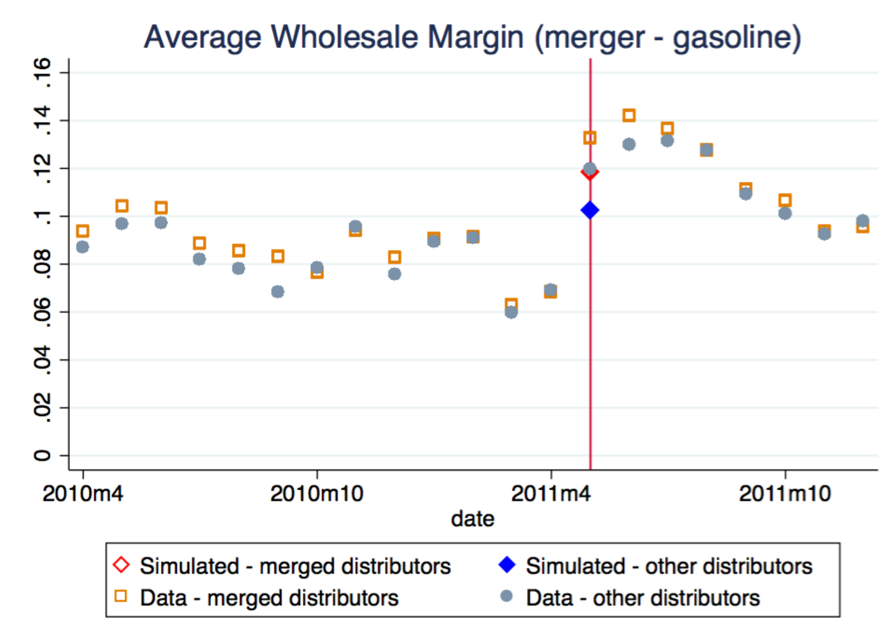


Figure 1.8: Observed and simulated wholesale prices post merger. Data from Vitoria Metropolitan Area (VMA).

Strategic complementarity at the downstream level implies that the unbranded retailers will eventually increase their retail prices in response to a price increase of the branded retailers. While the model correctly captures this response from the unbranded stations, it does not predict an increase of wholesale prices for independent retailers. The reason is that independent retailers can purchase from any distributor and in that case wholesale prices are assumed to be determined in a competitive way, depending on the cost structure of the distributors and not on the price charged by the independent retailer.

1.7 Conclusion

This paper studies upstream mergers in markets with ED contracts, where by ED I mean a contract between two firms in which one is prohibited from dealing with the competitors of the other. Under this type of arrangement, since the supplier is a monopolist with respect to the exclusive retailer, a merger between suppliers focusing on the input market has no clear effect on prices. The main result of the paper is to show that mergers under ED can produce sizable incentives for price increase. The network structure of exclusive retailers affects the bargaining position of distributors, creating a channel for wholesale price increase after a merger, with magnitude proportional to the diversion ratios within the network.

The assumption of simultaneous determination of wholesale and retail prices is used in this paper and in the literature on structural models of bargaining for tractability. I am currently working on relaxing this assumption, by allowing wholesale and retail prices to be determined sequentially. In order to do that, we need to understand how changes in the wholesale prices affect retail prices, which involves the computation of passthrough.

Although this paper provides an important step in understanding the vertical relations between suppliers and exclusive dealers, one limitation of the approach is that networks of exclusive retailers are taken as given. This assumption can be important when wholesale pricing influences the decision of retailers to become exclusive or to renew an existing exclusivity contract. In a related paper, I am extending the model to endogenize the network of exclusivity which will allow to access the incentives of retailers and distributors after a merger. This is important because antitrust authorities might want to consider the long run effects of a merger, including the incentives to become exclusive as well as which other mergers would be induced by the proposed merger.

The fuel industry is perhaps an extreme case because of product homogeneity. However, the key feature discussed in this paper is how the combination of the networks of exclusive dealers post merger can affect the incentives of the upstream firm to raise prices. Hence, the mechanism discussed here does not rely on the homogeneity of the goods sold. To a large extent, mergers in other industries where ED is common such as soft beverages, beer and automobiles would have similar incentives for price increase in the short run. In some cases, a long period is needed to realize the efficiencies from the merger as pointed in

Focarelli and Panetta (2003). Moreover, in the long run, since products in these industries can be highly differentiated, other aspects must be taken into consideration because of the possibility of changing product variety following a merger. Further study is needed to quantify the effects of upstream mergers in such contexts.

Chapter 2

Delayed Purchases in Markets with Network Effects

Consumers can find beneficial to delay their purchases when faced with uncertainty about quality or the adoption decisions by other consumers as well as in situations in which prices are decreasing over time. When quality is unknown at the time of the purchase decision, delay can be motivated by information acquisition in future periods. In the case of goods with network effects, consumer's valuation of a good is increasing in the number of other consumers buying the same good. This implies that prices alone do not convey sufficient information for consumers to make their decisions and, before doing that, they need to infer about the number of other consumers that are joining the network. Delaying purchases in this case will allow consumers to gain information about the popularity of the network and reduce ex-post regret. Finally, in the case of pure durable goods, the trade-off faced by consumers is basically early consumption versus lower future price.

In each of the cases described above, the strategic behavior of consumers can make the optimal decisions of the firm fairly complicated. Equilibrium analysis of consumer behavior in network markets is typically problematic because of the multiplicity of equilibria related to the coordination game played by consumers. Expectations of other consumers' behavior usually lead to non-monotonic demand functions and self-fulfilling multiple equilibria: a network that looks like succeeding (or failing) will as a result do so.

In this paper we consider a good with network effects which consumers can purchase in any of two periods (if at all). We model this problem as a dynamic game where a monopolist sells a network good to a population of consumers, who have private correlated intrinsic values for the good produced by the monopolist. In the beginning, each consumer receives a signal which is partially informative about both his intrinsic value and the fundamental representing the network quality. Then the firm announces a price for the first period and consumers decide whether or not to buy the product. Before the second period, adoption in the first period is realized and becomes public information. The monopolist will then post a second-period price and those consumers who have not joined the network in the first period make the final purchase decision. In this model, consumers face a trade-off regarding the delay of adoption choice. On one hand, waiting is costly because of late consumption and can also be costly because prices may raise when the network is successful. On the other hand, it is beneficial due to the more accurate information about adoption by other consumers and quality of the product. The option to delay might affect the demand function significantly, which in turn would affect the firm's profit-maximizing price.

The coordination problem among consumers is at the heart of models of goods with network effects and are also the source of the multiplicity of equilibria. Our analysis of the two period game with both consumers and firms taking actions requires further extension of the current state of the art. Our model builds on the literature on global games (see Carlsson and van Damme (1994); Morris and Shin (1998, 2001 and 2003)) in accounting for the fact that information structure is a crucial element in coordination models. We show that there exists a unique equilibrium of the game under some conditions on the parameters of the model. Taste heterogeneity, correlation of the valuations and imperfectly informed consumers generate a well-defined demand curve and profit-maximizing problem. Correlation in the valuation of consumers will "educate" their guesses about adoption and is an important feature to deliver uniqueness. With consumer heterogeneity about the intrinsic value of the good, the most enthusiastic consumers will typically become the early adopters, and those who are less enthusiastic will delay their purchase decision to future periods, when they will learn more about the product or about what other consumers are doing.

The general model that we consider nests the cases of herding and pure durable goods as special cases. In what we call the "herding version" of our model, we shut down the network effects and are stressing the importance of others' actions in the decision process. In the "full version" or the model, which includes information acquisition and network effects, the coordination problem associated to the decision of consumers stresses the complementarities in actions. In the "pure durable goods" version of the model, there are no informational gains (no herding) nor complementarities in actions (no network effects) and the basic choice weights early consumption versus lower price.

We illustrate the behavior that can emerge using a baseline parameterization of our model. We find two main results: first, the "critical mass effect" from the consumption behavior; the second-period consumption can be very sensitive to the adoption in the first-period. There exists a threshold fundamental quality below which there is no consumption in the second-period even when the price is equal to the monopolist's marginal cost. When the fundamental is above this threshold, the second period equilibrium demand rapidly increases in the fundamental. Second, the model features the "introductory price" phenomenon: with positive probability the monopolist increases its price over time. If the first-period consumption is high, then the positive feedback effect shifts the demand curve up. Therefore, the monopolist finds it optimal to announce a high price in the terminal period. This phenomenon does not happen in the case of pure durable goods, where prices are strictly decreasing over time.

Our paper relates to several strands in the literature on network effects.[1] Following Farrell and Saloner (1986), the literature on dynamic market with network effects mainly used sequential choice models where each consumer makes the adoption decision at each time.[2] This paper allows the continuum of consumers to simultaneously choose whether to join the network or wait, and the heterogeneous value and incomplete information lead to a unique equilibrium. As a result, our model generates richer dynamics such as positive feedback and critical mass effect.

There are papers studying delayed purchases for the case of unsponsored network(s), i.e., when firms problem are not incorporated in the model. Choi (1997) shows in the model with two consumers that allowing option to delay leads to less adoption since each consumer would want the other to experiment on the new technology (which he calls a "penguin effect"). We show in the model with a continuum of consumers and a monopolist pricing that adoption in the case with delay can be substantially larger than in the static benchmark case (no delay). Rysman (2003) analyzes adoption delay in a standards war (targeting the 56K modem case), where the delay occurs as the result of differentiated standards and because one type of consumer is not much enthusiastic about any specific technology. Contrary to these papers, the present paper considers the case of endogenous pricing, rich set of consumer types and does not suffer from multiplicity of equilibria.

This paper also contributes to the literature on monopoly price dynamics and Coase conjecture. Cabral, Salant, and Woroch (1999) study monopoly penetration pricing of durable network goods when buyers have rational expectations. In certain classes of example, they find that Coase conjecture price dynamics tend to predominate over penetration pricing: prices fall rather than rise over time, especially when there is complete information.

[1]See Farrell and Klemperer (2007) for an excellent survey.

[2]For more recent paper, see Chen, Doraszelski, and Harrington (2009) and Cabral (2011).

Radner *et al* (2014) study a network monopolist's dynamic pricing problem when adopters expect each period's network size to be equal to last period's; they find extreme bargain-then-ripoff pricing (the monopolist prices at zero until the network reaches its desired size). The present paper exhibits increasing price path over time with positive probability when network effects are considered.

Cabral (2011) studies dynamic price competition between two proprietary networks in an infinite period model with overlapping generations of consumers. He states conditions for equilibrium existence and uniqueness. Due to difficulty of getting analytical results, the main results of the paper are derived for restricted sets of parameter values and some simulations are performed in order to verify the robustness for a broader range of parameter values. Cabral's model assumes complete market coverage, meaning that consumers are not allowed to make delayed purchases. Moreover, in Cabral's paper there is no role for consumers' coordination since only one consumer is taking a decision per period.

Methodologically, the model of this paper is a global game with private correlated values. To the best of our knowledge, Argenziano (2008) and Julien and Pavan (2016) are the only papers in the literature to apply the ideas of global games to problems related to industrial organization. Argenziano (2008) considers competition between two firms in a one-sided market with network effects and signals perfectly informative about the private values. Julien and Pavan (2016) consider a static model of two-sided markets in which agents have incomplete information about quality. Compared to these papers, we model a dynamic game with an option to delay and focus on the adoption decisions and price dynamics. Additionally, we allow for information acquisition about the intrinsic value of the good. In the context of investment games, Dasgupta (2007) allows the players to delay their choice and shows that the delay leads to more efficient coordination: the present paper considers the market with network effects and adds the monopolist's endogenous pricing decision.[3]

The rest of the paper is organized as follows. Section 2 describes the model. Section 3 analyzes the static version of the model as a benchmark. In Section 4 we analyze the equilibrium of the model and discuss the uniqueness conditions. Section 5 presents a numerical analysis of the model, starting with a description of the algorithm used to compute the equilibrium and proceeding with a simulation of the model for a given parameterization. There we compare the full specification with the particular cases of pure durable goods, herding and the benchmark static model. Section 6 concludes and discusses questions for future work.

2.1 Model

The model consists of two periods $t = 1, 2$. A monopolistic firm sells a network good to a continuum of consumers of mass normalized to one, indexed by $i \in [0, 1]$. In the beginning of the first period, each consumer receives a signal which is informative about her value for the good. Next, the firm announces a price p_1, and consumers simultaneously decide whether or not to join the network. In the beginning of the second period, the firm and consumers observe adoption in the first-period. The firm will then announce the second-period price p_2 and consumers who did not purchase in the first period choose whether to join the network. Adoption decisions are irreversible.

If consumer i joins the network at time t, his payoff is

$$v_i + \alpha(n_1 + n_2) - p_t,$$

where v_i is consumer i's intrinsic value[4] of the network good, n_t is the measure of consumers

[3] See also Heidhues and Melissas (2006).

[4] Other terminologies for v_i from the literature are intrinsic benefit (Armstrong and Wright (2007)), stand-alone valuation (Julien and Pavan (2016)), and membership benefit (Weyl (2010)).

who join the network at time t, and p_t is the price of the network at time t.[5] The utility obtained from not joining any network is normalized to zero. We also assume that the cost of production is zero. Consumers and firm discounts the future at rate $\delta \in (0, 1)$.

The intrinsic value v_i is given by

$$v_i = \theta + h_i,$$

where θ is the fundamental quality of the network good and h_i is an idiosyncratic valuation. The fundamental θ is common across all consumers and is distributed normally with mean μ and variance σ_θ^2. The idiosyncratic valuation h_i is distributed according to a normal distribution with mean zero and variance σ_h^2. θ and h_i are independent of each other as well as across i. Consumers' valuations are correlated due to the presence of the common component θ.

In the beginning of the game, consumer i privately observes x_i which is partially informative about both θ and h_i. Specifically, we assume that

$$x_i = \theta + \eta_i,$$

where $\eta_i \sim \mathcal{N}(0, \sigma_h^2)$.[6] and $corr\,(\eta_i, h_i) = \tau$. The $\eta_i' s$ are independent across consumers and are independent of θ. Each consumer observes x_i but cannot identify θ and η_i separately.

[5] Even though the functional form of the consumer's payoff is particularly well suited to describe direct network effects, one can also think of these specifications as a reduced form representation of indirect network effects. Suppose for example that m is the number of applications (auxiliary/complementary products) for the network. Then $U_{i,t}$ will be a positive function of m, say $g_i\,(m)$. Since m depends on n, through, say $m = h\,(n)$, then $U_{i,t}$ depends on $g_i\,(h\,(n))$, i.e., $U_{i,t}$ is ultimately a function of n. Under indirect network effects it is also possible to justify the linearity of $U_{i,t}$ on n_1 and n_2 under some assumptions about the cost structure of firms providing complementary products.

[6] For simplicity we assume that the variances of h_i and η_i are the same.

The correlation of signals associated to consumers i and i' is given by

$$\rho \equiv corr\left(x_i, x_{i'}\right) = \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_h^2}.$$

Finally, the correlation of the signal of consumer i with his intrinsic value is

$$\kappa \equiv corr\left(x_i, v_i\right) = \frac{cov\left(x_i, v_i\right)}{\sqrt{var\left(x_i\right) var\left(v_i\right)}} = \frac{cov\left(\theta + \eta_i, \theta + h_i\right)}{\sigma_\theta^2 + \sigma_h^2} = \frac{\sigma_\theta^2 + \tau\sigma_h^2}{\sigma_\theta^2 + \sigma_h^2}.$$

When $\alpha \neq 0$, consumer i's utility depends on the purchase decision of other consumers. Because of this dependence, consumer i's beliefs about the decision of other consumers will ultimately affect her decision of whether or not to join the network. The common component θ is the source of correlation between the private valuations and is also the relevant element to characterize the beliefs (consumers' expectations) about the decision of other consumers.

The environment just described implies that the coordination game played by the consumers is a global game with private correlated values. The intrinsic value v_i^j is private in the sense that it tells the value of good j to consumer i excluding the network effect. Correlation comes from the common component θ, as discussed before.

This framework nests two interesting cases. First, when $\alpha = 0$ (no network effects) and $\tau < 1$ (information acquisition) we have the case of herding. Second, when $\alpha = 0$ and $\tau = 1$ we have the case of pure durable goods.

2.2 Benchmark: Static Model

As a benchmark, consider the case in which the consumers make the adoption decision only once. The payoff of the consumer who adopts is given by

$$v_i + \alpha n - p,$$

where n is the mass of consumers joining the network. We are interested in the equilibrium with monotone strategies: if a consumer with signal x_i chooses to join the network, then all consumers with $x_{i'} > x_i$ will also join. Any monotone strategy of the consumers can be characterized by a cutoff $\bar{x}$ where the consumer joins the network if and only if $x_i \geq \bar{x}$.

We start with the consumer's problem. Given a price p, consumer i will join the network if

$$U_i \equiv \mathbb{E}\left[v_i + \alpha n | x_i\right] - p \geq 0.$$

Under monotone strategies, the mass of consumers joining the network is

$$n = N(\bar{x}, \theta) \equiv \Phi\left(\frac{\theta - \bar{x}}{\sigma_h}\right).$$

Moreover, $\mathbb{E}\left[v_i|x_i\right] = (1-\kappa)\mu + \kappa x_i$, where $corr\left(v_i, x_i\right) = \kappa$ and $\mathbb{E}\left[n|x_i\right] = \mathbb{E}\left[\Phi\left(\frac{\theta-\bar{x}}{\sigma_h}\right)|x_i\right] = \Phi\left(\left(\mu - \overline{x}\right) z_1\right)$, with $z_1 = \sqrt{\frac{(1-\rho)}{(1+\rho)\left(\sigma_\theta^2+\sigma_h^2\right)}}$ and $\rho = \frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_h^2}$.

The marginal consumer $\bar{x}$ must be indifferent between adopting and not adopting since U_i is continuous in x_i. Therefore, her indifference condition is given by

$$(1-\kappa)\mu + \kappa\bar{x} + \alpha\Phi\left(\left(\mu - \overline{x}\right) z_1\right) = p. \tag{2.2.1}$$

Proposition 1. *Suppose $z_1 \frac{\alpha}{\kappa} < \sqrt{2\pi}$. Then $\bar{x}$ is uniquely determined for any p. Moreover, $\bar{x}$ is increasing in p.*

Proof. Given any p, The function

$$(1-\kappa)\mu+\kappa\bar{x}+\alpha\Phi\left(\left(\mu-\overline{x}\right)z_1\right)-p$$

will be negative for sufficiently negative values of $\overline{x}$ and will be positive for sufficiently high values of $\overline{x}$. Since this function is continuous in $\overline{x}$, Intermediate Value Theorem implies that there exists a solution to above indifference condition. Uniqueness is obtained by showing monotonicity of the left-hand side of the indifference condition:

$$\frac{\partial(LHS)}{\partial\overline{x}}=\kappa-\alpha z_1\phi\left(\left(\mu-\overline{x}\right)z_1\right)\geq\kappa-\frac{\alpha z_1}{\sqrt{2\pi}},$$

which is strictly positive when $\frac{\alpha z_1}{\kappa}<\sqrt{2\pi}$, completing the proof of uniqueness. Monotonicity also follows. $\square$

The uniqueness condition is satisfied in two cases. First, for sufficiently small σ_h. In this case, we have $\rho\to 1$ so z_1 becomes small. This corresponds to the case where the degree of heterogeneity is small relative to the uncertainty about the quality. Since there is high correlation between the x_i's, the degree of strategic uncertainty becomes insensitive to the signal. Second, for sufficiently high σ_h and small σ_θ. In this case, we have $\rho\to 0$ with high σ_h^2, so z_1 becomes sufficiently small. This corresponds to the case where the degree of heterogeneity is large. Since idiosyncratic payoffs are almost uncorrelated, the degree of strategic uncertainty again becomes insensitive to the signal.

The next proposition shows that the threshold consumer has virtually no information about the aggregate choices. This is called Laplacian property and was introduced by Morris and Shin (2003, p. 62) for the case of improper uniform prior beliefs.

Proposition 2. *The distribution of* $n\left(\theta\right)$ *conditional on* $x_i=\overline{x}$ *is uniform on* $[0,1]$.

Proof. Since $x_i = \theta + h_i$ and $h_i \sim \mathcal{N}(0, \sigma_h^2)$,

$$\begin{aligned}
\Pr[n(\theta) \leq z|\overline{x}] &= \Pr\left[1 - \Phi\left(\frac{(\overline{x} - \theta)}{\sigma_h}\right) \leq z|\overline{x}\right] \\
&= \Pr\left[\Phi\left(\frac{h_i}{\sigma_h}\right) \geq (1 - z)\right] \\
&= 1 - \Pr\left[\frac{h_i}{\sigma_h} < \Phi^{-1}(1 - z)\right] \\
&= 1 - \Phi\left(\Phi^{-1}(1 - z)\right) \\
&= 1 - (1 - z) = z
\end{aligned}$$

which proves the result. $\square$

Given the demand curve, the monopolist chooses the price p that maximizes the second-period profit $\Pi(p)$:

$$\max_{p_2 \geq c} \Pi(p) = \max_{p \geq c} (p - c)\mathbb{E}\left[N(\bar{x}(p), \theta)\right],$$

subject to (2.2.1). By plugging the constraint into the profit function, we have

$$\max_{\overline{x}} ((1 - \kappa)\mu + \kappa\bar{x} + \alpha\Phi\left((\mu - \overline{x})\, z_1\right) - c)\Phi\left(\beta(\bar{x})\right),$$

where $\beta(\bar{x}) = \frac{\mu - \bar{x}}{\sqrt{\sigma_\theta^2 + \sigma_h^2}}$.

First-order condition gives

$$\underbrace{(\kappa - \alpha z_1 \phi\left((\mu - \overline{x})\, z_1\right))\Phi\left(\beta(\bar{x})\right)}_{\text{benefit from price increase}} = \underbrace{\frac{(1 - \kappa)\mu + \kappa\bar{x} + \alpha\Phi\left((\mu - \overline{x})\, z_1\right) - c}{\sqrt{\sigma_\theta^2 + \sigma_h^2}} \phi\left(\beta(\bar{x})\right).}_{\text{cost from demand decrease}}$$

Solving, we have

$$(1 - \kappa)\mu + \kappa\bar{x} + \alpha\Phi\left((\mu - \overline{x})\, z_1\right) = c + \left(\kappa - \alpha z_1 \phi\left((\mu - \overline{x})\, z_1\right)\right) \sqrt{\sigma_\theta^2 + \sigma_h^2} \frac{\Phi\left(\beta(\bar{x})\right)}{\phi\left(\beta(\bar{x})\right)}. \quad (2.2.2)$$

which gives the optimal $\bar{x}$. Combining this with (2.2.1), we have the firm's optimal price

$$p = c + \underbrace{\left(\kappa - \alpha z_1 \phi\left(\left(\mu - \overline{x}\right) z_1\right)\right) \sqrt{\sigma_\theta^2 + \sigma_h^2} \frac{\Phi\left(\beta(\bar{x})\right)}{\phi\left(\beta(\bar{x})\right)}}_{\text{markup}}. \tag{2.2.3}$$

Note that the markup characterized in (2.2.3) is strictly positive if $\frac{\alpha z_1}{\kappa} > \sqrt{2\pi}$ (assumption for Proposition 1).

2.3 Equilibrium

We are interested in the equilibrium with monotone strategies: if a consumer with signal x_i chooses to join the network at period n, then all consumers with $x_{i'} > x_i$ join at period $t' \leq t$. Any monotone strategy of the consumers can be characterized by the sequence of cutoffs $\bar{x}_1$ and $\bar{x}_2$ ($\bar{x}_1 \geq \bar{x}_2$) where the remaining consumers at period t join the network if and only if $x_i \geq \bar{x}_t$. For a monotone strategy $(\bar{x}_1, \bar{x}_2)$ and a fundamental θ, the mass of consumers joining the network in each period is given by

$$\begin{aligned} n_1 &= N_1(\bar{x}_1, \theta) \equiv \Phi\left(\frac{\theta - \bar{x}_1}{\sigma_h}\right), \\ n_2 &= N_2(\bar{x}_1, \bar{x}_2, \theta) \equiv \Phi\left(\frac{\theta - \bar{x}_2}{\sigma_h}\right) - \Phi\left(\frac{\theta - \bar{x}_1}{\sigma_h}\right). \end{aligned}$$

Second Period: Consumer's Problem After the first period, the mass of consumers in the network n_1 is publicly observed, so the fundamental θ is perfectly known. After the firm announces a price p_2, All consumers who delayed their decision in the first period decide to join the network or not. The consumer i will join the network in the second period if

$$U_{i,2} = \mathbb{E}\left[v_i + \alpha(n_1 + n_2)|x_i, \theta\right] - p_2 \geq 0.$$

Given a monotone strategy $(\bar{x}_1, \bar{x}_2)$,

$$\begin{aligned} U_{i,2} &= \mathbb{E}\left[v_i | x_i, \theta\right] + \alpha\Phi\left(\frac{\theta - \bar{x}_2}{\sigma_h}\right) - p_2 \\ &= \theta + \mathbb{E}\left[h_i | \eta_i\right] + \alpha\Phi\left(\frac{\theta - \bar{x}_2}{\sigma_h}\right) - p_2 \\ &= (1-\tau)\theta + \tau x_i + \alpha\Phi\left(\frac{\theta - \bar{x}_2}{\sigma_h}\right) - p_2, \end{aligned}$$

since $h_i | \eta_i \sim N\left(\tau\eta_i, (1-\tau)\,\sigma_h^2\right)$ and $\eta_i = x_i - \theta$.

Let's first consider the equilibrium where $\bar{x}_2 < \bar{x}_1$, that is, a positive measure of consumers join the network in the second period. In this case, consumer $\bar{x}_2$ must be indifferent between buying and not buying since $U_{i,2}$ is continuous in x_i. Therefore, his indifference condition in the second period is given by

$$(1-\tau)\theta + \tau\bar{x}_2 + \alpha\Phi\left(\frac{\theta - \bar{x}_2}{\sigma_h}\right) = p_2. \tag{2.3.1}$$

Let $\hat{x}_2(p_2;\theta)$ be the solution of (2.3.1). This condition will characterize the second period threshold whenever $\hat{x}_2(p_2;\theta) < \overline{x}_1$. If $\hat{x}_2(p_2;\theta) \geq \overline{x}_1$, all consumers who waited in the first period are better off not buying the product, hence $\overline{x}_2 = \overline{x}_1$. This corresponds to the situation in which no consumer will join the platform in the second period for the given realization of θ and the price charged by the existing platform, p_2. Therefore, the cutoff function $\overline{x}_2\left(p_2;\bar{x}_1,\theta\right)$ is defined as

$$\overline{x}_2\left(p_2;\bar{x}_1,\theta\right) = \min\{\bar{x}_1, \hat{x}_2\left(p_2;\theta\right)\}.$$

A unique threshold in the second period relies on sufficiently large consumer heterogeneity, in the terms of the assumption below:

Assumption 3. $\frac{\tau\sigma_h}{\alpha} > \frac{1}{\sqrt{2\pi}}$.

This assumption will be satisfied under sufficiently large heterogeneity or a mild contemporaneous network externality in the second period[7].

Proposition 4. *Suppose Assumption 3 holds. Then given any* $\bar{x}_1$ *and* θ*,* $\bar{x}_2(p_2; \bar{x}_1, \theta)$ *is uniquely determined for any* $p_2 \geq c$*. Moreover,* $\bar{x}_2$ *is nondecreasing in* p_2*.*

Proof. Given $\bar{x}_1$, θ and $p_2 \geq c$, all terms in (2.3.1) are bounded, except for $\overline{x}_2$. Then,

$$(1-\tau)\theta + \tau\overline{x}_2 + \alpha\Phi\left(\frac{\theta - \bar{x}_2}{\sigma_h}\right) - p_2$$

will be negative for sufficiently negative values of $\overline{x}_2$ and will be positive for sufficiently high values of $\overline{x}_2$. Since this function is continuous in $\overline{x}_2$, Intermediate Value Theorem implies that there exists a solution to (2.3.1). Uniqueness is obtained by showing monotonicity of the left-hand side of (2.3.1):

$$\frac{\partial(LHS)}{\partial\overline{x}_2} = \tau - \frac{\alpha}{\sigma_h}\phi\left(\frac{\theta - \bar{x}_2}{\sigma_h}\right) \geq \tau - \frac{1}{\sqrt{2\pi}}\frac{\alpha}{\sigma_h},$$

which is strictly positive under Assumption 3, completing the proof of uniqueness. Monotonicity also follows. □

Second Period: Network's Problem For any $\bar{x}_1$ and θ, (2.3.1) characterizes the second-period demand function $N_2(\bar{x}_1, \overline{x}_2(p_2), \theta)$ as a function of p_2. Given this, the monopolist chooses the price p_2 that maximizes the second-period profit $\Pi_2(p_2; \bar{x}_1, \theta)$:

$$\begin{aligned}\max_{p_2 \geq c} \Pi_2(p_2; \bar{x}_1, \theta) &= \max_{p_2 \geq c} (p_2 - c)\, N_2(\bar{x}_1, \overline{x}_2(p_2), \theta) \\ & \quad s.t.\, \bar{x}_2(p_2; \bar{x}_1, \theta) = \min\{\bar{x}_1, \hat{x}_2(p_2; \bar{x}_1, \theta)\}.\end{aligned}$$

[7](Add a brief discussion based on Morris and Shin (2006) and Herrendorf et al (2000)). We will see that the condition for uniqueness in the first period relies on the opposite: sufficiently low heterogeneity. This is important there because it leads to a high correlation of the signals and also very precise information about the fundamental θ.

Depending on the realization of θ, the monopolist may or may not be able to make a positive profit. To understand why this is the case, note that if θ turns out to be sufficiently low, then the second-period network size n_2 can be zero even if the network sets the price equal to its marginal cost. Specifically, given p_2 the second-period network size n_2 would be zero when consumer $\bar{x}_1$ would find optimal not to join the network. That is, from (2.3.1),

$$(1-\tau)\theta + \tau\bar{x}_1 + \alpha\Phi\left(\frac{\theta - \bar{x}_1}{\sigma_h}\right) < p_2.$$

Define $\underline{\theta}(\bar{x}_1)$ be the solution of the equation

$$(1-\tau)\theta + \tau\bar{x}_1 + \alpha\Phi\left(\frac{\theta - \bar{x}_1}{\sigma_h}\right) = c. \tag{2.3.2}$$

Note that $\underline{\theta}(\bar{x}_1)$ is uniquely defined, and under Assumption 3 $\underline{\theta}(\bar{x}_1)$ is decreasing in $\bar{x}_1$. Then (2.3.1) implies that, $\overline{x}_2 = \bar{x}_1$ (and hence $n_2 = 0$) for any $p_2 \geq c$ if and only if $\theta < \underline{\theta}$. In this case the monopolist cannot make a positive profit. We assume that for $\theta \leq \underline{\theta}$, the monopolist charges its marginal cost c.

Now consider the case where $\theta > \underline{\theta}$. In this case, there exists a price $p_2 > c$ at which the demand is positive and the network earns positive profit. Therefore, the firm's optimal price must be strictly greater than c, and $\bar{x}_2$ must satisfy (2.3.1), that is, $\bar{x}_2 = \hat{x}_2\left(p_2; \bar{x}_1, \theta\right)$. Then the optimal p_2 is the solution to

$$\begin{aligned}\max_{p_2 \geq c} \Pi_2(p_2; \bar{x}_1, \theta) \quad &\equiv \quad \max_{p_2 \geq c} \; (p_2 - c)\, N_2(\bar{x}_2(p_2)) \\ s.t. \quad & p_2 = (1-\tau)\theta + \tau\bar{x}_2 + \alpha(N_1 + N_2(\bar{x}_2(p_2)))\end{aligned}$$

By plugging the constraint into the profit function, we have

$$\max_{\bar{x}_2 \leq \bar{x}_1} ((1-\tau)\theta + \tau\bar{x}_2 + \alpha(N_1 + N_2(\bar{x}_2)) - c) N_2(\bar{x}_2).$$

First-order condition gives

$$\underbrace{\left(\tau - \frac{\alpha}{\sigma_h}\phi\left(\frac{\theta - \bar{x}_2}{\sigma_h}\right)\right) N_2(\overline{x}_2)}_{\text{benefit from price increase}} = \underbrace{((1-\tau)\theta + \tau\bar{x}_2 + \alpha(N_1 + N_2(\bar{x}_2)) - c)\,\frac{1}{\sigma_h}\phi\left(\frac{\theta - \bar{x}_2}{\sigma_h}\right)}_{\text{cost from demand decrease}}$$

Solving, we have
$\left(\frac{\tau\sigma_h}{\phi\left(\frac{\theta-\bar{x}_2}{\sigma_h}\right)} - \alpha\right) N_2(\overline{x}_2) = (1-\tau)\theta + \tau\bar{x}_2 + \alpha(N_1 + N_2(\bar{x}_2)) - c$, or

$$(1-\tau)\theta + \tau\bar{x}_2 + \alpha N_1 = c + \left(\frac{\tau\sigma_h}{\phi\left(\frac{\theta-\bar{x}_2}{\sigma_h}\right)} - 2\alpha\right) N_2(\overline{x}_2). \tag{2.3.3}$$

combining this with (2.3.1), equilibrium second period price is characterized by

$$p_2 = c + \underbrace{\left(\frac{\tau\sigma_h}{\phi\left(\frac{\theta-\bar{x}_2}{\sigma_h}\right)} - \alpha\right) N_2(\overline{x}_2)}_{\text{markup}}. \tag{2.3.4}$$

Assumption 3 guarantees that the markup presented in (2.3.4) will be strictly positive.

Assumption 5. $\frac{\tau\sigma_h}{\alpha} > \sqrt{\frac{2}{\pi}}$.

Assumption 5 is a stronger version of Assumption 3 and is used in the next proposition to show single peakedness of the second period profit function.

Proposition 6. *Suppose Assumption 5 holds. Then given any $\bar{x}_1$ and θ, the network's problem in the second period has unique solution, which is determined by the first-order condition (2.3.3).*

Proof. It is sufficient to show that for any $\bar{x}_1$ and θ, the second-period profit as a function of $\bar{x}_2$,

$$\Pi_2(\bar{x}_2) = ((1-\tau)\theta + \tau\bar{x}_2 + \alpha\left(N_1 + N_2(\bar{x}_2)\right) - c)N_2(\bar{x}_2),$$

is single-peaked for $\bar{x}_2 \in (-\infty, \bar{x}_1)$. It is clear that $\Pi_2(\bar{x}_1) = 0$ and $\lim_{\bar{x}_2 \to -\infty} \Pi_2(\bar{x}_2) < 0$. Since there exists $\bar{x}_2 \in (-\infty, \bar{x}_1)$ such that $\Pi_2(\bar{x}_2) > 0$, it suffices to show that the first-order condition (2.3.3) has a unique solution. Rearranging (2.3.3), we have

$$\tau\bar{x}_2 + \left(2\alpha - \frac{\tau\sigma_h}{\phi\left(\gamma(\bar{x}_2)\right)}\right) N_2(\bar{x}_2) = -(1-\tau)\theta - \alpha N_1 + c. \tag{2.3.5}$$

Taking the derivative of the left-hand side of (2.3.5) we have

$$\frac{\partial(LHS)}{\partial\bar{x}_2} = 2\tau - 2\frac{\alpha}{\sigma_h}\phi\left(\gamma(\bar{x}_2)\right) + \tau\gamma(\bar{x}_2)\frac{N_2(\bar{x}_2)}{\phi\left(\gamma(\bar{x}_2)\right)},$$

since $\phi'(x) = -x\phi(x)$. To find the lower bound of the third term, we use a simple algebra to have

$$\min \gamma(\bar{x}_2)\frac{N_2}{\phi\left(\gamma(\bar{x}_2)\right)} > \min_{\bar{x}_2} \gamma(\bar{x}_2)\frac{\Phi(\gamma(\bar{x}_2))}{\phi\left(\gamma(\bar{x}_2)\right)} = \min_{\xi} \xi\frac{\Phi(\xi)}{\phi(\xi)}.$$

What is the lower bound on $\min_\xi \xi\frac{\Phi(\xi)}{\phi(\xi)}$? Let $\xi^* = \arg\min_\xi \xi\frac{\Phi(\xi)}{\phi(\xi)}$. The first-order condition gives

$$\frac{\Phi(\xi^*)}{\phi(\xi^*)} + \xi^* + (\xi^*)^2\frac{\Phi(\xi^*)}{\phi(\xi*)} = 0,$$

Hence

$$\xi^*\frac{\Phi(\xi^*)}{\phi(\xi^*)} = -\frac{(\xi^*)^2}{1+(\xi^*)^2} > -1.$$

We conclude that

$$\begin{aligned}\frac{\partial(LHS)}{\partial\bar{x}_2} &> \tau - 2\frac{\alpha}{\sigma_h}\phi\left(\gamma(\bar{x}_2)\right) \\ &\geq \tau - \sqrt{\frac{2}{\pi}}\frac{\alpha}{\sigma_h}.\end{aligned}$$

and then under Assumption 5, the left-hand side of (2.3.5) is strictly increasing in $\bar{x}_2$, which guarantees the uniqueness of the solution of (2.3.5). $\square$

The equilibrium analysis in the second period gives a unique pair $(p_2, \bar{x}_2)$ as functions of $\bar{x}_1$ and θ. We now turn to the equilibrium analysis in the first period.

First Period: Consumer's Problem Now we consider the equilibrium behavior in the first period given the second-period behavior $p_2(\bar{x}_1, \theta)$ and $\bar{x}_2(\bar{x}_1, \theta)$. In the first period, consumer i receives a signal x_i, observes the first-period price p_1, and then chooses whether to join the network. Given a monotone strategy $\bar{x}_1$, if consumer x_i joins the network in the first period, his expected payoff is

$$(1-\kappa)\mu + \kappa\overline{x}_1 + \alpha\mathbb{E}_\theta\left[N_1(\overline{x}_1, \theta) + N_2(\overline{x}_1, \overline{x}_2(\overline{x}_1, \theta), \theta)|x_i\right] - p_1.$$

On the other hand, if he delays his purchase decision his payoff is discounted by δ but he has the option to purchase or not after θ is realized. In this case, his expected payoff is given by

$$\delta\mathbb{E}_\theta\left[\max\left\{0, (1-\tau)\theta + \tau x_i + \alpha(N_1(\overline{x}_1, \theta) + N_2(\overline{x}_1, \overline{x}_2(\overline{x}_1, \theta), \theta)) - p_2(\bar{x}_1, \theta)\right\}|x_i\right],$$

or

$$\delta \Pr{}_\theta(\bar{x}_2(\bar{x}_1, \theta) < x_i|x_i)\left\{(1-\tau)\theta + \tau x_i + \mathbb{E}_\theta\left[\alpha(N_1 + N_2) - p_2(\bar{x}_1, \theta)|x_{i,}\bar{x}_2(\bar{x}_1, \theta) < \bar{x}_i\right]\right\},$$

since consumer i joins the network in the second period if and only if $x_i \geq \bar{x}_2(\bar{x}_1, \theta)$. The consumer who receives a signal $\bar{x}_1$ must be indifferent between joining the network and delaying the purchase. Therefore, his indifference condition is given by

$$\begin{aligned}(1-\kappa)\mu+\kappa\overline{x}_1 &+ \alpha\mathbb{E}_\theta\left[N_1+N_2|\bar{x}_1\right]-p_1 \\ &= \delta\Pr{}_\theta(\bar{x}_2<\bar{x}_1|\bar{x}_1)\left\{(1-\tau)\theta+\tau\bar{x}_1+\mathbb{E}_\theta\left[\alpha(N_1+N_2)-p_2|\bar{x}_1,\bar{x}_2<\bar{x}_1\right]\right\}\end{aligned}$$

Since (2.3.1) holds whenever $\bar{x}_2(\bar{x}_1,\theta)<\bar{x}_1$, we have

$$(1-\kappa)\mu+\kappa\overline{x}_1+\alpha\mathbb{E}_\theta\left[N_1+N_2|\bar{x}_1\right]-p_1=\delta\Pr{}_\theta(\bar{x}_2<\bar{x}_1|\bar{x}_1)\tau\left\{\bar{x}_1-\mathbb{E}_\theta\left[\bar{x}_2|\bar{x}_1,\bar{x}_2<\bar{x}_1\right]\right\}.$$

We have shown that $\overline{x}_2<\bar{x}_1$ if $\theta>\underline{\theta}$, and $\overline{x}_2=\bar{x}_1$ otherwise, with $\underline{\theta}$ defined in (2.3.2). The indifference condition can then be rewritten as

$$p_1=\underbrace{(1-\kappa)\mu+\kappa\overline{x}_1}_{\text{expected quality}}+\alpha\underbrace{\left(\bar{N}_1\left(\overline{x}_1\right)+\bar{N}_2\left(\overline{x}_1\right)\right)}_{\text{expected network size}}-\delta\tau(1-\bar{Q}(\bar{x}_1))\left[\bar{x}_1-\bar{X}_{2c}(\bar{x}_1)\right] \qquad (2.3.6)$$

where, conditional on the consumers being the marginal type in period 1, $\overline{x}_1$, the sum $\left(\bar{N}_1\left(\overline{x}_1\right)+\bar{N}_2\left(\overline{x}_1\right)\right)$ represents the expected adoption:

$$\begin{aligned}\bar{N}_1\left(\overline{x}_1\right) &= \mathbb{E}_\theta\left[\Phi\left(\frac{\theta-\bar{x}_1}{\sigma_h}\right)|\overline{x}_1\right]=\Phi\left(\left(\mu-\overline{x}_1\right)z_1\right), \\ \bar{N}_2\left(\overline{x}_1\right) &= \mathbb{E}_\theta\left[\Phi\left(\frac{\theta-\bar{x}_2(\bar{x}_1,\theta)}{\sigma_h}\right)|\overline{x}_1\right]-\bar{N}_1\left(\overline{x}_1\right)=\int_{-\infty}^{\infty}\Phi\left(\frac{\theta-\bar{x}_2(\bar{x}_1,\theta)}{\sigma_h}\right)f\left(\theta|\overline{x}_1\right)d\theta-\bar{N}_1\left(\overline{x}_1\right),\end{aligned}$$

the term $\bar{Q}\left(\overline{x}_1\right)$ represents the probability of no consumption in the second period

$$\bar{Q}\left(\overline{x}_1\right) = \Pr\left[\bar{x}_2(\bar{x}_1,\theta)=\bar{x}_1|\overline{x}_1\right]=\Pr\left[\theta\leq\underline{\theta}|\overline{x}_1\right],$$

which means that $1-\bar{Q}\left(\overline{x}_1\right)$ is the probability that the network will not die in the second period. $\bar{X}_{2c}(\bar{x}_1)$ the expected second period threshold:

$$\bar{X}_{2c}(\bar{x}_1) = \mathbb{E}_\theta\left[\bar{x}_2(\bar{x}_1,\theta)|\overline{x}_1,\bar{x}_2<\bar{x}_1\right]=\frac{1}{(1-\bar{Q}(\bar{x}_1))}\int_{\underline{\theta}}^{\infty}\bar{x}_2(\bar{x}_1,\theta)f\left(\theta|\overline{x}_1\right)d\theta$$

where $z_1 = \sqrt{\frac{(1-\rho)}{(1+\rho)\left(\sigma_\theta^2+\sigma_h^2\right)}}$.

The next proposition shows that there are ranges of parameter values for which demand is a well defined function of prices.

Proposition 7. *There exists* $\underline{\rho}$ *and* $\bar{\sigma_h}$ *such that for any* $\rho < \underline{\rho}$ *and* $\sigma_h > \bar{\sigma_h}$*,* $\bar{x}_1$ *and* p_1 *have one-to-one relationship.*

Proof. Let $H(\bar{x}_1)$ be the right-hand side of (2.3.6), that is,

$$\begin{aligned} H(\bar{x}_1) &= (1-\kappa)\mu + \kappa\overline{x}_1 + \alpha\left(\bar{N}_1\left(\overline{x}_1\right) + \bar{N}_2\left(\overline{x}_1\right)\right) - \delta\tau(1-\bar{Q}(\bar{x}_1))\left[\bar{x}_1 - \bar{X}_{2c}(\bar{x}_1)\right] \\ &= (1-\kappa)\mu + \kappa\overline{x}_1 + \alpha\Phi\left(\left(\mu - \overline{x}_1\right)z_1\right) + \int_{\underline{\theta}}^{\infty} A(\bar{x}_1,\theta) f\left(\theta|\overline{x}_1\right)d\theta, \end{aligned}$$

where $A(\bar{x}_1,\theta) = \alpha N_2\left(\overline{x}_1,\theta\right) - \delta\tau(\bar{x}_1 - \bar{x}_2\left(\overline{x}_1,\theta\right))$. Then there exists unique equilibrium cutoff $\bar{x}_1$ *for any* p_1 if and only if $H(\bar{x}_1)$ is monotonic in $\bar{x}_1$. Since $A(\bar{x}_1,\underline{\theta}) = 0$, by Leibniz's rule, we have

$$\frac{dH(\bar{x}_1)}{d\bar{x}_1} = \kappa - z_1\alpha\phi((\mu-\bar{x}_1)z_1) + \int_{\underline{\theta}}^{\infty} \frac{\partial\left[A(\bar{x}_1,\theta)f\left(\theta|\overline{x}_1\right)\right]}{\partial\bar{x}_1}d\theta, \tag{2.3.7}$$

where

$$\begin{aligned} \frac{\partial\left[A(\bar{x}_1,\theta)f\left(\theta|\overline{x}_1\right)\right]}{\partial\bar{x}_1} &= \left[\frac{\alpha}{\sigma_h}\left\{-\phi_2\frac{\partial\bar{x}_2}{\partial\bar{x}_1} + \phi_1\right\} - \delta\tau\left\{1 - \frac{\partial\bar{x}_2}{\partial\bar{x}_1}\right\}\right] f\left(\theta|\overline{x}_1\right) + \\ &\quad + \left[\alpha\left\{\Phi_2 - \Phi_1\right\} - \delta\tau\left\{\bar{x}_1 - \bar{x}_2\right\}\right]\frac{\rho(\theta - ((1-\rho)\mu + \rho\bar{x}_1)}{(1-\rho)\sigma_\theta^2} f\left(\theta|\overline{x}_1\right), \end{aligned} \tag{2.3.8}$$

where $\Phi_t = \Phi\left(\frac{\theta-\bar{x}_t}{\sigma_h}\right)$ and $\phi_t = \phi\left(\frac{\theta-\bar{x}_t}{\sigma_h}\right)$.When $\rho \to 0$, the second term of the right-hand side of (2.3.8) converges to zero. This is because both $\Phi_2 - \Phi_1$ and $\bar{x}_1 - \bar{x}_2$ are bounded above.Define $Z(\bar{x}_1,\bar{x}_2) = c - (1-\tau)\theta - \tau\bar{x}_2 - \alpha N_1(\bar{x}_1) - \left(2\alpha - \frac{\tau\sigma_h}{\phi_2}\right)N_2(\bar{x}_1,\bar{x}_2)$. (This is from the first-order condition (2.3.3)). Then

$$\frac{\partial \bar{x}_2}{\partial \bar{x}_1} = -\frac{\frac{\partial Z}{\partial \bar{x}_1}}{\frac{\partial Z}{\partial \bar{x}_2}} = -\frac{\left(\frac{\tau}{\phi_2} - \frac{\alpha}{\sigma_h}\right)\phi_1}{2\tau - 2\frac{\alpha}{\sigma_h}\phi_2 + \tau\gamma(\bar{x}_2)\frac{N_2(\bar{x}_2)}{\phi_2}}.$$

When $\rho \to 0$ and $\sigma_h \to \infty$, its limit is

$$\lim_{\sigma_h \to \infty} \frac{\partial \bar{x}_2}{\partial \bar{x}_1} = \lim_{\sigma_h \to \infty} \frac{\phi_1}{2\phi_2} = \frac{1}{2}.$$

Hence the first term of the right-hand side of (2.3.8) converges to zero as well. Therefore, (2.3.7) is positive, which completes the proof.

. $\square$

First Period: Network's problem Given the first-period demand $\bar{x}_1(p_1)$, and second-period behavior $(\bar{x}_2(\bar{x}_1, \theta), p_2(\bar{x}_1, \theta))$, the firm chooses the first-period price which maximizes the discounted sum of expected profit, that is,

$$\begin{aligned} p_1^* &= \arg\max_{p_1} \mathbb{E}_\theta \left[\Pi_1(p_1) + \delta\Pi_2(\bar{x}_1(p_1), \bar{x}_2(\bar{x}_1, \theta), \theta)\right], \\ &= \arg\max_{p_1} (p_1 - c)\mathbb{E}_\theta \left[N_1(p_1)\right] + \delta\mathbb{E}_\theta \left[(p_2(\bar{x}_1(p_1), \theta) - c)N_2(\bar{x}_1(p_1), \bar{x}_2(\bar{x}_1, \theta), \theta)\right] \end{aligned}$$

subject to (2.3.3), (2.3.4), and (2.3.6).

Equivalently, by Proposition 7, we can rewrite the firm's problem as one in which the firm chooses the optimal cutoff $\bar{x}_1$:8

$$\begin{aligned} \bar{x}_1^* &= \arg\max_{\bar{x}_1} \mathbb{E}_\theta \left[\Pi_1(\bar{x}_1) + \delta\Pi_2(\bar{x}_1, \theta)\right] \\ &= \arg\max_{\bar{x}_1} (p_1(\bar{x}_1) - c)\Phi\left(\frac{\mu - \bar{x}_1}{\sqrt{\sigma_\theta^2 + \sigma_h^2}}\right) + \delta\mathbb{E}_\theta \left[(p_2(\bar{x}_1, \theta) - c)N_2(\bar{x}_1, \bar{x}_2(\bar{x}_1, \theta), \theta)\right], \end{aligned}$$

since $\mathbb{E}_\theta\left[N_1(\bar{x}_1)\right] = \mathbb{E}_\theta\left[\Phi\left(\gamma(\bar{x}_1)\right)\right] = \Phi\left(\frac{\mu-\bar{x}_1}{\sqrt{\sigma_\theta^2+\sigma_h^2}}\right)$.

2.4 Numerical Simulation

In this Section we provide a numerical exercise to illustrate the rich dynamics of the full specification and compare with the benchmark cases of herding, pure durable goods and static models. We start with a description of the algorithm used to compute the equilibrium.

2.4.1 Algorithm for the computation of the equilibrium

The algorithm used to compute the equilibrium profile $[(p_1(\bar{x}_1^*), p_2(\bar{x}_1^*, \theta)), (\bar{x}_1^*, \bar{x}_2(\bar{x}_1^*, \theta)]$ involves the following steps:

1. Given a first-period cutoff $\overline{x}_1$ and fundamental θ, find the second-period cutoff $\overline{x}_2\left(\overline{x}_1, \theta\right)$ in the following way: let $\tilde{x}_2(\bar{x}_1, \theta)$ be the solution to the first-order condition (2.3.3):

$$(1-\tau)\theta + \tau\tilde{x}_2 = c - \alpha N_1 + \left(\frac{\tau\sigma_h}{\phi\left(\frac{\theta-\tilde{x}_2}{\sigma_h}\right)} - 2\alpha\right) N_2(\tilde{x}_2).$$

Then $\bar{x}_2\left(\theta, \overline{x}_1\right)$ is given by

$$\overline{x}_2\left(\bar{x}_1, \theta\right) = \begin{cases} \overline{x}_1 & \text{if } \theta \leq \underline{\theta}, \\ \tilde{x}_2\left(\bar{x}_1, \theta\right) & \text{if } \theta > \underline{\theta}, \end{cases}$$

where $\underline{\theta}$ is given by (2.3.2):

$$(1-\tau)\underline{\theta} + \tau\bar{x}_1 + \alpha\Phi\left(\frac{\underline{\theta} - \overline{x}_1}{\sigma_h}\right) = c.$$

2. The second-period price $p_2(\bar{x}_1, \theta)$ is given by (2.3.4):

$$p_2 = c + \underbrace{\left(\frac{\tau \sigma_h}{\phi\left(\frac{\theta - \overline{x}_2}{\sigma_h}\right)} - \alpha \right) N_2(\overline{x}_2)}_{\text{markup}}.$$

Note that $p_2 = c$ if and only if $\theta \leq \underline{\theta}$.

3. From the indifference condition (2.3.6), determine p_1 as a function of $\bar{x}_1$:

$$\begin{aligned} p_1 &= (1-\kappa)\mu + \kappa\overline{x}_1 + \alpha(\bar{N}_1(\overline{x}_1) + \bar{N}_2(\overline{x}_1)) - \delta\tau(1 - \bar{Q}(\bar{x}_1))\left[\bar{x}_1 - \bar{X}_{2c}(\bar{x}_1)\right] \\ &= (1-\kappa)\mu + \kappa\overline{x}_1 \\ &\quad + \int_{\underline{\theta}}^{\infty} \left[\alpha\Phi\left(\frac{\theta - \bar{x}_2(\bar{x}_1, \theta)}{\sigma_h}\right) - \delta\tau(\bar{x}_1 - \bar{x}_2(\bar{x}_1, \theta))\right] f(\theta|\overline{x}_1)\, d\theta. \end{aligned}$$

4. Finally, find the firm's optimal first-period cutoff $\bar{x}_1$ that maximizes the lifetime expected profit of the network: (this is equivalent to find the price p_1, given that p_1 and $\bar{x}_1$ have one-to-one relationship)

$$\begin{aligned} \bar{x}_1^* &= \arg\max_{\bar{x}_1} \mathbb{E}_\theta\left[\Pi_1(\bar{x}_1) + \delta\Pi_2(\bar{x}_1, \theta)\right] \\ &= \arg\max_{\bar{x}_1} (p_1(\bar{x}_1) - c)\Phi\left(\frac{\mu - \bar{x}_1}{\sqrt{\sigma_\theta^2 + \sigma_h^2}}\right) + \delta\mathbb{E}_\theta[(p_2(\bar{x}_1, \theta) - c)N_2(\bar{x}_1, \theta)] \end{aligned}$$

subject to (2.3.3), (2.3.4), and (2.3.6).

2.4.2 Simulation

In this section we discuss the results of simulations for a specific calibration of the parameters. Recall that the marginal cost for the monopolist is assumed to be zero throughout this paper and that the utility of consumer i at time t is

$$U_{i,t} = v_i + \alpha\,(n_1 + n_2) - p_t.$$

We consider the following parameterization:

Parameter	Value
μ	0.5
σ_θ	2
σ_h	4
α	2.5
δ	0.7

Table 2.1: Parameterization of the model.

We set $\tau = 0.5$ unless in the cases where signals are perfectly informative about the intrinsic value (i.e., $\tau = 1$).

2.4.2.1 Model with Network Effects

First, we present the equilibrium outcome of the dynamic model with both network effects (characterized by $\alpha = 2.5$) and the informational (herding) effect (characterized by $\tau < 1$). Optimal first period price for the parameterization above is $p_1^* = 2.42$. Figure (2.1) shows the second period price as a function of θ, where we can see that second period prices will be larger than p_1^* with positive probability. If the fundamental is below $\underline{\theta}(\overline{x}_1) \simeq -3$, we have $\bar{x}_2 = \bar{x}_1$, which implies that prices are set to the marginal cost level and there will be zero demand in the second period.

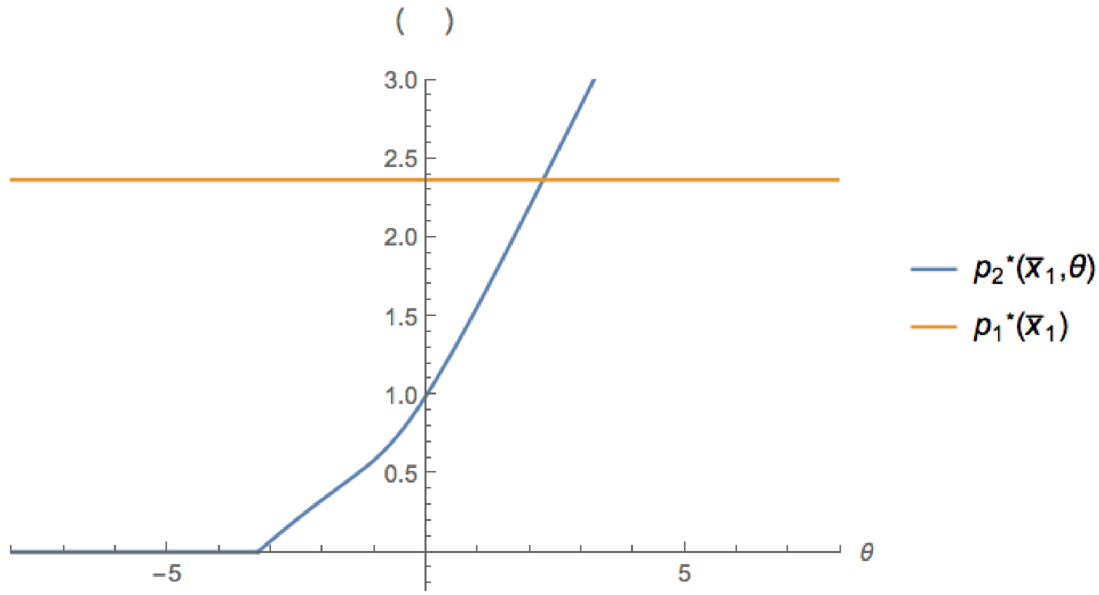


Figure 2.1: Equilibrium prices p_1 and $p_2\,(\theta)$ for the model with network effects and information acquisition ($\tau = 0.5$)

Figure (2.2) shows the adoption in both periods as a function of θ. First period consumption is a positive function of θ and for values of θ around zero, there is substantial delay. We see the occurrence of tipping: there is a minimum level of adoption in the first period that will induce consumption in the second period, otherwise the network will die in period two. This is the case even with the monopolist posting a price at its marginal cost (zero in the current case). The total demand shows a kink at $\theta = \underline{\theta}$, with the "critical mass effect": the total demand is very sensitive to θ once $\theta \geq \underline{\theta}$. Consumer's collective behavior shows the positive feedback effect, in the sense that the second-period reaction crucially depends on the size of the first-period consumption.

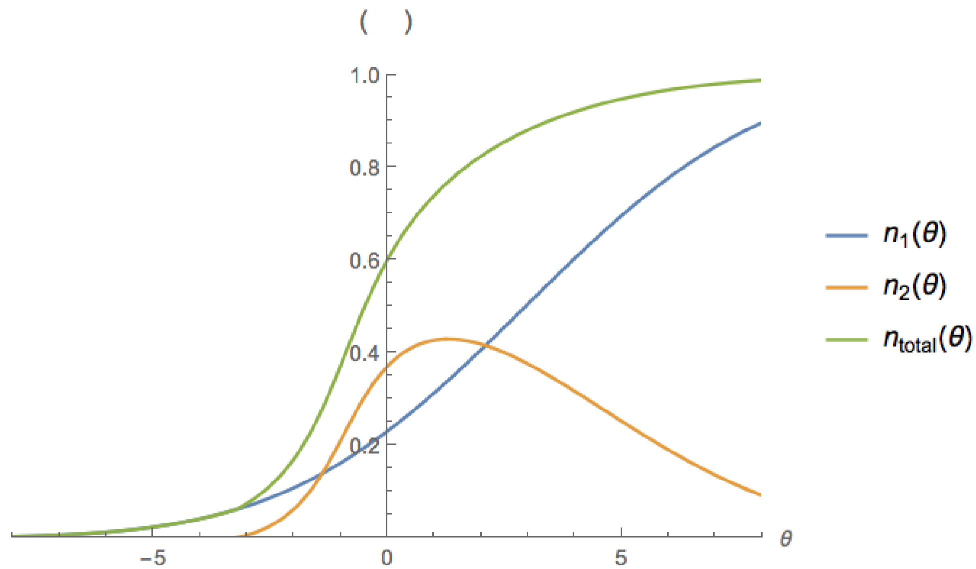


Figure 2.2: Equilibrium share of consumers joining the network in each period and the total share of consumers joining the network as a function of the fundamental θ for the cases the full model with network effects and information acquisition ($\tau = 0.5$).

We now turn to the case of the dynamic model with network effects and full information about v_i (i.e., $\tau = 1$). Figure (2.3) describes the equilibrium price in both periods. The first period price in this case is $p_1^* = 3.38$, nearly 40% larger than in the case without full information about v_i. On the other hand, the function describing second period prices is flatter than the one in the former case. Second period prices can still be larger than p_1^*, but with lower probability than in the case with $\tau < 1$. It is worth noting that the model with complete information about τ does not exhibit the occurrence of tipping, as we se in Figure (2.4).

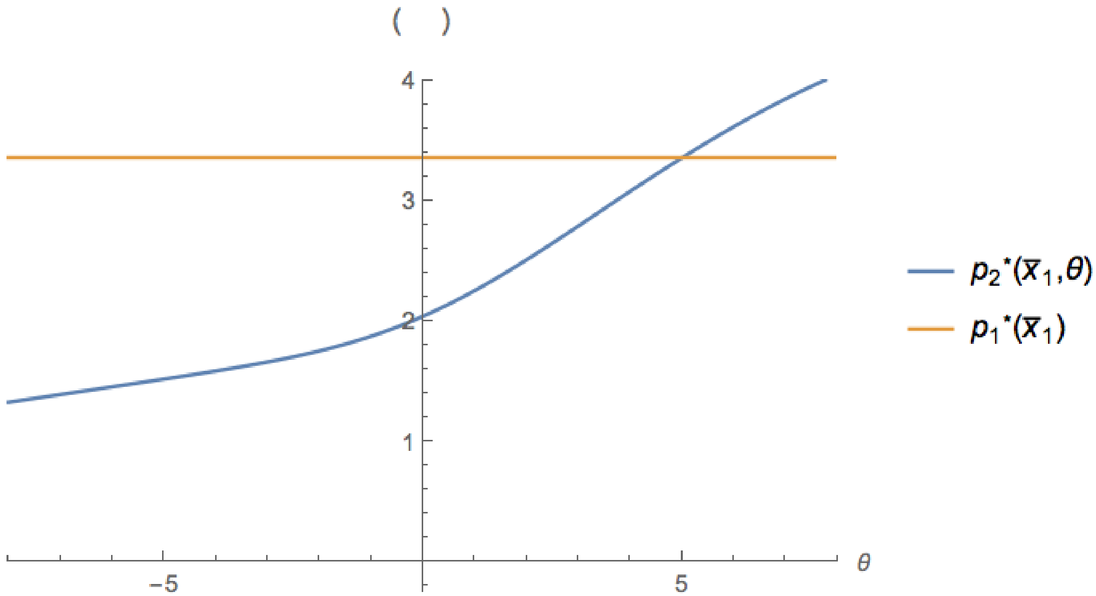


Figure 2.3: Equilibrium prices p_1 and $p_2(\theta)$ for the model with network effects and signals perfectly informative about intrinsic value of the good ($\tau = 1$).

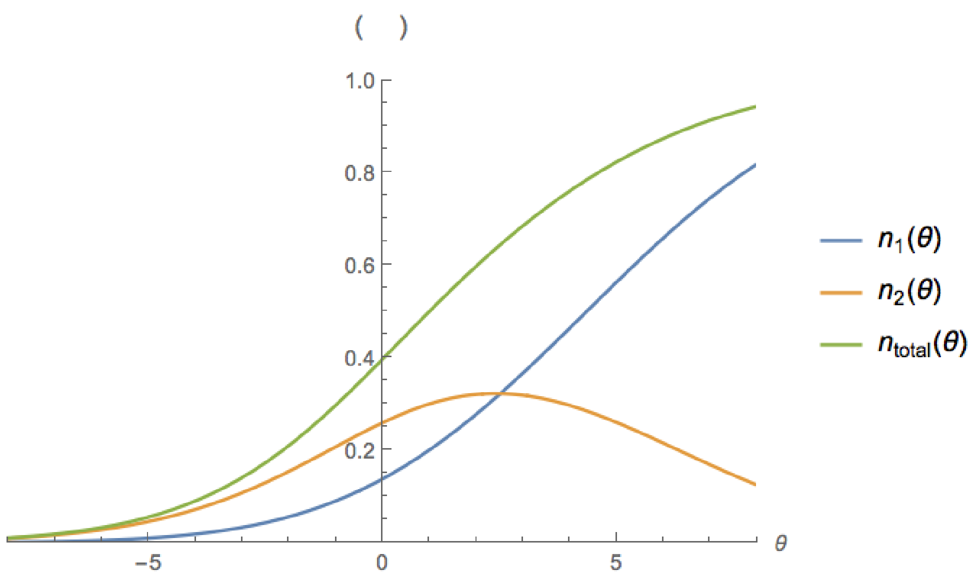


Figure 2.4: Equilibrium share of consumers joining the network in each period and the total share of consumers joining the network as a function of the fundamental θ for the model with network effects and signals perfectly informative about intrinsic value of the good ($\tau = 1$).

2.4.2.2 Static Model with Network Effects

In the static version of the model, the optimal monopolist's price is $p^{static} = 3.37$. Only for low values of θ, adoption in the static model will be higher than that of the full specification presented in Figure (2.2).

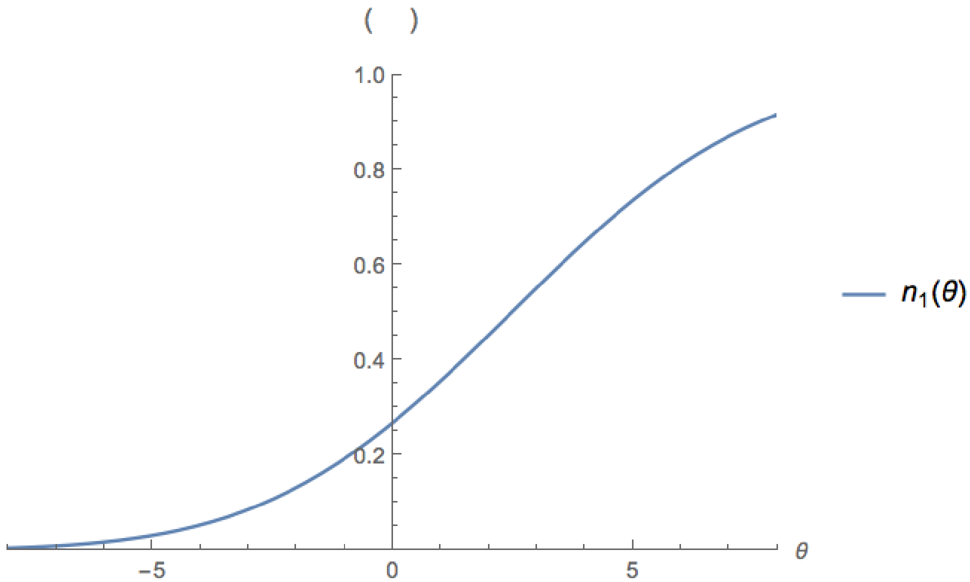


Figure 2.5: Equilibrium share of consumers joining the network in as a function of the fundamental θ in the static model with network effects and signals perfectly informative about the intrinsic value of the good ($\tau = 1$).

2.4.2.3 Herding Model

Next we consider the case in which there is no direct network effect ($\alpha = 0$), but still consumers do not have complete information about their intrinsic value ($\tau < 1$) so there is a informational herding effect.

Figure (2.6) shows the equilibrium prices p_1 and $p_2(\theta)$. First period price is similar to the one charged in the model with network effects, but generates much lower consumption. In terms of second period prices, for low values of the fundamental, the firm posts a price equal to its marginal cost (zero); if the fundamental high enough, then the firm posts a price higher than p_1.

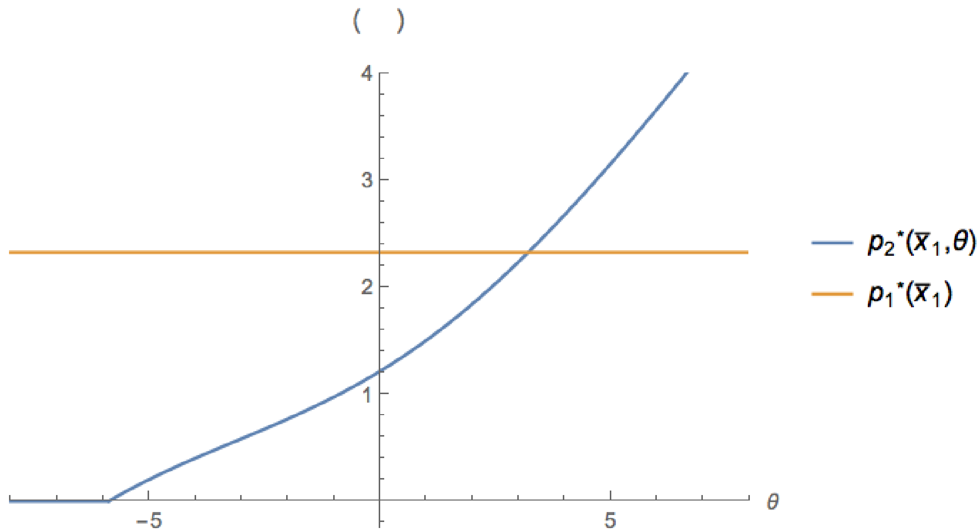


Figure 2.6: Equilibrium prices p_1 and $p_2\left(\theta\right)$ for the herding model.

Figure (2.7) describes the equilibrium demand in each period, and the sum of the demand, as a function of the fundamental θ. Delayed purchases happen more when consumers have "good news" about the quality of the good and adoption is strictly lower than in the case with network effects for the entire range of θ. Also, there is no feature related to critical mass as in the case with network effects. Comparing the results from the pure herding model with the versions with network effects, we conclude that information interact with network effects in order to generate "tipping" in this model.

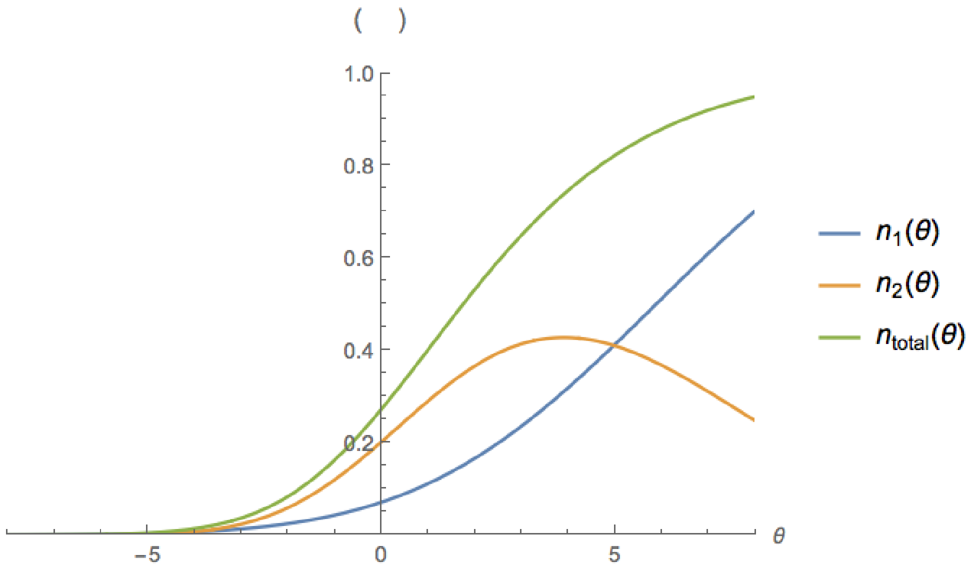


Figure 2.7: Equilibrium share of consumers purchasing the good in each period and the total share of consumers joining the network as a function of the fundamental θ for the case of the herding model ($\tau < 1$ and $\alpha = 0$).

2.4.2.4 Pure Durable Good

The final version examined in this simulation is the case of pure durable goods monopoly, in which there is no network effect ($\alpha = 0$) and consumers are fully informed about their intrinsic value ($\tau = 1$). In this case, the equilibrium behavior is very different from the one with the network effects. Figure (2.8) shows the equilibrium prices p_1 and $p_2(\theta)$. First we observe that there is no increasing price path: the second period price is lower than first period price with probability one. Furthermore, the optimal first-period price in the pure durable goods monopoly is lower than the one with network effects. Figure (2.9) describes the equilibrium demand in each period as a function of the fundamental θ. In this case, delay occurs only due to expectation of lower future prices. There is no feature related to critical mass as in the case with network effects.

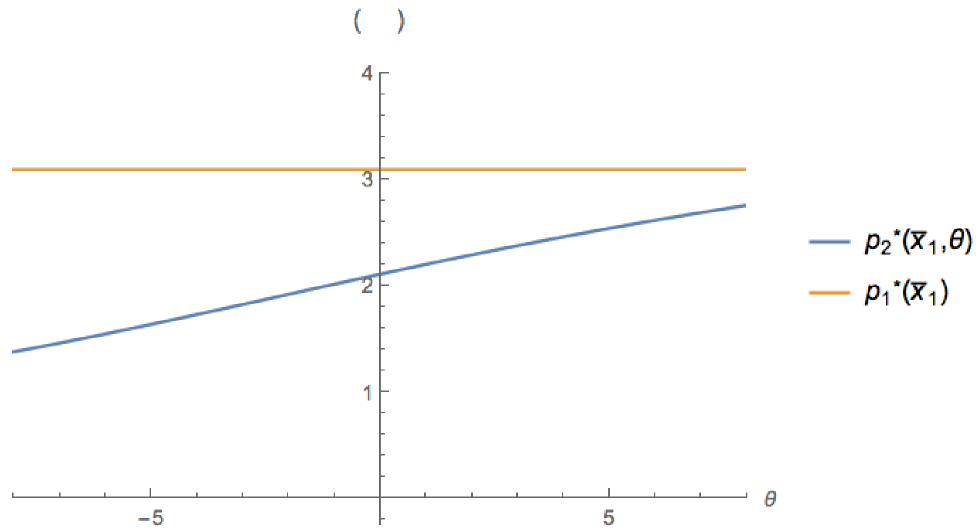


Figure 2.8: Equilibrium prices p_1 and $p_2\,(\theta)$ for the pure durable goods model.

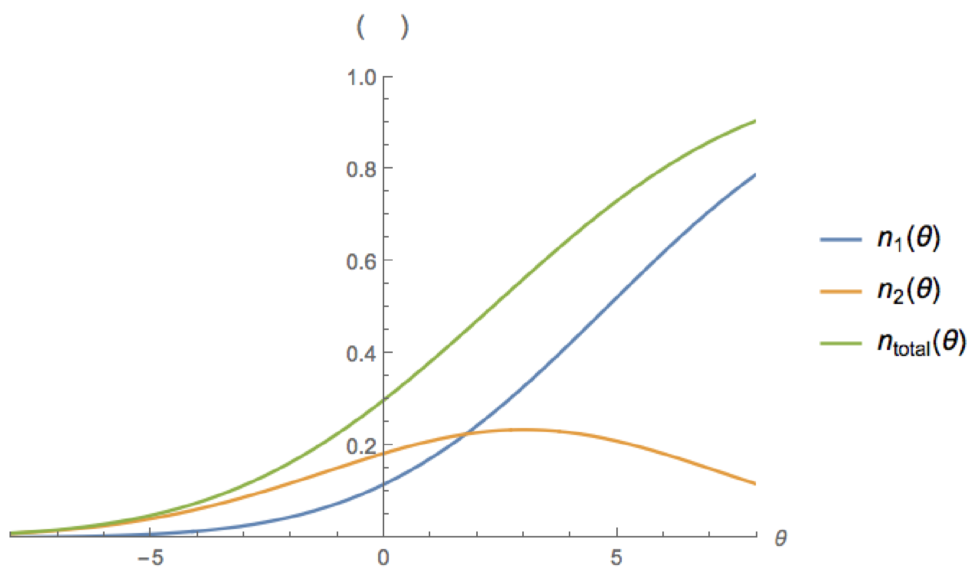


Figure 2.9: Equilibrium share of consumers purchasing the good in each period and the total share of consumers joining the network as a function of the fundamental θ for the case of the pure durable goods model ($\tau = 1$ and $\alpha = 0$).

2.4.2.5 Adoption and profits in the different versions of the model

Table 2 shows that the different assumptions about α and τ can generate potentially very different prices in the first period. Uncertain about the intrinsic value v_i seems to induce lower first period prices. Moreover, when v_i is unknown, the existence of network effects permit the firm to charge a slightly higher price compared to the pure herding model. Interestingly, the higher prices for the case with network effects are associated to a much larger adoption in the first period (see Figures 2 and 7).

Model	Price in the first period
Network effects ($\alpha > 0$) with unknown v_i ($\tau < 1$)	2.41
Network effects ($\alpha > 0$) with v_i known ($\tau = 1$)	3.38
Static model with v_i unknown ($\tau < 1$)	2.06
Static model with v_i known ($\tau = 1$)	3.37
Herding model ($\alpha = 0, \tau < 1$)	2.37
Pure durable goods ($\alpha = 0, \tau = 1$)	3.08

Table 2.2: First period equilibrium prices in the different versions of the model.

Figure (2.10) presents the total adoption for different specifications of the model as a function of the fundamental parameter θ. The model with network effects and uncertainty about v_i generates a considerably larger consumer base than the different benchmarks for a wide range of values of θ. The only specification than is dominated by the full model in terms of adoption is the case of herding. In the other three cases (network effects with v_i known, pure durable goods and static model with network effects and v_i known), total adoption will be larger than that in the full specification when the realized fundamental is sufficiently low.

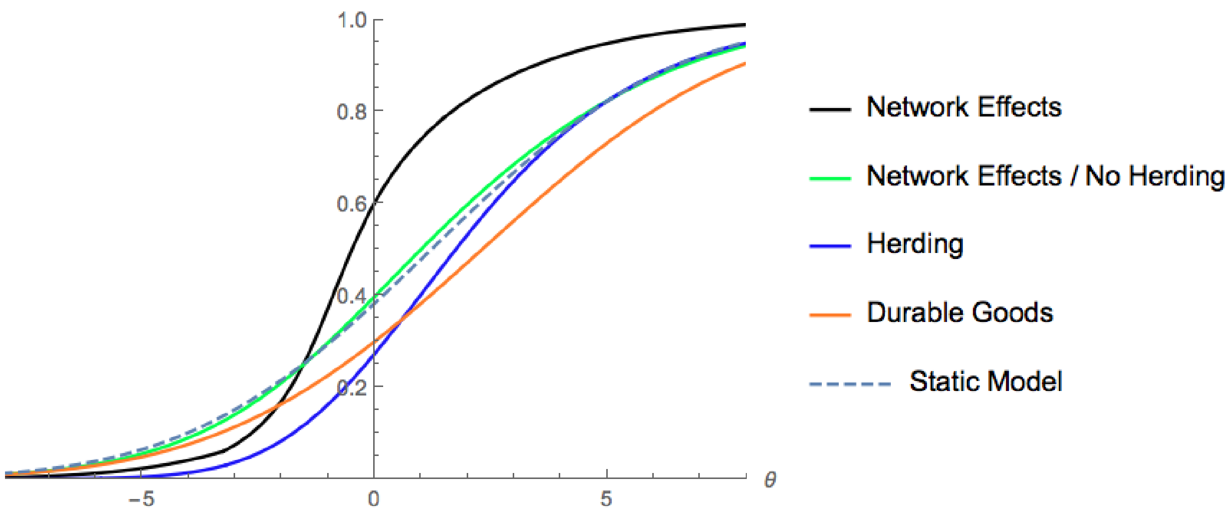


Figure 2.10: Total adoption in the different specifications of the model.

The larger adoption in the case of the dynamic model with network effects does not translate into higher profits when we compare with the benchmark static model. Only for intermediate values of θ, the monopolist profit in the model with delay will be higher than in the static model. For extreme realizations of θ, the static model will be more profitable for different reasons depending in which tail we are located. When θ is low, then the lower profits in the full specification are simply because of small number of adopters. On the other hand, when θ is high, since first period price is set based on expectations, the realized adoption in the first period will be high and there will be just not many consumers left in the second period in order to allow the monopolistic to extract a high rent.

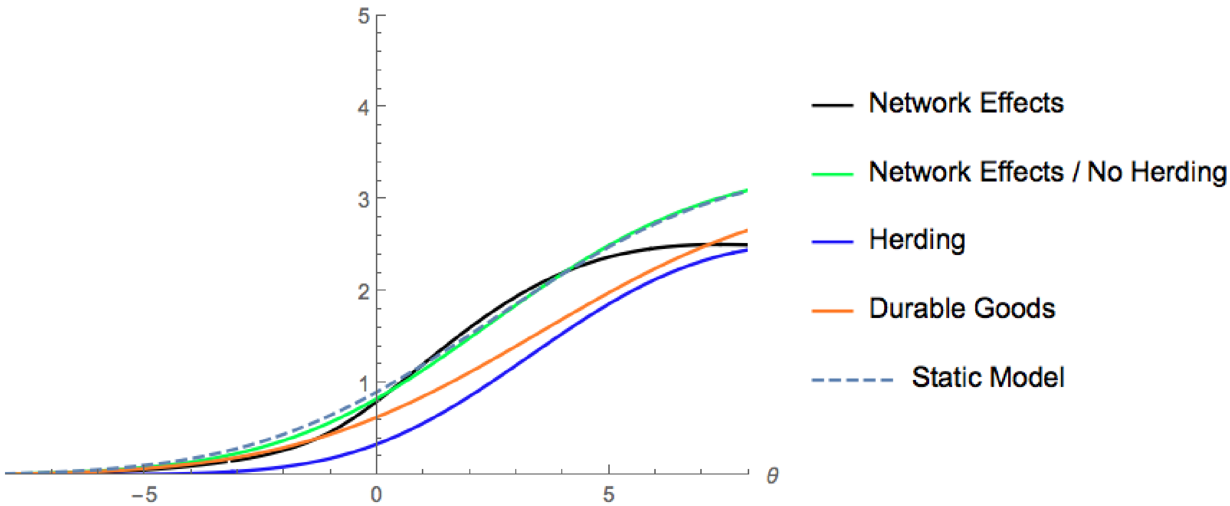


Figure 2.11: Total profits of the monopolist in the different specifications of the model.

Given the shape of the profit function for the case with network effects with complete and incomplete information about v_i, one can imagine that investment in advertisement which would increase the information about v_i acts as an insurance, increasing the profits of the firm in the case of extreme realizations of the fundamental.

2.5 Concluding Remarks

We consider a market with network effects in which consumers can decide if and when to purchase a good with network effects. We develop a simple yet rich model that is capable of exploring many issues in the economics of network effects such as introductory pricing and early critical mass for platform survival. Our specification nests the pure durables goods and herding models. We derive uniqueness conditions for equilibrium of the dynamic model with endogenous delayed purchases.

Comparing the adoption in the versions with network effects and in the pure herding model, we conclude that incomplete information about the intrinsic value of the good is important in order to generate "tipping" in this model. Numerical results illustrate the richness of possible outcomes in the dynamic model and highlight the differences relative to the benchmark cases in terms of pricing, adoption and profits.

We leave for future research an extension of this model to the case of platform competition in the same environment, where consumers can choose which network to join and when.

Chapter 3

Endogenous Vertical Differentiation

Product improvement is an important competitive strategy. It differentiates products, and sets the stage for price competition. Twin considerations of market uncertainty and strategic uncertainty shape incentives for product improvement. First, firms are unclear about how consumers value their products. Second, firms are unsure about rivals' quality decisions. The information structure of the market determines the interplay of these kinds of uncertainty.

The residential market for broadband Internet service is a case in point. A typical local market is a duopoly served by a cable TV company (e.g. Comcast or Time Warner) and a telephone company (e.g. AT&T or Verizon). Cable companies and telephone companies recently pursued strategies of upgrading their networks in selected markets. Cable companies widely adopted the DOCSIS 3.0 standard permitting higher bandwidth data transmission, while telephone companies more selectively invested to extend fiber optic plant to customer premises (e.g. Verizon FIOS) or to a neighborhood interface (e.g. AT&T U-Verse). These network upgrades dramatically improved download and upload speeds. The fiber optic upgrades of the telephone companies arguably were more costly but provided higher-quality service than the improved cable broadband service. In deciding where and when to invest, telephone companies like AT&T and Verizon were uncertain about market acceptance, and in some cases unsure of how quickly cable rivals would deploy the new DOCSIS 3.0 standard. Similarly, in deciding the speed and scope of DOCSIS 3.0 deployment, cable TV companies needed to assess market conditions and rivalry with telephone companies.

We develop a model of product quality competition that focuses on how information structure determines equilibrium outcomes. We focus on the case of duopoly entry into a market with the option to upgrade the product quality. Market uncertainty is captured by the idea that, while firms are symmetric ex ante, one or the other will have an ex post brand advantage of uncertain magnitude, Γ. Strategic uncertainty is captured by the idea that firms make their product quality decisions with privately-observed noisy information about brand advantage. We consider the case of limited preference heterogeneity, which leads to a corner solution in the Bertrand price equilibrium following the quality upgrade stage. In this equilibrium, the firm selling higher quality good sets its price equal to the quality difference, and the rival firm sets its price to zero.

We investigate equilibrium strategies under three information structures associated to Γ: (i) complete information; (ii) incomplete information with no signal; and (iii) incomplete information with each firm receiving privately correlated signals about Γ.

In the case of complete information about Γ, there is an interval around zero with two pure strategy Nash equilibria each having one firm investing in quality and the other not, as well as a strictly mixed strategy Nash equilibrium. The profits in the pure strategy equilibria are very different even when the quality difference is arbitrarily small. In the mixed strategy equilibrium, miscoordination results either in wasteful investment or failure to improve.

The case of incomplete information deals with the criticism that firms might be unsure about who has the advantage and its size. Except in a non-generic special case, sufficient uncertainty about brand advantage will lead to a unique equilibrium. However, the equilibrium outcomes in the regions of unique equilibrium result in either wasteful investment or failure to improve. In the equilibrium with investment by the two firms, the loser firm will always have negative profits. Interestingly, the winning firm can also end up with negative profit, which will occur when the realized brand advantage is sufficiently small. Since price charged by the winning firm is proportional to the excess quality, a small brand advantage might not cover the investment cost. This provides some intuition for why we need σ_Γ high for uniqueness: firms need to believe that there is enough chance to get favorable values of Γ. The mixed strategy BNE converges to the complete information MSNE as market uncertainty vanishes. Most qualitative results remain unchanged when we modify the game imposing prices to be defined before the realization of Γ.

In the case of incomplete information with each firm receiving privately correlated signals about Γ, this strategic interaction becomes a global game[1]. We show that equilibrium product improvement decisions are unique if and only if market uncertainty is sufficiently high relative to strategic uncertainty. One important thing about uniqueness of equilibrium in this case is the ability that one firm has to use its own signal to update the distribution of signals of the other firm. The higher the correlation between the signals of the two firms, the more accurate will be this prediction and the lower will be the probability of miscoordination.

This paper is related to the literature on entry games, interpreted in a broad sense to include the introduction of a new product or the adoption of a new technology. We can divide the entry models in two groups: identical firms (e.g. Mankiw and Whinston (1986), Bresnahan and Reiss (1988, 1990, 1991), Berry and Waldfogel (1999)) and heterogeneous firms (e.g. Berry (1992), Mazzeo (2003) and Seim (2005)). In our model, firms are ex ante identical but one or the other will have an ex post brand advantage. In the appendix we consider a variation in which firms have asymmetric payoffs. Multiple equilibria is a

[1]See Carlsson and Van-Damme (1993) and Morris and Shin (2003).

typical problem in entry models, which has motivated new estimation methods in the form of moment inequalities (e.g. Ciliberto and Tamer (2009), Andrews and Jia (2012) and Pakes, Porter, Ho and Ishii (2015)).

We build on and contribute to the literature on investments in quality upgrade. Shaked and Sutton (1982) show that one benefit of quality differentiation is that it relaxes price competition. Gabszewicz and Thisse (1980) as well as Shaked and Sutton (1982, 1983 and 1984) are the seminal works in vertical product differentiation. These models of quality choice assume that firms invest in quality before the market competition stage, which implies that the cost of product improvement is sunk when competition takes place. Gilbert and Riordan (2007) consider a winner-take-all duopoly model in which firms invest in product improvement. They discuss different equilibria that may arise depending on the vertical relations in the market and the options for equilibrium selection. In this paper we revisit the problem that the coordination game played by the firms may have multiple equilibria with very different equilibrium outcomes in a winner-take-all market. Multiplicity of equilibria is problematic for counterfactuals and policy analysis in the absence of a reliable selection criterion[2]. Investment decisions in these circumstances rely to a large extent on coordinated behavior of the firms, which implies that the information structure plays a crucial role in the determination of which equilibrium will be played. We focus on the interplay between information and actions in a quality improvement game, exploring how beliefs are determined and the conditions under which equilibrium is unique.

Methodologically, the game in our main specification is a global game with common values and private correlated signals. Morris and Shin (2003) survey the applied literature using global games approach in finance and macroeconomics. The games used to study models of bank runs, currency crises and herding behavior all feature strategic comple-

[2]Gilbert and Riordan (2007) and Besanko, Doraszelski and Kryukov (2014) examine different solutions to reduce the number of equilibria.

mentarities. Global games with strategic substitution like the one studied in this paper have been the subject of much fewer studies. Applications of global games to industrial organization is still scarce and the only existing work that we are aware are Argenziano (2008) and Julien and Pavan (2016).

Carlsson and van Damme (1993), hereafter CvD, show for the case of 2×2 games that if there are strictly dominant regions associated to each possible action , then if there are two pure strategy Nash equilibria in the underlying complete information game, for sufficiently small noise iterative elimination of strictly dominated strategies selects a unique equilibrium, which is the Harsanyi and Selten (1988) risk dominant equilibrium.[3] Some assumptions used in the main theorem in CvD are not satisfied in our main specification[4], which motivates our approach for a tailored proof.

The remainder of this paper is organized as follows. Section 2 provides a simple example of a quality upgrade game and its link to global games to introduce the problem studied in the paper. Section 3 presents the model and equilibrium analysis under different information structures. Section 4 concludes. Some of the proofs and detailed computations are contained in the Appendix.

3.1 Illustrative Example

The purpose of this Section is to provide a simple example of a model of quality upgrade and global game. Let $g(\Gamma)$ be the $2x2$ game depicted in Figure 1 and consider the class of games $\{g(\Gamma)\}_{\Gamma \in \mathbb{R}}$.

[3] See footnote 5 for explanation of risk dominant equilibrium.

[4] More specifically, Assumption 1 in CvD requires the payoffs to be of class C^1, which is not the case in our model. Assumption 4 in CvD consider the support of the noise to be bounded. This assumption can be dispensed with as long as payoffs are bounded. None of these hold in the model considered in this paper.

		Firm 2	
		$q_2 = 0$	$q_2 = 1$
Firm 1	$q_1 = 0$	1, 1	0, Γ
	$q_1 = 1$	Γ, 0	$\Gamma - c$, $\Gamma - c$

Figure 3.1: Simple quality upgrade game.

The individual payoff is normalized to one when both firms decide not to upgrade their products, Γ when the firm upgrades alone and $\Gamma - c$ when both firms upgrade. We interpret Γ as the monopoly profit generated when the firm invests in quality upgrade and $c \in (0, 1)$ as the discount in the profit that comes from competition when both firms are upgrading.

In the complete information game, invest $((q_1, q_2) = (1, 1))$ is a dominant strategy for both firms when $\Gamma > 1$. If $\Gamma < c$, then not invest $((q_1, q_2) = (0, 0))$ is dominant. For values of Γ between c and 1, there will be two pure strategy Nash equilibria and one equilibrium in mixed strategies.

Now consider the incomplete information game, in which each firm will observe the realized game with some noise and then chooses whether or not to invest in quality upgrade. More specifically, assume that the fundamental parameter Γ is normally distributed with mean μ and variance σ_Γ^2. We interpret the distribution of Γ as uncertainty about the success of quality improvement. Moreover, firms do not observe Γ directly and instead, each firm receives a signal $x_i = \Gamma + \varepsilon_i$, where $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$ with ε_1 and ε_2 independent of each other and of Γ. The signals are jointly normally distributed with correlation coefficient $\rho = \frac{\sigma_\Gamma^2}{\sigma_\Gamma^2 + \sigma_\varepsilon^2}$. This model is a global game with private correlated signals and common values (Carlsson and van Damme (1994)).[5]

[5]In Appendix B we show a similar model with private correlated values.

This game features both fundamental uncertainty (Γ is not known) and strategic uncertainty (a firm does not know the opponent's action). Similarly to the game of complete information, if the signals are precise about the realization of Γ (i.e., σ_ε^2 is low), high signals will induce firms to invest as a dominant strategy and low signals will lead not to invest. We start by considering equilibria with cutoff strategies, in which there is a threshold $\overline{x}_i$ below which firm i will not invest and above which the firm will invest.

The expected payoff to Firm 1 from not investing is $Pr\left(x_2 \leq \overline{x}_2 | x_1\right)$. When investing, the expected payoff to Firm 1 is

$$\mathbb{E}\left[\Gamma | x_1\right] - c\left[1 - Pr\left(x_2 \leq \overline{x}_2 | x_1\right)\right]$$

where

$$\mathbb{E}\left[\Gamma | x_1\right] = \rho x_1 + (1-\rho)\,\mu$$

and

$$Pr\left(x_2 \leq \overline{x}_2 | x_1\right) = \Phi\left(\frac{\overline{x}_2 - \rho x_1 - (1-\rho)\,\mu}{\sigma_\Gamma \sqrt{\frac{1-\rho^2}{\rho}}}\right).$$

Firm 1 is indifferent between investing or not if

$$\rho \overline{x}_1 + (1-\rho)\,\mu - c - (1-c)\,\Phi\left(\frac{\overline{x}_2 - \rho \overline{x}_1 - (1-\rho)\,\mu}{\sigma_\Gamma \sqrt{\frac{1-\rho^2}{\rho}}}\right) = 0.$$

The left hand side will be negative for $\overline{x}_1$ sufficiently low and positive tor $\overline{x}_1$ sufficiently high. Hence, for any $\overline{x}_2$, existence of $\overline{x}_1$ satisfying this indifference condition is guaranteed from continuity of the left hand side.

If $\overline{x}_2 \to \infty$, this indifference condition becomes $\overline{x}_1^U = \frac{1-(1-\rho)\mu}{\rho}$ and when $\overline{x}_2 \to -\infty$, $\overline{x}_1^L = \frac{c-(1-\rho)\mu}{\rho}$, where it is straightforward to see that $\overline{x}_1^U > \overline{x}_1^L$. The slope of the best response of Firm 1 is always positive (i.e., a game with strategic complementarity):

$$b_1'(x_2) = \frac{\frac{(1-c)}{\sigma_\Gamma\sqrt{\frac{1-\rho^2}{\rho}}}\phi\left(\frac{x_2-\rho\overline{x}_1-(1-\rho)\mu}{\sigma_\Gamma\sqrt{\frac{1-\rho^2}{\rho}}}\right)}{\rho+\frac{(1-c)\rho}{\sigma_\Gamma\sqrt{\frac{1-\rho^2}{\rho}}}\phi\left(\frac{x_2-\rho\overline{x}_1-(1-\rho)\mu}{\sigma_\Gamma\sqrt{\frac{1-\rho^2}{\rho}}}\right)} > 0.$$

The condition for uniqueness in this case requires that the slope of the best response $b_1(x_2)$ is smaller than 1, i.e, the best response $b_1(x_2)$ crosses the 45^o line only once. This is true if

$$\frac{\sigma_\Gamma}{(1-c)} > \frac{1}{\sqrt{2\pi\rho}}\sqrt{\frac{1-\rho}{1+\rho}}. \tag{3.1.1}$$

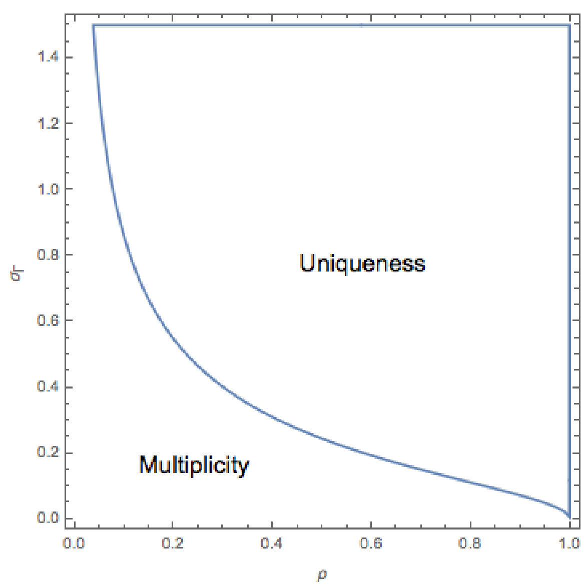


Figure 3.2: Combinations of (ρ, σ_Γ) satisfying uniqueness condition for different values of σ_Γ for the symmetric model (Figure (a)) and asymmetric model (Figure (b)). Shaded area represents the uniqueness region for the case of $c = 0.25$.

Condition (3.1.1) states that in this simple duopoly model of quality competition, equilibrium product improvement decisions are unique if and only if market uncertainty (measured by σ_Γ) is sufficiently high relative to strategic uncertainty (measured by ρ). If equilibrium is unique, because of the symmetry of the game, we must have that $\overline{x}_1 = \overline{x}_2 = \overline{x}$. If we rewrite the indifference condition in this case we have that

$$\rho\overline{x} + (1-\rho)\,\mu - c - (1-c)\,\Phi\left((\overline{x}-\mu)\,\frac{\sqrt{\rho}}{\sigma_\Gamma}\sqrt{\frac{1-\rho}{1+\rho}}\right) = 0.$$

The limit case for $\rho \to 1$ is $\overline{x} = \frac{1+c}{2}$. This threshold value coincides with the value that separates the two regions of risk dominance[6] (Harsanyi and Selten (1988)): invest is risk dominant if $\Gamma > \frac{1+c}{2}$ and not invest is risk dominant if $\Gamma < \frac{1+c}{2}$.

3.2 Information structure and product improvement

The previous Section illustrates the use of global games using a simple game of quality upgrade. The simplicity of the payoff structure carries one inconvenient implication: even the monopolistic payoff can be unboundedly negative. Another aspect that is noteworthy in the simplified example is the resulting strategic complementarity: greater opposing action make greater actions more appealing. When we deal with cases of limited consumer

[6]Consider that the strategy pairs (H, H) and (G, G) are the only pure Nash equilibria in the following game:

		Firm 2	
		H	G
Firm 1	H	A,a	C,b
	G	B,c	D,d

We say that strategy pair (G, G) **risk dominates** (H, H) if the product of the deviation losses is highest for (G, G) (Harsanyi and Selten, 1988, Lemma 5.4.4), i.e., if the following inequality holds:

$$(C–D)(c–d) \geq (B–A)(b–a).$$

In our application, we say that strategy profile $(q_1, q_2) = (0, 0)$ risk dominates $(q_1, q_2) = (1, 1)$ if

$$(1-\Gamma)^2 > (\Gamma - c)^2\,,$$

which holds for $\Gamma < \frac{1+c}{2}$. When $\Gamma > \frac{1+c}{2}$, $(q_1, q_2) = (1, 1)$ risk dominates $(q_1, q_2) = (0, 0)$.

heterogeneity, the winner-take-all nature of these markets is more conveniently modeled as a game of strategic substitution. The specification developed in this Section deals with these two aspects.

3.2.1 Market structure

Consider a market with two symmetric firms, indexed by $i \in \{1, 2\}$. Firm i decides whether to improve its product by choosing $q_i \in \{0, 1\}$. The realized profit of Firm 1 is

$$\pi_1 (q_1, q_2) = max \{0, q_1 - q_2 + \Gamma\} - rq_1, \tag{3.2.1}$$

and the profit of Firm 2 is

$$\pi_2 (q_1, q_2) = max \{0, q_2 - q_1 - \Gamma\} - rq_2. \tag{3.2.2}$$

The parameter $r \in (0, 1)$ is the cost of investment in high quality, and $\Gamma \in \mathbb{R}$ is an exogenous advantage of Firm 1.

We offer two interpretations of these profit functions. One interpretation that two upstream firms are competing to supply component input to a downstream customer, and can invest in design improvement prior to price competition. The customer has a gross value of $\gamma_i + q_i$ for Firm i's design, and $\Gamma = \gamma_1 - \gamma_2$ is Firm 1's advantage over Firm 2. The upstream firms offer prices p_i, and customer selects the component with a higher net value. If $\Gamma = \gamma_1 - \gamma_2$, then the profit functions describe equilibrium profits for alternative investment configurations. A second interpretation is that the two firms are posting prices to sell a final product in a consumer market with limited preference diversity. Suppose, for example, that a type θ consumer enjoys utility $(1 + \theta) Q - P$ from consuming a quality Q product at price P, and chooses whichever product offers a higher utility. If the consumer population is distributed exponentially i.e. $\theta \sim F(\theta) \equiv 1 - e^{-\lambda\theta}$ with $\lambda \geq 1$, then the

Bertrand price equilibrium yields a corner solution, in which the higher quality firm sets its price equal to the quality difference, and the rival firm sets its price to zero. According to this interpretation the quality difference between Firm 1 and Firm 2 is $q_1 - q_2 + \Gamma$. [7]

Given the reduced form profit functions, we model duopoly product improvement as a Bayesian game with the following timing. At Stage 1, Nature chooses a realization of $\Gamma \sim \mathcal{N}(0, \sigma_\Gamma^2)$. At Stage 2, the firms privately observe signals $x_i = \Gamma + \sigma_\varepsilon \epsilon_i$, with $\sigma_\varepsilon > 0$ and $\epsilon_1, \epsilon_2 \overset{iid}{\sim} \mathcal{N}(0, 1)$, and then simultaneously choose q_i. A (pure) strategy for Firm i is a function $q_i : \mathbb{R} \to \{0, 1\}$, which defines a product improvement decision for each possible realization of x_i. At Stage 3, the firms earn profits as function of Γ and (q_1, q_2). The limiting cases of perfectly informative signals (full information) and completely noisy signals (no information) are two benchmark models. In the full information model, both firms observe Γ at the beginning of stage two prior to product improvement decisions. In the imperfect information model, neither firm learns about Γ until Stage 3 when profits are realized. Our main interest, however, is the intermediate Bayesian model, in which firms base product improvement decision on noisy and correlated signals about brand advantage.

3.2.2 Full information model

The full information limiting case generates a class of games indexed by Γ, $\{g(\Gamma)\}_{\Gamma \in \mathbb{R}}$, with the following form:

The full-information model has either a unique dominant strategy equilibrium, or three Nash equilibria, depending on r and Γ. The following proposition characterizes these equilibria when Firm 1 has the advantage ($\Gamma \geq 0$). The opposite case is symmetric, with the firms switching roles.

[7] Details are in Appendix A. While the exponential distribution simplifies the computation of equilibrium prices, it is not crucial for a winner-take-all outcome. Tirole (1988) derives a similar result for a uniform distribution.

		Firm 2	
		$q_2 = 0$	$q_2 = 1$
Firm 1	$q_1 = 0$	$max\{\Gamma, 0\}$, $max\{-\Gamma, 0\}$	$max\{\Gamma - 1, 0\}$, $max\{1 - \Gamma, 0\} - r$
	$q_1 = 1$	$max\{1 + \Gamma, 0\} - r$, $max\{-1 - \Gamma, 0\}$	$max\{\Gamma, 0\} - r$, $max\{-\Gamma, 0\} - r$

Figure 3.3: Full-information payoff matrix

Proposition 8. *In the full-information model with $\Gamma \geq 0$: (1) If $\Gamma > r$, then the unique dominant strategy equilibrium is for Firm 1 to choose $q_1 = 1$ and for Firm 2 to choose $q_2 = 0$. (2) If $\Gamma \leq r$, there are two pure strategy Nash equilibria, and a strictly mixed strategy Nash equilibrium. In a pure strategy equilibrium only one of the two firms chooses product improvement. In a mixed strategy equilibrium, Firm 1 improves its product with probability $\frac{1-\Gamma-r}{1-\Gamma}$, and Firm 2 improves its product with probability $\frac{1-r}{1-\Gamma}$.*

Proof. When $\Gamma > r$, Firm 1 will have $q = 1$ as a dominant strategy and Firm 2 will have $q = 0$ as a dominant strategy, proving (1). To show (2) note that if $0 \leq \Gamma \leq r$, Firms are playing best responses to each other only when they choose different quality levels; hence the two pure strategy Nash equilibria: (I, NI) and (NI, I). In order to characterize a mixed strategy equilibrium, let $\alpha = Prob\,(q_1 = 1)$ and $\beta = Prob\,(q_2 = 1)$. The value of β that makes Firm 1 indifferent between investing or not is $\overline{\beta} = \frac{1-r}{1-\Gamma}$. For $\beta < \overline{\beta}$, Firm 1 plays $q_1 = 1$ and for $\beta > \overline{\beta}$, Firm 1 plays $q_1 = 0$. The value of α that makes Firm 2 indifferent is $\overline{\alpha} = \frac{1-r-\Gamma}{1-\Gamma}$. For $\alpha < \overline{\alpha}$, Firm 2 plays $q_2 = 1$ and for $\alpha > \overline{\alpha}$, Firm 2 plays $q_2 = 0$. □

The multiplicity of equilibria for $0 \leq \Gamma \leq r$ is troubling on two counts. In the pure strategy equilibria, the two firms earn very different profits even if the exogenous quality difference is arbitrarily small. In the mixed strategy equilibrium, while expected profits are continuous in Γ, coordination failure results either in wasteful investment or failure to improve. One way to remedy the multiplicity problem is to introduce sufficient uncertainty about brand advantage.

		Firm 2	
		$q_2 = 0$	$q_2 = 1$
Firm 1	$q_1 = 0$	$\frac{\sigma_\Gamma}{\sqrt{2\pi}}, \frac{\sigma_\Gamma}{\sqrt{2\pi}}$	$\Phi\left(\frac{1}{\sigma_\Gamma}\right) + \sigma_\Gamma \phi\left(\frac{1}{\sigma_\Gamma}\right) - 1, \Phi\left(\frac{1}{\sigma_\Gamma}\right) + \sigma_\Gamma \phi\left(\frac{1}{\sigma_\Gamma}\right) - r$
	$q_1 = 1$	$\Phi\left(\frac{1}{\sigma_\Gamma}\right) + \sigma_\Gamma \phi\left(\frac{1}{\sigma_\Gamma}\right) - r, \Phi\left(\frac{1}{\sigma_\Gamma}\right) + \sigma_\Gamma \phi\left(\frac{1}{\sigma_\Gamma}\right) - 1$	$\frac{\sigma_\Gamma}{\sqrt{2\pi}} - r, \frac{\sigma_\Gamma}{\sqrt{2\pi}} - r$

Figure 3.4: No-information payoff matrix

3.2.3 No information

A valid criticism of the full information model is that competitors may be unsure of who has the advantage and of how large is the advantage. To address this criticism, assume it is common knowledge that $\Gamma \sim \mathcal{N}\left(0, \sigma_\Gamma^2\right)$, and that the firms receive no additional information prior to investment. The payoffs of the game are now the corresponding expected values of payoffs of the full information game. Using that[8]

$$\mathbb{E}\left[max\left\{q+\Gamma, 0\right\}\right] = \mathbb{E}\left[max\left\{q-\Gamma, 0\right\}\right] = q\Phi\left(\frac{q}{\sigma_\Gamma}\right) + \sigma_\Gamma \phi\left(\frac{q}{\sigma_\Gamma}\right),$$

where $\phi\left(.\right)$ and $\Phi\left(.\right)$ respectively denote the standard normal density function and cumulative distribution function, we have the following matrix of expected payoffs:

Our characterization of the equilibrium set of this game uses the following definition:

$$\overline{r}\left(s\right) \equiv \Phi\left(\frac{1}{s}\right) + s\phi\left(\frac{1}{s}\right) - \frac{s}{\sqrt{2\pi}}.$$

Lemma 9. *For $s > 0$, $\overline{r}\left(s\right)$ is decreasing and bounded between 1/2 and 1, and $\overline{r}\left(s\right) > 1 - \overline{r}\left(s\right)$.*

[8] See proof in the appendix.

Proof. First, for $s > 0$

$$\frac{d\overline{r}\left(s\right)}{ds} = \phi\left(\frac{1}{s}\right) - \frac{1}{\sqrt{2\pi}} < 0.$$

Moreover,

$$Lim_{s\to 0}\overline{r}\left(s\right) = 1$$

and (by L'Hopital)

$$Lim_{s\to+\infty}\overline{r}\left(s\right) = \frac{1}{2} + Lim_{s\to+\infty}\left(\frac{\phi\left(\frac{1}{s}\right) - \frac{1}{\sqrt{2\pi}}}{\frac{1}{s}}\right) = \frac{1}{2}.$$

Finally, $\overline{r}\left(s\right)$ decreasing and $Lim_{s\to+\infty}\overline{r}\left(s\right) = \frac{1}{2}$ implies that $\overline{r}\left(s\right) > \frac{1}{2}$, i.e. $\overline{r}\left(s\right) > 1 - \overline{r}\left(s\right)$.

$\square$

The next proposition characterizes the equilibria of the game without information. The equilibrium is unique when the cost of product improvement is sufficiently high or sufficiently low. However, in those cases equilibrium will reflect a coordination problem with wasteful investment (when r is low) or failure to improve (when r is high). For intermediary values of r there will be two pure strategy equilibria with only one firm investing in product improvement and also one strictly mixed strategy.

Proposition 10. *In the no-information model: (1) There exists a unique equilibrium in which neither firm improves its product if and only if* $r > \overline{r}\left(\sigma_\Gamma\right)$. *(2) There exists a unique equilibrium in which both firms improve their products if and only if* $r < 1 - \overline{r}\left(\sigma_\Gamma\right)$. *(3) There exist two asymmetric pure strategy equilibria in which only one firm improves its product if and only if* $\overline{r}\left(\sigma_\Gamma\right) \geq r \geq 1 - \overline{r}\left(\sigma_\Gamma\right)$. *(4) There exists a symmetric strictly mixed strategy equilibrium, in which each firm improves its product with probability* $\frac{\overline{r}(\sigma_\Gamma) - r}{2\overline{r}(\sigma_\Gamma) - 1}$, *if and only if* $\overline{r}\left(\sigma_\Gamma\right) > r > 1 - \overline{r}\left(\sigma_\Gamma\right)$.

Proof. First, consider dominant strategy equilibria. It is straightforward to check that $q_1 = 0$ is a best response to $q_2 = 0$ if and only if

$$\frac{\sigma_\Gamma}{\sqrt{2\pi}} \geq \Phi\left(\frac{1}{\sigma_\Gamma}\right) + \sigma_\Gamma \phi\left(\frac{1}{\sigma_\Gamma}\right) - r,$$

or equivalently $r \geq \overline{r}(\sigma_\Gamma)$, and a best response if to $q_2 = 1$ if and only if $r \geq 1 - \overline{r}(\sigma_\Gamma)$. Therefore, by symmetry $(q_1, q_2) = (0, 0)$ is a dominant strategy equilibrium if and only if $r \geq \overline{r}(\sigma_\Gamma)$. Furthermore, $q_1 = 0$ is a strictly dominant strategy if the inequality is strict. Similarly, $(q_1, q_2) = (1, 1)$, is a dominant strategy equilibrium if and only if $r \leq 1 - \overline{r}(\sigma_\Gamma)$, and a unique equilibrium if the inequality is strict. Second, $(q_1, q_2) = (1, 0)$ or $(q_1, q_2) = (0, 1)$ are asymmetric equilibria if and only if

$$\Phi\left(\frac{1}{\sigma_\Gamma}\right) + \sigma_\Gamma \phi\left(\frac{1}{\sigma_\Gamma}\right) - r \geq \frac{\sigma_\Gamma}{\sqrt{2\pi}}$$

and

$$\Phi\left(\frac{1}{\sigma_\Gamma}\right) + \sigma_\Gamma \phi\left(\frac{1}{\sigma_\Gamma}\right) - 1 \geq \frac{\sigma_\Gamma}{\sqrt{2\pi}} - r$$

both hold. These inequalities are equivalent to $\overline{r}(\sigma_\Gamma) \geq r \geq 1 - \overline{r}(\sigma_\Gamma)$. Finally, consider the possibility of a mixed strategy equilibrium. Let $\beta = Prob(q_2 = 1)$. The expected payoff of Firm 1 when playing $q_1 = 0$ is

$$(1-\beta)\,\mathbb{E}\left[max\left\{\Gamma, 0\right\}\right] + \beta\mathbb{E}\left[max\left\{\Gamma - 1, 0\right\}\right] \;=\; \frac{\sigma_\Gamma}{\sqrt{2\pi}} + \beta\overline{r}(\sigma_\Gamma) - \beta,$$

and the expected payoff of Firm 1 when playing $q_1 = 1$ is

$$(1-\beta)\,\mathbb{E}\left[max\left\{1 + \Gamma, 0\right\}\right] + \beta\mathbb{E}\left[max\left\{\Gamma, 0\right\}\right] - r \;=\; \overline{r}(\sigma_\Gamma) + \frac{\sigma_\Gamma}{\sqrt{2\pi}} - \beta\overline{r}(\sigma_\Gamma) - r.$$

Therefore, the value of β that makes Firm 1 indifferent between investing or not is

$$\beta = G(\sigma_\Gamma, r) \equiv \frac{\overline{r}(\sigma_\Gamma) - r}{2\overline{r}(\sigma_\Gamma) - 1},$$

which is strictly between 0 and 1 for investment cost satisfying $\overline{r}(\sigma_\Gamma) > r > 1 - \overline{r}(\sigma_\Gamma)$. $\square$

Proposition 2 fully characterizes all equilibria, as illustrated in the next figure.

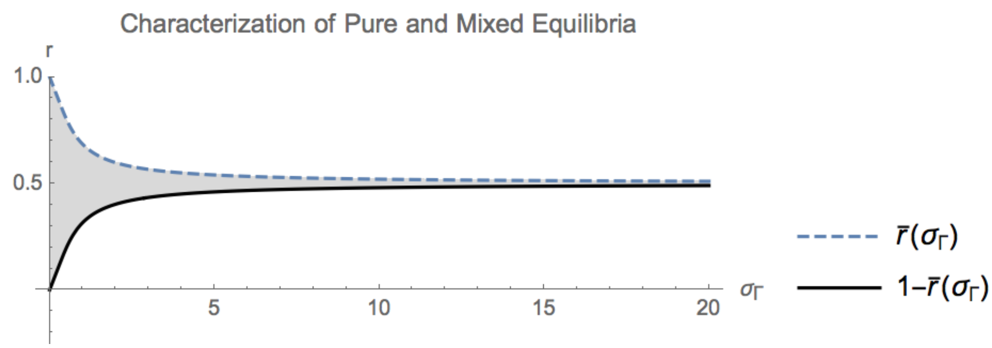


Figure 3.5: Functions $\overline{r}(\sigma_\Gamma)$, $1 - \overline{r}(\sigma_\Gamma)$ and characterization of regions of pure and mixed equilibria. The gray area, within $\overline{r}(\sigma_\Gamma)$ and $1 - \overline{r}(\sigma_\Gamma)$, consists of all combinations of r and σ_Γ for which we have two pure (asymmetric) equilibria $(q_1, q_2) = (1, 0)$ and $(q_1, q_2) = (0, 1)$ as well as a strictly mixed equilibrium, characterized in Proposition 2. The region above $\overline{r}(\sigma_\Gamma)$ has a unique pure equilibrium in which none of the firms invest in quality improvement. Finally, the region below $1 - \overline{r}(\sigma_\Gamma)$ characterizes all combinations of r and σ_Γ for which there is a unique pure equilibrium with both firms investing in quality improvement.

Notice that $r = 1/2$ is a special case for which multiple equilibria always exist, even as σ_Γ grows large. The reason is that, as σ_Γ grows large, the probability of "winning" converges to $1/2$ independently of product improvement. Consequently, for $r > 1/2$, the expected return to product improvement is negative if market uncertainty is sufficiently great, and, conversely, for $r < 1/2$, the expected return to product improvement is positive. For $r \neq 1/2$, the equilibrium is unique if market uncertainty is sufficiently great.

3.2.4 Noisy information

3.2.4.1 Information structure

We return to our main model in which the firms base their product improvement decisions on noisy information about brand advantage. In the noisy-information model, firms share a common prior belief, $\Gamma \sim N\left(0, \sigma_\Gamma^2\right)$. and observe private signals $x_i = \Gamma + \sigma_\varepsilon \epsilon_i$ with $\epsilon_1, \epsilon_2 \overset{iid}{\sim} N\left(0, 1\right)$ and $\sigma_\varepsilon > 0$. The pair of signals (x_1, x_2) are correlated through the common component Γ. The correlation coefficient is $\rho \equiv \sigma_\Gamma^2 / \left(\sigma_\varepsilon^2 + \sigma_\Gamma^2\right)$. We will parameterize the information structure by (σ_Γ, ρ), interpreting σ_Γ as a measure of market uncertainty, and ρ as a measure of strategic uncertainty. The full-information and no-information benchmarks correspond respectively to $\rho = 1$ and $\rho = 0$.

The marginal distributions of x_{-i} and Γ conditional on x_i are respectively $N\left(\rho x_i, \frac{1-\rho^2}{\rho}\sigma_\Gamma^2\right)$ and $N\left(\rho x_i, (1-\rho)\,\sigma_\Gamma^2\right)$, with correlation coefficient $\sqrt{\frac{\rho}{1+\rho}}$, and the distribution of Γ conditional on (x_1, x_2) is $N\left(\frac{\rho}{1+\rho}\left(x_1 + x_2\right), \frac{1-\rho}{1+\rho}\sigma_\Gamma^2\right)$.

3.2.4.2 Bayes-Nash equilibrium

A (pure) strategy for Firm i is a function $q_i : \mathbb{R} \to \{0, 1\}$, which defines a quality decision for each possible realization of x_i. The product improvement game proceeds as follows. Firms privately observe their signals, decide product improvement investments, and then learn the value of Γ. After that, the higher quality firm earns a profit equal to the realized total quality difference. Letting

$$\Pi_1\left(\Gamma, q_1, q_2\right) \equiv max\left\{\Gamma + q_1 - q_2, 0\right\}$$

denote the realized profit of Firm 1, and $\Pi_2(\Gamma, q_2, q_1) \equiv \Pi_1(-\Gamma, q_2, q_1)$ the realized profit of Firm 2, a Bayes Nash Equilibrium (BNE) is a strategy profile $[q_1(x_1), q_2(x_2)]$ such that

$$q_i(x_i) \in \arg\max_{q_i} \mathbb{E}\left[\Pi_i\left(\Gamma, q_i, q_{-i}(x_{-i})\right)|x_i\right]$$

for all x_i, for $i = 1, 2$. The equilibrium concept captures the interplay of market uncertainty and strategic uncertainty in that equilibrium beliefs about Γ and $q_{-i}(x_{-i})$ are jointly determined by the Firm i's signal x_i.

3.2.4.3 Threshold equilibrium

A natural strategy is for a firm to invest in product improvement if and only if its information about expected returns is sufficiently favorable. The firms have opposing interests: a high value of x_1 is good news for Firm 1, while a low value of x_2 is good news for Firm 2. Therefore, a natural candidate for equilibrium is a a strategy profile defined by a pair of critical values $(\bar{x}_1, \bar{x}_2)$ such that

$$q_1(x_1) = \begin{cases} 1 & if\ x_1 \geq \bar{x}_1 \\ 0 & if\ x_1 < \bar{x}_1 \end{cases} \text{ and } q_2(x_2) = \begin{cases} 1 & if\ x_2 \leq \bar{x}_2 \\ 0 & if\ x_2 > \bar{x}_2 \end{cases}.$$

That is, each firm follows a "threshold strategy" defined by a critical value of a signal for product improvement. The following proposition lays the foundation for consideration of such a "threshold equilibrium".

Proposition 11. *A threshold strategy is a best response to a threshold strategy.*

Proof. We show that Firm 1 optimally uses a threshold strategy if Firm 2 uses a threshold strategy; the converse is proved similarly. Suppose Firm 2 follows a threshold strategy, $q_2(x_2)$, defined by a critical value $\overline{x}_2$ such that Firm 2 invests only if $x_2 < \overline{x}_2$. Firm 1's expected payoff conditional on observing x_1 is

$$\Pi\left(q_1, x_1; \overline{x}_2\right) \equiv \mathbb{E}\left[max\left\{q_1 + \Gamma - q_2\left(x_2\right), 0\right\}|x_1\right]$$

and the expected return from investing is

$$\Delta\left(x_1; \overline{x}_2\right) \equiv \Pi\left(1, x_1; \overline{x}_2\right) - \Pi\left(0, x_1; \overline{x}_2\right).$$

5Let $F\left(x_2|\Gamma\right) = F\left(x_2|\Gamma, x_1\right)$ denote the cumulative distribution function for x_2 given Γ, and let $G\left(\Gamma|x_1\right)$ denote the distribution of Γ given x_1. These conditional normal distributions satisfy first-order stochastic dominance. Since

$$\begin{aligned}
\Pi\left(1, x_1; \overline{x}_2\right) &= \mathbb{E}\left[max\left\{1 + \Gamma - q_2\left(x_2\right), 0\right\}|x_1\right] \\
&= \int_0^{+\infty}\left[\int_{-\infty}^{\overline{x}_2}\Gamma dF\left(x_2|\Gamma\right)\right]dG\left(\Gamma|x_1\right) + \int_{-1}^{+\infty}\left[\int_{\overline{x}_2}^{+\infty}\left(1+\Gamma\right)dF\left(x_2|\Gamma\right)\right]dG\left(\Gamma|x_1\right) \\
&= \int_0^{+\infty}\Gamma F\left(\overline{x}_2|\Gamma\right)dG\left(\Gamma|x_1\right) + \int_{-1}^{+\infty}\left(1+\Gamma\right)\left[1 - F\left(\overline{x}_2|\Gamma\right)\right]dG\left(\Gamma|x_1\right)
\end{aligned}$$

and

$$\begin{aligned}
\Pi\left(0, x_1; \overline{x}_2\right) &= \mathbb{E}\left[max\left\{\Gamma - q_2\left(x_2\right), 0\right\}|x_1\right] \\
&= \int_1^{+\infty}\left[\int_{-\infty}^{\overline{x}_2}\left(\Gamma - 1\right)dF\left(x_2|\Gamma\right)\right]dG\left(\Gamma|x_1\right) + \int_0^{+\infty}\left[\int_{\overline{x}_2}^{+\infty}\Gamma dF\left(x_2|\Gamma\right)\right]dG\left(\Gamma|x_1\right) \\
&= \int_1^{+\infty}\left(\Gamma - 1\right)F\left(\overline{x}_2|\Gamma\right)dG\left(\Gamma|x_1\right) + \int_0^{+\infty}\Gamma\left[1 - F\left(\overline{x}_2|\Gamma\right)\right]dG\left(\Gamma|x_1\right),
\end{aligned}$$

It follows that

$$\Delta\left(x_1; \overline{x}_2\right) = \int \delta\left(\Gamma, \overline{x}_2\right)dG\left(\Gamma|x_1\right),$$

with

$$\delta(\Gamma, \overline{x}_2) = \begin{cases} 0 & \Gamma \leq -1 \\ (1+\Gamma)\left[1 - F(\overline{x}_2|\Gamma)\right] \quad \text{if} & -1 < \Gamma \leq 0 \\ 1 - (1-\Gamma) F(\overline{x}_2|\Gamma) & 0 < \Gamma \leq 1 \\ 1 & \Gamma \geq 1 \end{cases}$$

$\delta(\Gamma, \overline{x}_2)$ is the expected return from investing (gross of investment cost) conditional on (Γ, x_1), given Firm 2's threshold strategy. Since $\delta(\Gamma, \overline{x}_2)$ is increasing in Γ, the first-order stochastic dominance implies $\Delta(x_1; \overline{x}_2)$ is increasing in x_1. Furthermore,

$$\lim_{x_1 \to +\infty} \Delta(x_1; \overline{x}_2) = 1$$

and

$$\lim_{x_1 \to -\infty} \Delta(x_1; \overline{x}_2) = 0$$

imply that $\Delta(\overline{x}_1; \overline{x}_2) = r$ has a unique interior solution if $0 < r < 1$. This solution defines a best response threshold strategy for Firm 1, i.e. Firm 1 invests only if $x_1 > \overline{x}_1$. $\square$

Equilibrium threshold strategies are determined by symmetric indifference conditions for $\overline{x}_1$ and $\overline{x}_2$: $\Delta(\overline{x}_1; \overline{x}_2) = r$ and $\Delta(-\overline{x}_2; -\overline{x}_1) = r$. Solution defines a best response for Firm 1: $\overline{x}_1 = b_1(\overline{x}_2) \equiv b(\overline{x}_2)$. By symmetry the best response for Firm 2 is $\overline{x}_2 = b_2(\overline{x}_1) \equiv -b(-\overline{x}_1)$. It follows that $\overline{x}_1 = b(-b(-\overline{x}_1)) \equiv B(\overline{x}_1)$ at equilibrium.

The following proposition motivates consideration of conditions for a unique threshold equilibrium that is symmetric and stable. Symmetry requires that $\bar{x}_1 = -\bar{x}_2$ at equilibrium, i.e. Firm 1's best response curve crosses the negative diagonal in R^2, and stability requires that it crosses from above. Figure 2 illustrates a unique, symmetric, and stable equilibrium for the parameter configuration with a standardized degree of market uncertainty, a relatively low degree of strategic uncertainty, and several values of of the cost of product improvement.

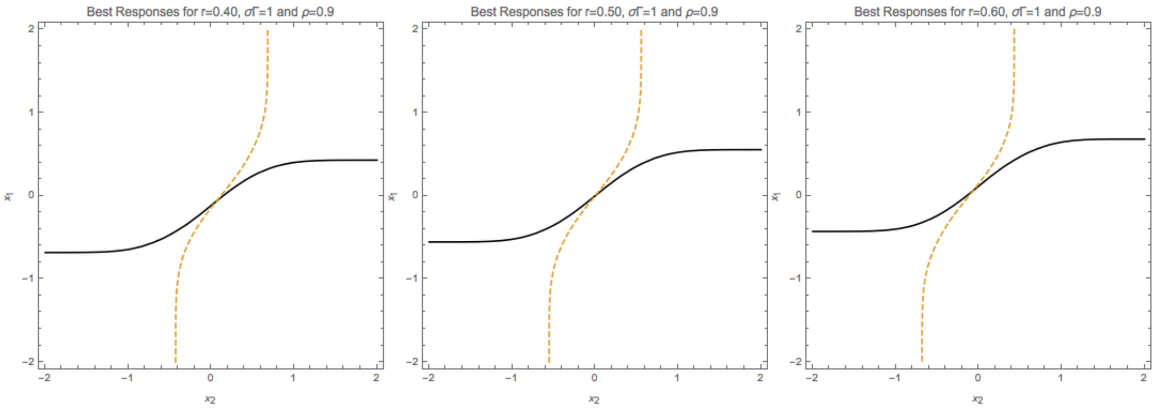


Figure 3.6: Best responses for Firm 1 (solid) and Firm 2 (dashed) for the case of $\sigma_\Gamma = 1$, $\rho = 0.9$ and $r \in \{0.4, 0.5, 0.6\}$. All cases generate a unique equilibrium.

The proposition states that a unique and stable threshold equilibrium is sufficient for a unique dominance solvable BNE.

Proposition 12. *If there is a unique stable equilibrium in threshold strategies, then no other equilibrium survives the iterated elimination of dominated strategies.*

Proof. The best response functions $b_1(\overline{x}_2)$ and $b_2(\overline{x}_1)$ are continuous, strictly increasing, and bounded above and below. Assume $(\overline{x}_1^*, \overline{x}_2^*)$ uniquely satisfy $\overline{x}_1^* = b_1(\overline{x}_2^*)$ and $\overline{x}_2^* = b_2(\overline{x}_1^*)$. By stability, $b_2(\overline{x}_1) < b_1^{-1}(x_1)$ for $b_1(+\infty) > \overline{x}_1 > \overline{x}_1^*$, and $b_2(\overline{x}_1) > b_1^{-1}(\overline{x}_1)$ for $b_1(-\infty) < \overline{x}_1 < \overline{x}_1^*$. Let $B(\overline{x}_2) \equiv b_2(b_1(\overline{x}_2))$; so $\overline{x}_2^* = B(\overline{x}_2^*)$. Define $\overline{x}_2^1 = b_2(\infty)$ and $\overline{x}_2^{k+1} = B\left(x_2^k\right)$. Given $\overline{x}_2^k > x_2^*$ and stability, we have $\overline{x}_2^{k+1} < \overline{x}_2^k$ and $\underset{x_1 \to -\infty}{lim} \overline{x}_2^k = \overline{x}_2^*$. Thus, in any equilibrium that survives the iterated elimination of iterated strategies, Firm 2 does not invest if $x_2 > \overline{x}_2^*$. A similar argument starting at $\overline{x}_2^1 = b_2(-\infty)$ implies that Firm 2 does not invest if $x_2 < \overline{x}_2^*$. Therefore, in any equilibrium surviving the iterative elimination of dominated strategies, Firm 2 uses a threshold strategy. By the Monotonicity Lemma, Firm 1 responds with a threshold strategy. Therefore, if $(\overline{x}_1^*, \overline{x}_2^*)$ is a unique threshold equilibrium and threshold best replies satisfy the stability property, then there is no other equilibrium. $\square$

Multiple threshold equilibria cannot be ruled out in general. Figure 3 illustrates multiple intersections of response curves for a lower value of ρ compared to the case in Figure 2. Now there is an unstable symmetric equilibrium, and two stable asymmetric equilibria. The comparison suggests the conjecture that a unique equilibrium exists if $\rho$is sufficiently high, i.e. strategic uncertainty is sufficiently low.

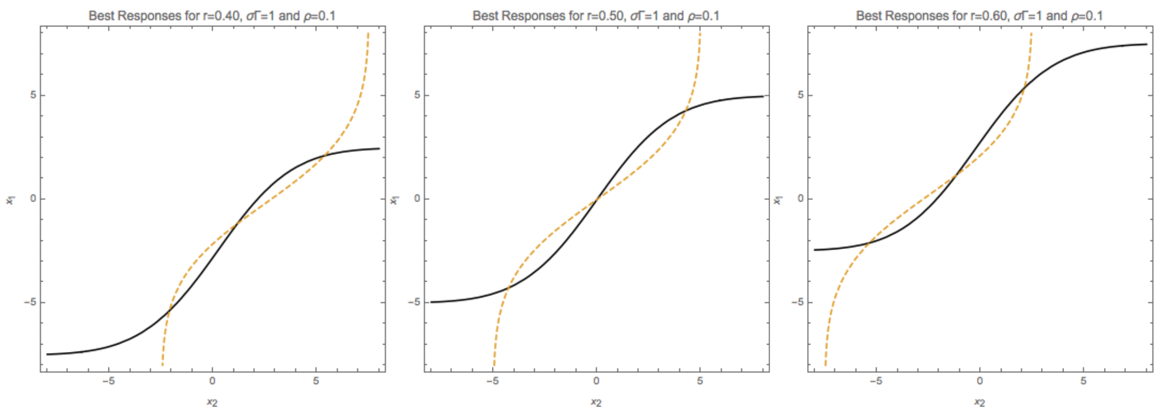


Figure 3.7: Best responses for Firm 1 (solid) and Firm 2 (dashed) for the case of $\sigma_\Gamma = 1$, $\rho = 0.1$ and different values of r. All cases have 3 threshold equilibria.

3.2.4.4 Special case: $\mathbf{r = 1/2}$

Multiple equilibria are intrinsic in the no-information model for the $r = 1/2$ special case. The multiplicity problem, however, vanishes if firms have a sufficiently small amount of noisy information, i.e. strategic uncertainty is great. In this case threshold strategy best response curves cross only once in the neighborhood of the origin. In the limit the unique equilibrium is a threshold strategy profile with $(\overline{x}_1, \overline{x}_2) = (0, 0)$.

Proposition 13. *For $r = 1/2$ and ρ sufficiently close to 1, there exists a unique and stable threshold equilibrium.*

Proof. For $x_1 \leq 0$, we have

$$\underset{\rho \to 1}{lim} \Delta(x_1; \bar{x}_2) = \begin{cases} 0 & if\ x_1 < \overline{x}_2 \\ \frac{1+x_1}{2} & if\ x_1 = \overline{x}_2 \\ 1 + x_1 & if\ x_1 > \overline{x}_2 \end{cases} .$$

and, for or $x_1 \geq 0$,

$$\underset{\rho \to 1}{lim} \Delta(x_1; \bar{x}_2) = \begin{cases} x_1 & if\ x_1 < \overline{x}_2 \\ \frac{1+x_1}{2} & if\ x_1 = \overline{x}_2 \\ 1 & if\ x_1 > \overline{x}_2 \end{cases} .$$

For $0 < \rho < 1$, $\Delta(x_1; \bar{x}_2)$ is continuously increasing in x_1, even though it is discontinuous at the limit as $\rho \to 1$. Consequently, given $r = 1/2$, if $\overline{x}_2 > 0$, then $\frac{1+\overline{x}_2}{2} > r$ and $b(\overline{x}_2) \leq \overline{x}_2$ for ρ sufficiently close to 1; $\overline{x}_2 < 0$, then similarly implies that $b(\overline{x}_2) \geq \overline{x}_2$. Therefore, $b(0) = 0$ is the unique and stable intersection of best response curves for ρ close to 1. □

This uniqueness result for $r = 1/2$ is striking, because multiple equilibria always exist in the no-information model for this case. The multiplicity problem is resolved by introducing an arbitrarily small amount of private information about market conditions, consistent with the global game literature.

Proposition 14. *For any given σ_Γ and $r = 1/2$, if equilibrium is unique for $\overline{\rho}$, then it is also unique for any $\rho > \overline{\rho}$.*

Proof. The proof is organized in 3 steps. Some properties are proved numerically whenever the analytical results are not tractable.

Step 1: If $r = 1/2$, $b_1(0) = 0$.

To prove this it suffices to show that $\Delta(0,0) = \frac{1}{2}$.

$$\begin{aligned}\Delta(0;0) &= 1-\Phi\left(\frac{1}{\sigma\sqrt{\rho}}\right)+\sigma\sqrt{\rho}\left(\phi(0)-\phi\left(-\frac{1}{\sigma\sqrt{\rho}}\right)\right)+\Phi\left(\frac{1}{\sigma\sqrt{\rho}}\right)-\frac{1}{2}-\frac{1}{\sigma\sqrt{\rho}}\int_0^1\phi\left(\frac{\Gamma}{\sigma\sqrt{\rho}}\right)\Gamma d\Gamma \\ &= \frac{1}{2}+\sigma\sqrt{\rho}\left(\phi(0)-\phi\left(-\frac{1}{\sigma\sqrt{\rho}}\right)\right)-\frac{1}{\sigma\sqrt{\rho}}\int_0^1\phi\left(\frac{\Gamma}{\sigma\sqrt{\rho}}\right)\Gamma d\Gamma\end{aligned}$$

Then, using the result

$$\int_0^1\phi\left(\frac{\Gamma}{\sigma\sqrt{\rho}}\right)\Gamma d\Gamma=\sigma^2\rho\left(\phi(0)-\phi\left(\frac{-1}{\sigma\sqrt{\rho}}\right)\right)$$

we have

$$\Delta(0;0) = \frac{1}{2}+\sigma\sqrt{\rho}\left(\phi(0)-\phi\left(-\frac{1}{\sigma\sqrt{\rho}}\right)\right)-\sigma\sqrt{\rho}\left(\phi(0)-\phi\left(\frac{-1}{\sigma\sqrt{\rho}}\right)\right)=\frac{1}{2}.$$

Step 2: For $r=1/2$, $b_1(x_2)$ is concave when $x_2>0$ and convex when $x_2<0$.

Step 3: If $b_1'(0)<1$ for $\overline{\rho}$, then $b_1'(0)<1$ for every $\rho>\overline{\rho}$. In order to prove this result we use the fact that at $(0,0)$, $b_1'(0)$ is

$$\frac{2\int_0^1(1-\Gamma)\phi\left(\frac{\Gamma}{\sigma_\Gamma\sqrt{1-\rho}}\right)\phi\left(\frac{\Gamma}{\sigma_\Gamma\sqrt{\frac{1-\rho}{\rho}}}\right)d\Gamma}{\sigma_\Gamma^2(1-\rho)\sqrt{\rho}\left(0.5-\Phi\left(\frac{-1}{\sigma_\Gamma\sqrt{1-\rho}}\right)\right)+\int_0^1(1-\Gamma)\left(\frac{\Gamma}{\sigma_\Gamma\sqrt{\frac{1-\rho}{\rho}}}\right)\phi\left(\frac{\Gamma}{\sigma_\Gamma\sqrt{1-\rho}}\right)\left[2\Phi\left(\frac{\Gamma}{\sigma_\Gamma\sqrt{\frac{1-\rho}{\rho}}}\right)-1\right]d\Gamma}$$

which, for any $\sigma_\Gamma>0$, is concave in ρ and converges to 1 as $\rho\to1$. Hence, if $b_1'(0)<1$ for $\overline{\rho}$, then $b_1'(0)<1$ for every $\rho>\overline{\rho}$. □

We study the (nonempty) set

$$\mathbf{H}^{-1}=\{(\overline{x}_1,\overline{x}_2,\rho,\sigma_\Gamma,r)\,|\,\mathbf{H}(\overline{x}_1,\overline{x}_2;\rho,\sigma_\Gamma,r)=\mathbf{0},\rho\in(0,1),\sigma_\Gamma>0,r\in(0,1)\},$$

where $\mathbf{H}(.)$ is the system of equations that characterizes the threshold equilibria, which depends on the parameters ρ, σ_Γ and r. The goal is to explore whether this set is a singleton or not, which will characterize the regions of uniqueness and multiplicity of equilibria. In principle we could try the approach suggested by Besanko et al (2010) and Doraszelski and Pakes (2007) and use the homotopy algorithm to search for all equilibria. Instead, since the goal is to characterize the uniqueness region and not search for all equilibria, we use a simpler approach relying on the fact that the slope of the best response for Firm 1 should be greater than one under multiplicity and smaller than one under uniqueness. The graphic describing the uniqueness and multiplicity regions of $\mathbf{H}^{-1}$ as a function of the parameters is a surface. In Figures 5 and 6 we explore this graphic by taking slices of it for different values of σ_Γ.

The range of ρ for which equilibrium is unique depends on the degree ex ante market uncertainty, as shown in Figure 4. Equilibrium is unique if either ρ is sufficiently small, or if σ_Γ is sufficiently large. Note that the boundary between the uniqueness region and the multiplicity region is very steep near the extreme case of no-information ($\rho = 0$) and full-information ($\rho = 1$). For intermediate cases of strategic uncertainty, a substantial amount of market uncertainty is necessary for unique equilibria.

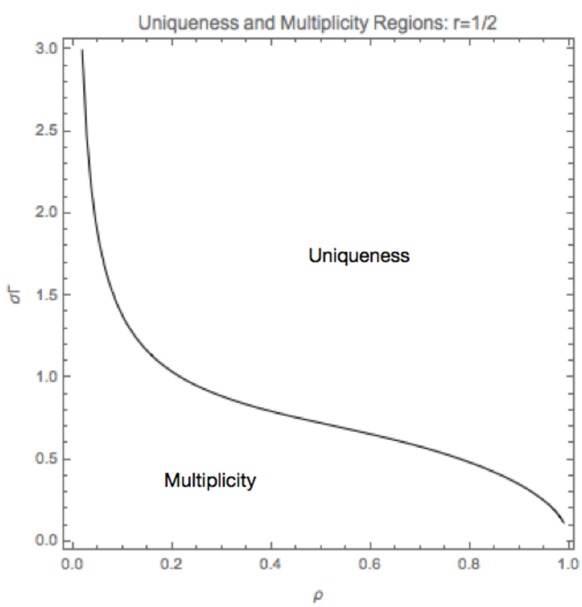


Figure 3.8: Combinations of (ρ, σ_Γ) satisfying uniqueness condition for $r = 1/2$.

3.2.4.5 Uniqueness more generally

Numerical analysis confirms that uniqueness is obtained for a larger range of (ρ, σ_Γ) beyond the $r = 1/2$ special case. Figure 7 characterizes the uniqueness region in the (ρ, r) space for various values of σ_Γ. The solid line is the critical threshold above which there exists a unique and stable threshold equilibrium. There are multiple threshold equilibria in the parameter range below the threshold. For lower (higher) values of σ_Γ, the critical threshold shifts down (up). The multiplicity region shrinks as σ_Γ grows large. Uniqueness can be reached either through enough correlation (i.e., ρ sufficiently high, but $\rho < 1$) or variance σ_Γ^2 sufficiently high. In Figures 4 and 5 we illustrate this by fixing $\sigma_\Gamma = 1$ and increasing the correlation from $\rho = 0.1$ (case of multiple equilibria) to $\rho = 0.9$ (case of unique equilibrium).

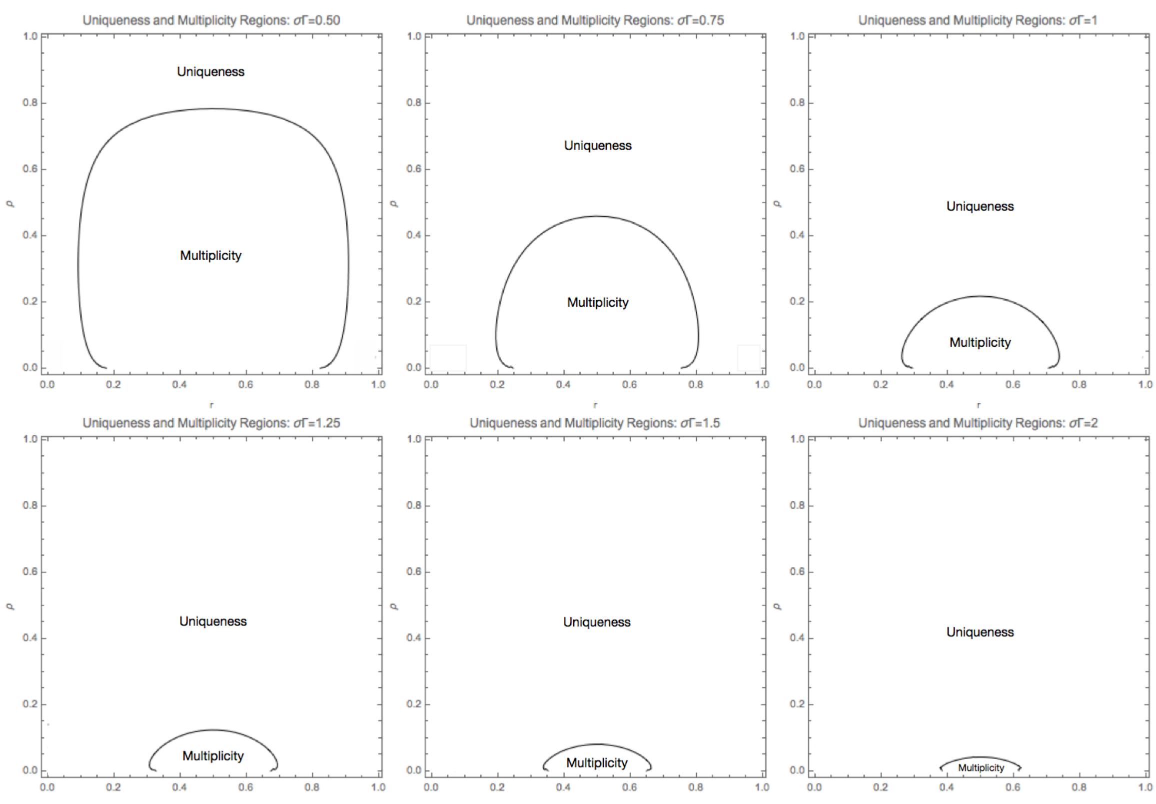


Figure 3.9: Combinations of (r, ρ) satisfying uniqueness condition for different values of σ_Γ.

3.2.5 Miscoordination

Miscoordination in the current setup happens whenever both firms take the same action. In the case with no information, the probability that both firms will take the same action will increase to 1 as σ_Γ grows, except in the non-generic case where $r = 1/2$. For the complete information case, miscoordination happens only under mixed strategies, which requires $|\Gamma| < r$. Figure 8 describes the probability of miscoordination for the complete information case when $r = 1/2$.

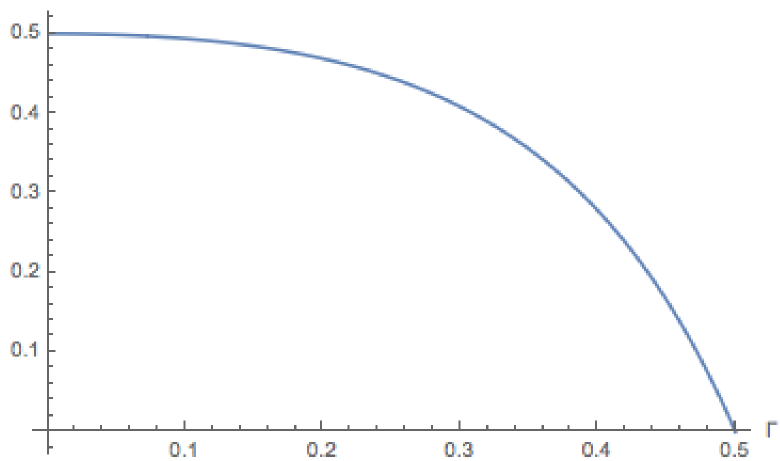


Figure 3.10: Probability of miscoordination in the complete information case as a function of Γ for $r = \frac{1}{2}$.

3.2.5.1 "Ex-ante" Probability of Miscoordination

We call "ex-ante probability" the case that considers the distribution of Γ, as opposed to a given realization of Γ, to be discussed in the next subsection. The "ex ante" probability of miscoordination is

$$Pr\left(q_1 = 1, q_1 = 1\right) + Pr\left(q_1 = 0, q_1 = 0\right) \quad = \quad Pr\left(x_1 > \overline{x}_1, x_2 < \overline{x}_2\right) + Pr\left(x_1 < \overline{x}_1, x_2 > \overline{x}_2\right)$$

$$= \int_{x_2<\bar{x}_2}\int_{x_1>\bar{x}_1} f(x_1,x_2)\,dx_1dx_2 + \int_{x_2>\bar{x}_2}\int_{x_1<\bar{x}_1} f(x_1,x_2)\,dx_1dx_2. \qquad (3.2.3)$$

Since $x_1, x_2 \sim N(0, \sigma_\Gamma^2 + \sigma_\epsilon^2)$ with $corr(x_1, x_2) = \rho$, the joint density $f(x_1, x_2)$ is given by

$$f(x_1,x_2) = \frac{1}{2\pi\left(\sigma_\Gamma^2+\sigma_\epsilon^2\right)\sqrt{1-\rho^2}} exp\left[-\frac{1}{2\left(1-\rho^2\right)}\left[\frac{x_1^2+x_2^2-2\rho x_1x_2}{\sigma_\Gamma^2+\sigma_\epsilon^2}\right]\right].$$

The next figure illustrates how the ex-ante probability of miscoordination varies for different combinations of σ_Γ^2 and ρ. One interesting take away from this figure is that ρ close to one, which is a condition that guarantees uniqueness, also reduces the probability of miscoordination to nearly zero.

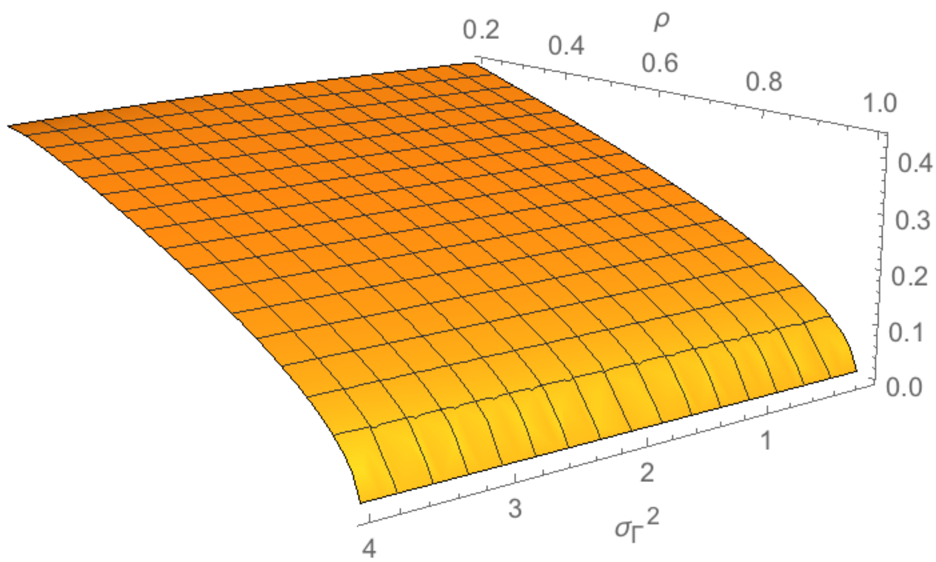


Figure 3.11: "Ex-ante" probability of miscoordination for different combinations of (σ_Γ^2, ρ) and $r = \frac{1}{2}$.

3.2.5.2 "Ex-post" Probability of Miscoordination

Now we consider the probability of miscoordination given a realization of Γ. This "ex-post" probability of miscoordination is the probability that both firms take the same action for any given realization of Γ. Since $x_1, x_2|\Gamma \overset{iid}{\sim} N(\Gamma, \sigma_\epsilon^2)$, then $f(x_1, x_2|\Gamma) = f(x_1|\Gamma)\, f(x_2|\Gamma) =$

$\frac{1}{\sigma_\epsilon^2}\phi\left(\frac{x_1-\Gamma}{\sigma_\epsilon}\right)\phi\left(\frac{x_2-\Gamma}{\sigma_\epsilon}\right)$.

$$Pr\left(q_1=1,q_1=1|\Gamma\right)+Pr\left(q_1=0,q_1=0|\Gamma\right) \quad = \quad Pr\left(x_1>\overline{x}_1,x_2<\overline{x}_2|\Gamma\right)+Pr\left(x_1<\overline{x}_1,x_2>\overline{x}_2|\Gamma\right)$$

$$= \int_{x_2<\overline{x}_2}\int_{x_1>\overline{x}_1} f\left(x_1,x_2|\Gamma\right)dx_1dx_2+\int_{x_2>\overline{x}_2}\int_{x_1<\overline{x}_1} f\left(x_1,x_2|\Gamma\right)dx_1dx_2 \tag{3.2.4}$$

The next figure illustrates how the ex-post probability of miscoordination varies for different combinations of σ_ε^2 and Γ.

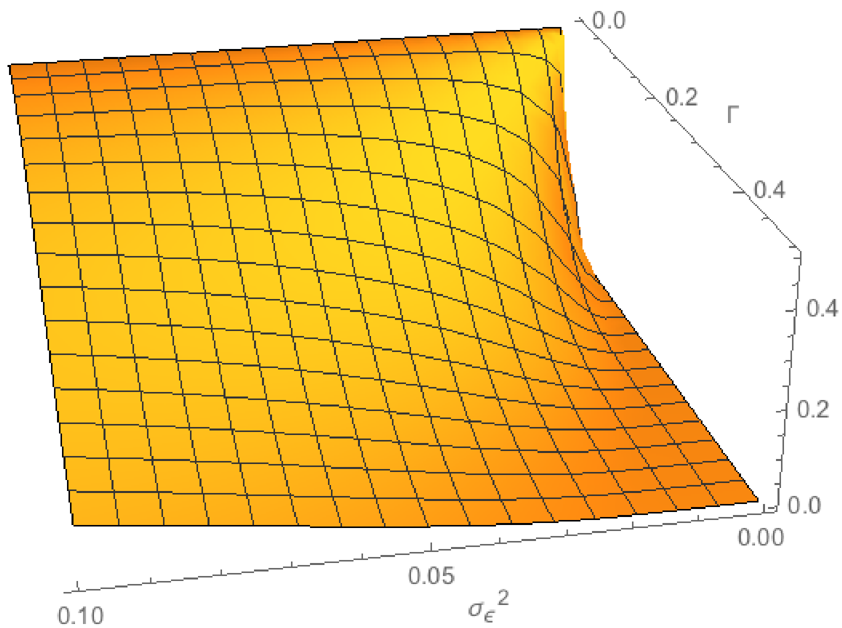


Figure 3.12: "Ex-post" probability of miscoordination for different combinations of $(\sigma_\varepsilon^2,\Gamma)$ and $r=\frac{1}{2}$.

We see from Figure 10 that, for any Γ, the ex-post probability of miscoordination vanished as $\sigma_\varepsilon \to 0$. It is useful to compare this probability with the benchmark complete information model, which we do in Figure 11. There, the orange surface represents the probability of miscoordination in the complete information game minus the ex-post probability of miscoordination for different combinations of $(\sigma_\varepsilon^2, \Gamma)$ and $r = \frac{1}{2}$. The zero plane is represented in blue. We see that the probability of miscoordination in the complete information game strictly dominates the ex-post probability of miscoordination, except when Γ is close to 0.5 and σ_ε^2 is not too close to zero.

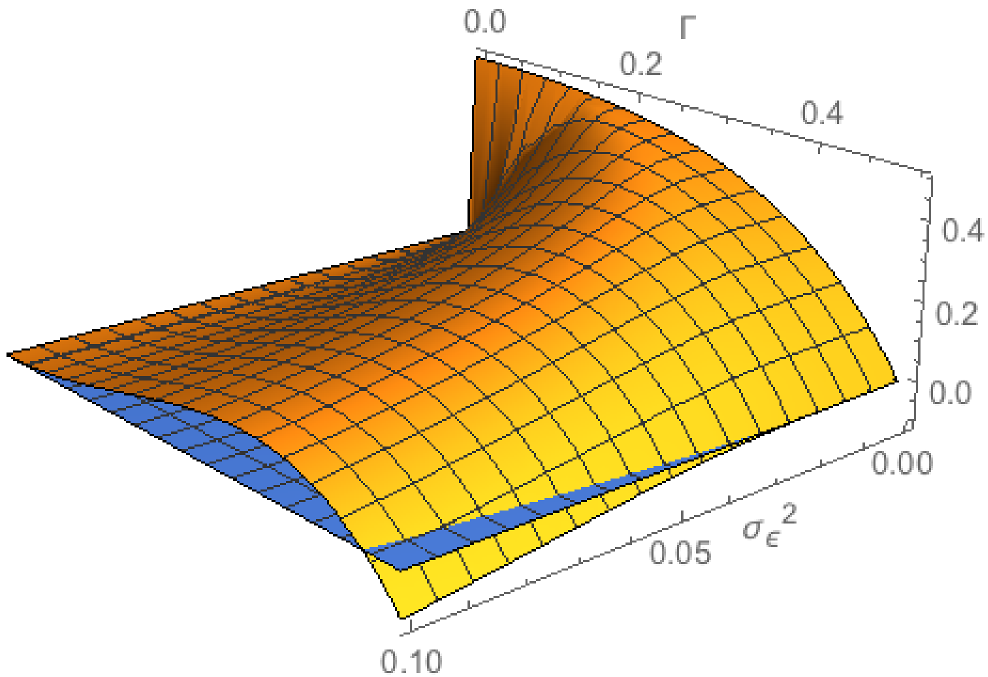


Figure 3.13: Probability of miscoordination in the complete information game minus the ex-post probability of miscoordination (orange) for $r = \frac{1}{2}$. The zero plane is represented in blue.

3.3 Conclusion

We have examined the interplay between information and actions in a duopoly game of quality upgrade. The underlying consumer market that motivates our payoff structure exhibits limited heterogeneity, resulting in a winner-take-all market. The resulting quality upgrade game is a (anti) coordination game of strategic substitution.

In the complete information version of the game, there are two pure strategy Nash equilibria and one equilibrium in mixed strategies. Uniqueness of equilibria can be reached by introducing uncertainty about the realization of the fundamental Γ. However, if firms have common knowledge of the game played and do not receive any additional information about the realization of Γ, the uniqueness region will feature miscoordination in investments with probability one. In these cases, not only the loser will have negative profits in the equilibrium in which both firms upgrade their products, but also there is a positive probability that the winning firm will also have negative profits. This result comes from the fact that in equilibrium prices are proportional to the quality differences, which in the unique equilibrium in the "no information" case is determined only by Γ since actions of both firms are the same.

With strategic uncertainty, each firm is uncertain as to what the other firm will do. In the type of competition studied here, firms have no incentive to choose the same quality as the competition arising in the marketplace would bring prices to equalize marginal cost. Actions will then depend on what each firm believes the opponent will do. When we introduce private and correlated signals about the fundamental uncertainty, each firm can form a more educated guess about what the opponent must be doing, which is the key for uniqueness of equilibria. Interestingly, this information structure alleviates substantially the problem of miscoordination observed in the no "information game" and also dominates the complete information game for a large range of parameters in the model.

Our analysis presumes that firms make their decisions simultaneously and in one single period. When these decisions are taken over time, early investments of one firm can preempt the competitor from upgrading in the future. Extending the model to a dynamic framework which can incorporate this preemption motive for investments is an important topic for future research.

Bibliography

[1] Aghion, P., and P. Bolton, "Contracts as a Barrier to Entry," *American Economic Review*, vol. 77, pp. 388-401, 1987.

[2] Anderson, S., "The Demand for Ethanol as a Gasoline Substitute," *NBER Working Paper*, No. 16371, 2010.

[3] Andrews, D. and P. Jia, "Inference for Parameters Defined by Moment Inequalities: A Recommended Moment Selection Procedure," *Econometrica*, vol. 80, No. 6, pp. 2805-2826, 2012.

[4] ANP, "Anuario Estatstico Brasileiro do Petroleo e do Gas Natural 2010," 2010.

[5] Argenziano, R. "Differentiated Networks: equilibrium and efficiency," *RAND Journal of Economics*, vol. 39, No. 3, pp. 747-769, 2008.

[6] Armstrong, M. and J. Wright, "Two-sided Markets, Competitive Botlenecks and Exclusive Contracts," *Economic Theory*, vol. 32, No. 2, pp. 353-380, 2007.

[7] Ashenfelter, O. and D. Hosken, "The Effects of Mergers on Consumer Prices: Evidence from Five Selected Case Studies," *NBER Working Paper*, No. 13859, 2008.

[8] Ashenfelter, A., D. Hosken, and M. Weinberg, "The Price Effects of a Large Merger of Manufacturers: a Case Study of Maytag-Whirlpool," *American Economic Journal: Economic Policy*, vol. 5, No. 1, pp. 239-261, 2013.

[9] Asker, J., "Diagnosing Foreclosure Due to Exclusive Dealing," Working Papers 04-36, New York University, Leonard N. Stern School of Business, Department of Economics, 2005.

[10] Bain, J. "Economics of Scale, Concentration and Entry," *American Economic Review*, vol. 44, pp. 15-39, 1954.

[11] Barron, J., B. Taylor, and J. Umbeck, "Number of Sellers, Average Prices, and Price Dispersion," *International Journal of Industrial Organization*, vol. 22, pp. 1041-1066, 2004.

[12] Barron, J. and J. Umbeck, "The Effects of Different Contractual Arrangements: the Case of Retail Gasoline Markets," *Journal of Law and Economics*, vol. 27, pp. 313-328, 1984.

[13] Berry, S., "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, vol. 60, No. 4, 1992.

[14] Berry, S., "Estimating Discrete Choice Models of Product Differentiation," *The RAND Journal of Economics*, vol. 25, pp. 242-262, 2004.

[15] Berry, S., J. Levinsohn, and A. Pakes, "Automobile Prices in Market Equilibrium," *Econometrica*, vol. 63, pp. 841-890, 1995.

[16] Berry, S. and J. Waldfogel, "Free entry and social inefficiency in radio broadcasting," *RAND Journal of Economics*, vol. 30, pp. 397–420, 1999.

[17] Besanko, D., U. Doraszelski, and Y. Kryukov, "The Economics of Predation: What Drives Pricing When There Is Learning-by-Doing?," *American Economic Review*, vol. 104, No. 3, pp. 868-897, 2014.

[18] Besanko, D., U. Doraszelski, Y. Kryukov, and M. Satterthwaite, “Learning-by-doing, organizational forgetting, and industry dynamics,” *Econometrica*, vol. 78, pp. 453-508, 2010.

[19] Bjornerstedt, J. and F. Verboven, ”Does Merger Simulation Work? A Natural Experiment in the Swedish Analgesics Market,” Katholieke Universiteit Leuven Discussion Paper Series 12, 2012.

[20] Borenstein, S. “Selling Costs and Switching Costs: Explaining Retail Gasoline Margins,” *The RAND Journal of Economics*, vol. 22, pp. 354-369, 1991.

[21] Borenstein, S. and A. Shepard, “Dynamic Pricing in Retail Gasoline Markets,” *The RAND Journal of Economics*, vol. 27, pp. 429-451, 1996.

[22] Borenstein, S., J. Bushnell, and M. Lewis, “Market Power in California’s Gasoline Market,” CSEM working paper 132, 2004.

[23] Bresnahan, T. and P. Reiss, “Do Entry Conditions Vary Across Markets?,” Booking Papers Econ. Activity, No. 3, pp 833-882, 1988.

[24] Bresnahan, T. and P. Reiss, “Entry and Competition in Concentrated Markets,” *Journal of Political Economy*, vol. 99, No. 5, pp. 977-1009, 1991.

[25] Bronnenberg, B., J.P. Dube, M. Gentzkow, and J. Shapiro, “Do pharmacists buy Bayer? Informed shoppers and the brand premium,” *Quarterly Journal of Economics*, vol. 130, No. 4, pp. 1669-1726, 2015.

[26] Cabral, L. “Dynamic Price Competition with Network Effects,” *Review of Economic Studies*, vol. 78, No. 1, pp. 83-111, 2011.

[27] Cabral, L. , D. J. Salant, and G. A. Woroch, "Monopoly Pricing with Network Externalities," *International Journal of Industrial Organization*, vol. 17, No. 2, pp. 199–214, 1999.

[28] Carlsson, H. and E. van Damme, "Global Games and Equilibrium Selection," *Econometrica*, vol. 61, pp. 989–1018, 1993.

[29] Chen, J., U. Doraszelski, and J. Harrington, "Avoiding Market Dominance: Product Compatibility in Markets with Network Effects," *RAND Journal of Economics*, vol. 40, No. 3, 455–485, 2009.

[30] Chen, Y. and M. H. Riordan, "Vertical Integration, Exclusive Dealing, and Ex Post Cartelization," *The RAND Journal of Economics*, vol. 38, No. 1, pp. 1-21, 2007.

[31] Choi, J. P. "Herd behavior, the" penguin effect," and the suppression of informational diffusion: an analysis of informational externalities and payoff interdependency," *RAND Journal of Economics*, pp. 407–425, 1997.

[32] Ciliberto, F. and E. Tamer, "Market Structure and Multiple Equilibria in Airline Markets," *Econometrica*, vol. 77, No. 6, 2009.

[33] Collard-Wexler, A., G. Gowrisankaran, and R. Lee, "An Alternating Offers Representation of the Nash-in-Nash Bargaining Solution," *Working paper*, 2015.

[34] Crawford, G. and A. Yurukoglu, "The Welfare Effects of Bundling in Multichannel Television Markets," *American Economic Review*, vol. 102, No. 2, 2012.

[35] Crawford, G., R. Lee, M. Whinston, and A. Yurukoglu, "The Welfare Effects of Vertical Integration in Multichannel Television Markets," *Working paper*, 2015.

[36] Dasgupta, A. "Coordination and Delay in Global Games," *Journal of Economic Theory*, vol. 134, No. 1, pp. 195–225, 2007.

[37] Davis, P., "Spatial competition in retail markets: movie theaters," *The RAND Journal of Economics*, vol. 37, No. 4, pp. 964–982, 2006.

[38] Doraszelski, U. and A. Pakes, "A Framework for Applied Dynamic Analysis in IO." Handbook of Industrial Organization, Edited by M. Armstrong and R. Porter, 2007.

[39] Draganska, M., D. Klapper, and S. Villas-Boas, "A Larger Slice or a Larger Pie? An Empirical Investigation of Bargaining Power in the Distribution Channel," *Marketing Science*, vol. 29, No. 1, 2010.

[40] Farrell, J. and P. Klemperer, "Coordination and Lock-In: Competition with Switching Costs and Network Effects," In M. A. Porter and R (eds.), Handbook of Industrial Organization, Elsevier, pp. 1967–2072, 2007.

[41] Farrell, J. and G. Saloner, "Installed Base and Compatibility: Innovation, Product Preannouncements, and Predation," *American Economic Review*, vol. 76, No. 5, pp. 940–955, 1986.

[42] Focarelli, D., and F. Panetta, "Are mergers beneficial to consumers?: Evidence from the market for bank deposits," *American Economic Review,* vol 93, pp. 1152–1172, 2003.

[43] Fumagalli, C. and M. Motta, "Exclusive Dealing and Entry, when Buyers Compete," *American Economic Review*, vol. 96, No. 3, pp. 785-795, 2006.

[44] Fumagalli, C., M. Motta, and L. Persson, "On the Anticompetitive Effect of Exclusive Dealing When Entry by Merger is Possible," *The Journal of Industrial Economics*, vol. XLVII, No. 4, pp. 785-811, 2009.

[45] Gabszewicz, J. and J-F. Thisse, "Price Competition, Quality and Income Disparities," *Journal of Economic Theory*, pp. 340-359, 1979.

[46] Gilbert, R., and J. Hastings, "Market Power, Vertical Integration and the Wholesale Price of Gasoline," *Journal of Industrial Economics*, vol. 53, No. 4, pp. 469-492, 2005.

[47] Gilbert, R., M. H. Riordan, "Product Improvement and Technological Tying in a Winner-Take-All Market," *The Journal of Industrial Economics*, vol. 55, No. 1, pp. 113-139, 2007.

[48] Gowrisankaran, G., A. Nevo, and R. Town, "Mergers When Prices Are Negotiated: Evidence from the Hospital Industry," *American Economic Review*, vol. 105, No. 1, pp. 172–203, 2014.

[49] Grennan, M., "Price Discrimination and Bargaining: Empirical Evidence from Medical Devices," *American Economic Review*, vol. 103, No. 1, 2013.

[50] Grout, P., "Investment and Wages in the Absence of Binding Contracts: A Nash Bargaining Approach," *Econometrica*, vol. 52, pp. 449-60, 1984.

[51] Harsanyi, J. and R. Selten, A General Theory of Equilibrium Selection in Games, MIT Press, 1988.

[52] Hastings, J., "Vertical Relationships and Competition in Retail Gasoline Markets: Empirical Evidence from Contract Changes in Southern California," *American Economic Review*, vol. 94, No.1 , pp. 317-328, 2004.

[53] Hastings, J., "Wholesale Price Discrimination and Regulation: Implications for Retail Gasoline Prices," *Working paper*, 2011.

[54] Heide, J., S. Dutta, and M. Bergen, "Exclusive Dealing and Business Efficiency: Evidence from Industry Practice," *Journal of Law and Economics*, vol. 41, pp. 387-407, 1998.

[55] Heidhues, P. and N. Melissas, "Equilibria in a dynamic global game: the role of cohort effects," *Economic Theory*, vol. 28, No. 3, pp. 531-557, 2006.

[56] Hellerstein, R., "Who Bears the Cost of a Change in the Exchange Rate? Pass-Through Accounting for the Case of Beer," *Journal of International Economics*, vol. 76, pp. 14-32, 2008.

[57] Herrendorf, B., A. Valentinyi and R. Waldmann. "Ruling out Multiplicity and Indeterminacy: The Role of Heterogeneity." *Review of Economic Studies*, vol. 67, pp. 295-308, 2000.

[58] Ho, K. and R. Lee, "Insurer Competition and Negotiated Hospital Prices," *Working paper*, 2015.

[59] Horn, H. and A. Wolinsky, "Bilateral Monopolies and Incentives for Merger," *The RAND Journal of Economics*, vol. 19, No. 3, 1988.

[60] Houde, J-F., "Gasoline Markets," *The New Palgrave Dictionary of Economics*, 2010.

[61] Houde, J-F., "Spatial Differentiation and Vertical Mergers in Retail Markets for Gasoline," *American Economic Review*, vol. 102, No. 5, 2012.

[62] Hortacsu, A., and C. Syverson, "Cementing Relationships: Vertical Integration, Foreclosure, Productivity, and Prices," *Journal of Political Economy*, vol. 115, No. 2, pp. 250-301, 2007.

[63] Johnson, J., "Anticompetitive Exclusive Dealing Without Exclusion," *Working paper*, 2014.

[64] Julien, B. and A. Pavan, "Platform Pricing Under Dispersed Information", *Working paper*, 2016.

[65] Knittel, C. and K. Metaxoglou, “Estimation of Random Coefficient Demand Models: Two Empiricists’ Perspective,” *Review of Economics and Statistics*, vol. 96, No. 1, pp. 34-59, 2014.

[66] Lafontaine, F., “Agency Theory and Franchising: Some Empirical Results,” *The RAND Journal of Economics*, vol. 23, pp. 263-283, 1992.

[67] Langer, A. and S. McRae, “Fueling Alternatives: Evidence from naturalistic Driving Data,” *Working paper*, 2013.

[68] Lee, R., “Vertical Integration and Exclusivity in Platform and Two-Sided Markets,” *American Economic Review*, vol. 103, No. 7, pp. 2960-3000, 2013.

[69] Lewis, M., “Price Dispersion and Competition with Differentiated Sellers,” *Journal of Industrial Economics*, vol. 56, pp. 654-678, 2006.

[70] Manuszak, M., “Predicting the impact of upstream mergers on downstream markets with an application to the retail gasoline industry,” *International Journal of Industrial Organization*, vol. 28, pp. 99-111, 2010.

[71] Mazzeo, M. J. ”Product choice and oligopoly market structure,” *RAND Journal of Economics*, vol. 33, No. 2, pp. 221-242, 2002.

[72] Milliou, C., and E. Petrakis, “Upstream Horizontal Mergers, Vertical Contracts, and Bargaining,” *International Journal of Industrial Organization*, 25, pp. 963-987, 2007.

[73] Morris, S. and H. Shin, “Global Games: Theory and Applications,” in Advances in Economics and Econometrics (Proceedings of the Eighth World Congress of the Econometric Society), edited by M. Dewatripont, L. Hansen and S. Turnovsky; Cambridge University Press, 2003.

[74] Morris, S. and H. Shin, "Heterogeneity and Uniqueness in Interaction Games," in The Economy as an Evolving Complex System III, edited by L. Blume and S. Durlauf; Oxford University Press, Santa Fe Institute Studies in the Sciences of Complexity, 2005.

[75] Mortimer J., "Vertical Contracts in the Video Rental Industry, " *Review of Economic Studies*, vol. 75, No. 1, pp. 165-199, 2008.

[76] Mussa, M. and S. Rosen, "Monopoly and Product Quality," *Journal of Economic Theory*, v. 18, pp. 301-317, 1978.

[77] Nash, J., "The Bargaining Problem," *Econometrica*, vol. 18, 1950.

[78] Nash, J., "Two-person Cooperative Games," *Econometrica*, vol. 21, 1953.

[79] Nevo, A., "Mergers with Differentiated Products: the Case of the Ready-to-Eat Cereal Industry," *The RAND Journal of Economics*, vol. 31, pp. 395-421, 2000.

[80] Nurski, L. and F. Verboven, "Exclusive Dealing as an Entry Barrier," *Working paper*, 2015.

[81] Pakes, A., J. Porter, K. Ho, and J. Ishii, "Moment Inequalities and Their Applications," *Econometrica*, vol. 83, No. 1, pp. 315-334, 2015.

[82] Pesendorfer, M., "Horizontal mergers in the paper industry," *The RAND Journal of Economics*, vol. 43, No. 3, pp. 495–515, 2003.

[83] Pessoa, J. P., L. Rezende, and J. Assuncao, "Flex cars and Competition in Ethanol and Gasoline Retail Markets," REAP working paper 008, 2011.

[84] Peters, C., "Evaluating the Performance of Merger Simulations: Evidence from the U.S. Airline Industry," *Journal of Law and Economics*, vol. 49, pp. 627-649, 2006.

[85] Pinske, J., M. Slade, and C. Brett., "Spatial Price Competition: A Semiparametric Approach," *Econometrica*, vol. 70, pp. 1111-53, 2002.

[86] Radner, R., A. Radunskaya, and A. Sundararajan, "Dynamic pricing of network goods with boundedly rational consumers," Proceedings of the National Academy of Sciences, 111 (1), pp. 99–104, 2014.

[87] Rubinstein, A., "Perfect Equilibrium in a Bargaining Model," *Econometrica*, vol. 50, No. 1, 1982.

[88] Rysman, M. "Adoption Delay in a Standards War," *SSRN Electronic Journal*, 2003.

[89] Salvo, A. and C. Huse, "Arbitrage Tying the Price of Ethanol to that of Gasoline? Evidence from the Uptake of Flexible-Fuel Technology," *The Energy Journal*, vol. 32, 2011.

[90] Salvo, A. and C. Huse, "Build it, but will they come? Evidence from consumer choice between gasoline and sugarcane ethanol," *Working paper*, 2013.

[91] Sass, T., "The competitive effects of exclusive dealing: Evidence from the U.S. beer industry," *International Journal of Industrial Organization*, vol. 23 , pp. 203-225, 2005.

[92] Segal, I., and M. Whinston, "Exclusive contracts and protection of investments," *The RAND Journal of Economics*, vol. 31, pp. 603–633, 2000.

[93] Shepard, A. "Contractual Form, Retail Price, and Asset Characteristics in Gasoline Retailing," *The RAND Journal of Economics*, 24: pp. 58-77, 1993.

[94] Shepard, A., "Price discrimination and retail configuration," *Journal of Political Economy*, vol. 99, pp. 30–53, 1991.

[95] Shriver, S., "Network Effects in Alternative Fuel Adoptions: Empirical Analysis of the Market for Ethanol," *Net Institute working paper*, No. 10-20, 2010.

[96] Sieg, H., "Estimating a Bargaining Model with Asymmetric Information: Evidence from Medical Malpractice Disputes," *Journal of Political Economy*, vol. 108, No. 5, 2000.

[97] Sinkinson, M., "Pricing and Entry Incentives with Exclusive Contracts: Evidence from Smartphones," *Working paper*, 2014.

[98] Slade, M., "Conjectures, Firm Characteristics, and Market Structure," *International Journal of Industrial Organization* vol. 4, pp. 347–69, 1986.

[99] Slade, M., "Interfirm Rivalry in a Repeated Game: An Empirical Test of Tacit Collusion," *Journal of Industrial Economics*, vol. 35, pp. 499-516, 1987.

[100] Slade, M., "Vancouver's Gasoline-Price Wars: An Empirical Exercise in Uncovering Supergame Strategies," *Review of Economic Studies*, vol. 59, pp. 257-76, 1992.

[101] Small, K., and H. Rosen, "Applied Welfare Economics with Discrete Choice Models," *Econometrica*, vol. 49, No. 1, pp. 105-130, 1981.

[102] Thomadsen, R., "The effect of ownership structure on prices in geographically differentiated industries," *The RAND Journal of Economics*, vol. 36, No. 4, pp. 908–929, 2005.

[103] Tirole, J., "The Theory of Industrial Organization," Cambridge, MA: MIT Press. 1988.

[104] Villas-Boas, S., "Vertical relationships between manufacturers and retailers: inference with limited data," *Review of Economic Studies*, vol. 74, No. 2, pp. 625–652, 2007.

[105] Weinberg, M., "More Evidence on the Performance of Merger Simulations," *American Economic review: Papers & Proceedings*, vol. 101, No. 3, pp. 51-55, 2011.

[106] Werden, G. and L. Froeb., "The Effects of Mergers in Differentiated Products Industries: Logit Demand and Market Policy," *Journal of Law, Economic and Organization*,vol. 10, No. 2, pp. 407-26, 1994.

[107] Weyl, G. "A Price Theory of Multi-Sided Platforms," *American Economic Review*, vol. 100, No. 4, pp. 1642–1672, 2010.

[108] Whinston, M., "Lectures on Antitrust Economics," MIT Press, 2004.

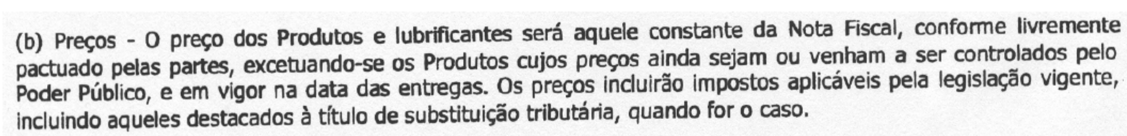
[109] Williamson, O., "The Economic Institutions of Capitalism," New York: Free Press, 1985. Katz, M. L. Vertical Contractual Relations, in R Schmalensee and R. D. Willing (eds.), Handbook of Industrial organization, New York: Elsevier Science, 1989.

Appendix A

Appendix to Chapter 1

The next figure brings a portion of the exclusivity contract of a major distributor mentioning how wholesale prices are determined. I translate the text below (my italics):

(b) Prices - the [wholesale] price of products and lubricants will be the one listed in the invoice, *freely agreed between the parties*, except for the products which prices still are or come to be regulated by the public authority, and in force at the delivery date. The [wholesale] prices will include taxes applicable according to the current legislation, including those related to tax substitution, whenever that is the case.



(b) Preços - O preço dos Produtos e lubrificantes será aquele constante da Nota Fiscal, conforme livremente pactuado pelas partes, excetuando-se os Produtos cujos preços ainda sejam ou venham a ser controlados pelo Poder Público, e em vigor na data das entregas. Os preços incluirão impostos aplicáveis pela legislação vigente, incluindo aqueles destacados à título de substituição tributária, quando for o caso.

Figure A.1: Portion of the contract of a major distributor mentioning how wholesale prices are determined.

Appendix B

Appendix to Chapter 2

In this appendix we consider the case of an unsponsored network. Suppose the price of the network is fixed at p_t in period t. We are interested in the equilibrium with monotone strategies: if a consumer with signal x_i choose to join the network at period n, then all consumers with $x_{i'} > x_i$ join at period $t' \leq t$. Any monotone strategy of the consumers can be characterized by the sequence of cutoffs $\bar{x}_1$ and $\bar{x}_2$ ($\bar{x}_1 \geq \bar{x}_2$) where the remaining consumers at period t join the network if and only if $x_i \geq \bar{x}_t$.

Second Period If a consumer does not join the network in the first period, her problem in the second period is to decide between joining the network or not, in which case she will get zero utility. Solving backwards, we know that consumer i will join the network in the second period if

$$U_{i,2} = \mathbb{E}\left[v_i + \alpha\left(n_1 + n_2\right)|x_i, \theta\right] - p_2 \geq 0.$$

For any given any monotone strategy $(\bar{x}_1, \bar{x}_2)$, The mass of consumers joining the network in each period is given by

$$n_1 = N_1(\bar{x}_1, \theta) \equiv \Phi\left(\frac{\theta - \overline{x}_1}{\sigma_h}\right),$$

and

$$n_2 = N_2(\bar{x}_1, \bar{x}_2, \theta) \equiv \Phi\left(\frac{\theta - \overline{x}_2}{\sigma_h}\right) - \Phi\left(\frac{\theta - \overline{x}_1}{\sigma_h}\right).$$

Since n_1 is publicly observed, the fundamental θ is perfectly known in the second period. Therefore

$$\begin{aligned} U_{i,2} &= \mathbb{E}\left[v_i | x_i, \theta\right] + \alpha N_1(\bar{x}_1, \theta) + \alpha N_2(\bar{x}_1, \bar{x}_2, \theta) - p_2 \\ &= \mathbb{E}\left[\theta + h_i | x_i, \theta\right] + \alpha N_1(\bar{x}_1, \theta) + \alpha N_2(\bar{x}_1, \bar{x}_2, \theta) - p_2 \\ &= \theta + \mathbb{E}\left[h_i | \eta_i\right] + \alpha N_1(\bar{x}_1, \theta) + \alpha N_2(\bar{x}_1, \bar{x}_2, \theta) - p_2 \\ &= (1-\tau)\theta + \tau x_i + \alpha N_1(\bar{x}_1, \theta) + \alpha N_2(\bar{x}_1, \bar{x}_2, \theta) - p_2, \end{aligned}$$

since $h_i | \eta_i \sim N\left(\tau\eta_i, (1-\tau)\,\sigma_h^2\right)$ and $\eta_i = x_i - \theta$.

Let's first consider the equilibrium where $\bar{x}_2 < \bar{x}_1$, that is, positive measure of consumers join the network in the second period. In this case, consumer $\bar{x}_2$ must be indifferent between buying and not buying since $U_{i,2}$ is continuous in x_i. Therefore, his indifference condition in the second period is given by

$$(1-\tau)\theta + \tau\bar{x}_2 + \alpha N_1(\bar{x}_1, \theta) + \alpha N_2(\bar{x}_1, \bar{x}_2, \theta) \quad = \quad p_2. \tag{B.0.1}$$

Let $\hat{x}_2(\bar{x}_1, \theta)$ be the solution of (B.0.1). This condition will characterize the second period threshold whenever $\hat{x}_2(\bar{x}_1, \theta) < \overline{x}_1$. If $\hat{x}_2(\bar{x}_1, \theta) \geq \overline{x}_1$, all consumers who waited in the first period are better off not buying the product, hence $\overline{x}_2 = \overline{x}_1$. This corresponds to the situation in which no consumer will join the platform in the second period for the given realization of θ and the price charged by the existing platform, p_2. Therefore, the cutoff function $\overline{x}_2\left(p_2; \bar{x}_1, \theta\right)$ is defined as

$$\overline{x}_2\left(\bar{x}_1, \theta\right) = \min\{\bar{x}_1, \hat{x}_2\left(\bar{x}_1, \theta\right)\}.$$

A unique threshold in the second period relies on enough consumer heterogeneity in tastes, in the terms of the assumption below:

Assumption 15. $\frac{\tau\sigma_h}{\alpha} > \frac{1}{\sqrt{2\pi}}$.

This assumption will be satisfied under sufficiently large heterogeneity or a mild contemporaneous network externality in the second period.

Proposition 16. *Suppose Assumption 15 holds. Then given any $\bar{x}_1$ and θ, $\bar{x}_2$ is uniquely determined for any p_2. Moreover, $\bar{x}_2$ is nondecreasing in p_2.*

Proof. Given $\bar{x}_1$, θ and $p_2 \geq c$, all terms in (B.0.1) are bounded, except for $\overline{x}_2$. Then,

$$(1-\tau)\theta + \tau\overline{x}_2 + \alpha N_1 + \alpha N_2(\overline{x}_2) - p_2$$

will be negative for sufficiently negative values of $\overline{x}_2$ and will be positive for sufficiently high values of $\overline{x}_2$. Since this function is continuous in $\overline{x}_2$, Intermediate Value Theorem implies that there exists a solution to (B.0.1). Uniqueness is obtained by showing monotonicity of the left-hand side of (B.0.1):

$$\frac{\partial(LHS)}{\partial\overline{x}_2} = \tau - \frac{\alpha}{\sigma_h}\phi\left(\gamma(\bar{x}_2)\right) \geq \tau - \frac{1}{\sqrt{2\pi}}\frac{\alpha}{\sigma_h},$$

which is strictly positive under Assumption 15, completing the proof of uniqueness. Monotonicity also follows. □

Define $\underline{\theta}(\bar{x}_1)$ be the solution of the equation

$$(1-\tau)\theta + \tau\bar{x}_1 + \alpha\Phi\left(\frac{\theta - \overline{x}_1}{\sigma_h}\right) = p_2. \tag{B.0.2}$$

Note that $\underline{\theta}(\bar{x}_1)$ is uniquely defined, and under Assumption 15 $\underline{\theta}(\bar{x}_1)$ is decreasing in $\bar{x}_1$.

Then (2.3.1) implies that

$$\overline{x}_2\left(\bar{x}_1,\theta\right)=\begin{cases}\hat{x}_2\left(\bar{x}_1,\theta\right) & \text{if } \theta>\underline{\theta}(\bar{x}_1)\\ \bar{x}_1 & \text{if } \theta\leq\underline{\theta}(\bar{x}_1).\end{cases}$$

First Period Now we consider the equilibrium behavior in the first period given the function $\bar{x}_2(\bar{x}_1,\theta)$. In the first period, consumer i receives a signal x_i, observes the first-period price p_1, and then chooses whether to join the network. Consider the monotone strategy in which consumers join the network if and only if $x_i\geq\bar{x}_1$. If consumer x_i joins the network in the first period, his expected payoff is

$$\mathbb{E}_\theta[v_i|x_i]+\mathbb{E}_\theta\left[\alpha N_1(\overline{x}_1)+\alpha N_2(\overline{x}_1,\overline{x}_2(\overline{x}_1,\theta),\theta)|x_i\right]-p_1.$$

On the other hand, if he delays the purchase, his payoff is discounted by δ but he has the option to purchase or not in the second period. In this case, his expected payoff is given by

$$\delta\cdot\mathbb{E}_\theta\left[\max\left\{0,(1-\tau)\theta+\tau x_i+\alpha N_1(\overline{x}_1)+\alpha N_2(\overline{x}_1,\overline{x}_2(\overline{x}_1,\theta),\theta)-p_2\right\}|x_i\right],$$

or

$$\delta\Pr{}_\theta(\bar{x}_2(\bar{x}_1,\theta)<x_i|x_i)\left\{(1-\tau)\theta+\tau x_i+\mathbb{E}_\theta\left[\alpha N_1+\alpha N_2-p_2|x_{i,}\bar{x}_2(\bar{x}_1,\theta)<\bar{x}_i\right]\right\},$$

since consumer i joins the network in the second period if and only if $x_i\geq\bar{x}_2(\bar{x}_1,\theta)$. The consumer whose intrinsic value equals the first-period cutoff $\bar{x}_1$ must be indifferent between joining the network and delaying the purchase. Therefore, his indifference condition is given by

$$\begin{aligned}\mathbb{E}_\theta[v_i|x_i] &+ \alpha\mathbb{E}_\theta\left[N_1+N_2|\bar{x}_1\right]-p_1\\ &= \delta\Pr{}_\theta(\bar{x}_2<\bar{x}_1|\bar{x}_1)\left\{(1-\tau)\theta+\tau\bar{x}_1+\mathbb{E}_\theta\left[\alpha N_1+\alpha N_2-p_2|\bar{x}_1,\bar{x}_2<\bar{x}_1\right]\right\}\end{aligned}$$

Since (2.3.1) holds whenever $\bar{x}_2(\bar{x}_1,\theta)<\bar{x}_1$, we have

$$\mathbb{E}_\theta[v_i|x_i]+\mathbb{E}_\theta\left[\alpha N_1+\alpha N_2|\bar{x}_1\right]-p_1=\delta\Pr{}_\theta(\bar{x}_2<\bar{x}_1|\bar{x}_1)\tau\left\{\bar{x}_1-\mathbb{E}_\theta\left[\bar{x}_2|\bar{x}_1,\bar{x}_2<\bar{x}_1\right]\right\}.$$

We have shown that $\overline{x}_2<\bar{x}_1$ if $\theta>\underline{\theta}$, and $\overline{x}_2=\bar{x}_1$ otherwise, with $\underline{\theta}$ defined in (2.3.2). The indifference condition can then be rewritten as

$$p_1=(1-\kappa)\mu+\kappa\overline{x}_1+\alpha\bar{N}_1\left(\overline{x}_1\right)+\alpha\bar{N}_2\left(\overline{x}_1\right)-\delta\tau(1-\bar{Q}(\bar{x}_1))\left[\bar{x}_1-\bar{X}_{2c}(\bar{x}_1)\right] \qquad \text{(B.0.3)}$$

where

$$\begin{aligned}\bar{N}_1\left(\overline{x}_1\right) &= \mathbb{E}_\theta\left[\Phi\left(\gamma(\bar{x}_1)\right)|\overline{x}_1\right]=\Phi\left(\left(\mu-\overline{x}_1\right)z_1\right),\\ \bar{N}_2\left(\overline{x}_1\right) &= \mathbb{E}_\theta\left[\Phi\left(\frac{\theta-\bar{x}_2(\bar{x}_1,\theta)}{\sigma_h}\right)|\overline{x}_1\right]-\bar{N}_1\left(\overline{x}_1\right)=\int_{-\infty}^{\infty}\Phi\left(\frac{\theta-\bar{x}_2(\bar{x}_1,\theta)}{\sigma_h}\right)f\left(\theta|\overline{x}_1\right)d\theta-\bar{N}_1\left(\overline{x}_1\right),\\ \bar{Q}\left(\overline{x}_1\right) &= \Pr\left[\bar{x}_2(\bar{x}_1,\theta)=\bar{x}_1|\overline{x}_1\right]=\Pr\left[\theta\le\underline{\theta}|\overline{x}_1\right],\\ \bar{X}_{2c}(\bar{x}_1) &= \mathbb{E}_\theta\left[\bar{x}_2(\bar{x}_1,\theta)|\overline{x}_1,\bar{x}_2<\bar{x}_1\right]=\frac{1}{(1-\bar{Q}(\bar{x}_1))}\int_{\underline{\theta}}^{\infty}\bar{x}_2(\bar{x}_1,\theta)f\left(\theta|\overline{x}_1\right)d\theta\end{aligned}$$

where $z_1=\sqrt{\frac{(1-\rho)}{(1+\rho)\left(\sigma_\theta^2+\sigma_h^2\right)}}$.

Appendix C

Appendix to Chapter 3

C.1 Consumer market

The following duopoly model illustrates how limited consumer heterogeneity results in a winner-take-all market. A type θ consumer enjoys net utility $(1+\theta)Q-P$ from consuming a good of quality Q at price P, and the population of consumers is distributed exponentially, i.e. $\theta \sim F(\theta) = 1 - e^{-\lambda\theta}$.

There are two firms, $i \in \{1, 2\}$, with given qualities $Q_1 > Q_2$, who set prices P_i in Bertrand-Nash equilibrium. At an interior solution, such that the firms share the market, the marginal consumer is

$$\widehat{\theta} = \frac{P_1 - P_2}{Q_1 - Q_2} - 1.$$

Firm 1 maximizes $P_1 e^{-\lambda\hat{\theta}}$ and has a dominant strategy

$$P_1 = \frac{Q_1 - Q_2}{\lambda}.$$

Firm 2 maximizes $P_2[1 - e^{-\lambda\hat{\theta}}]$ and sets a price satisfying the first-order condition

$$\left[1 - e^{-\lambda\hat{\theta}}\right] - \frac{\lambda P_2}{Q_1 - Q_2} e^{-\lambda\hat{\theta}} = 0$$

and the second-order condition

$$\left[\frac{-2\lambda}{Q_1 - Q_2} - \frac{\lambda^2 P_2}{(Q_1 - Q_2)^2}\right] e^{-\lambda\hat{\theta}} < 0.$$

At $P_2 = 0$ and $P_1 = \frac{Q_1 - Q_2}{\lambda}$, the left-hand side of Firm 2's first-order condition is equal to $1 - e^{-(1-\lambda)}$, which is positive if $\lambda < 1$. Therefore $P_2 > 0$ is a best response if $\lambda < 1$, and the two firms share the market. If $\lambda \geq 1$, then an interior solution fails to exist. The equilibrium is a corner solution in which Firm 1 sets $P_1 = Q_1 - Q_2$, Firm 2 sets $P_2 = 0$, and Firm 1 wins the entire market.

In case of an interior solution, the first-order conditions combine and imply that the condition defining the marginal consumer becomes

$$e^{\lambda\hat{\theta}} + \lambda(1 + \hat{\theta}) = 2.$$

Remarkably, the marginal consumer, and therefore market shares, depend only on λ. Consequently, the profits of both firms are proportional to the quality difference:

$$\pi_1 = \frac{Q_1 - Q_2}{\lambda} e^{-\lambda\hat{\theta}} \equiv \frac{Q_1 - Q_2}{\lambda} s_1;$$

$$\pi_2 = \frac{Q_1 - Q_2}{\lambda}[e^{\lambda\hat{\theta}} - 1][1 - e^{-\lambda\hat{\theta}}] \equiv \frac{Q_1 - Q_2}{\lambda} \frac{(1 - s_1)^2}{s_1}.$$

Clearly, there is an incentive for maximum quality differentiation. The issue is how do firms coordinate on quality. It is interesting that Firm 2 earns greater profit when s_1 is small, which occurs in equilibrium when λ is small.

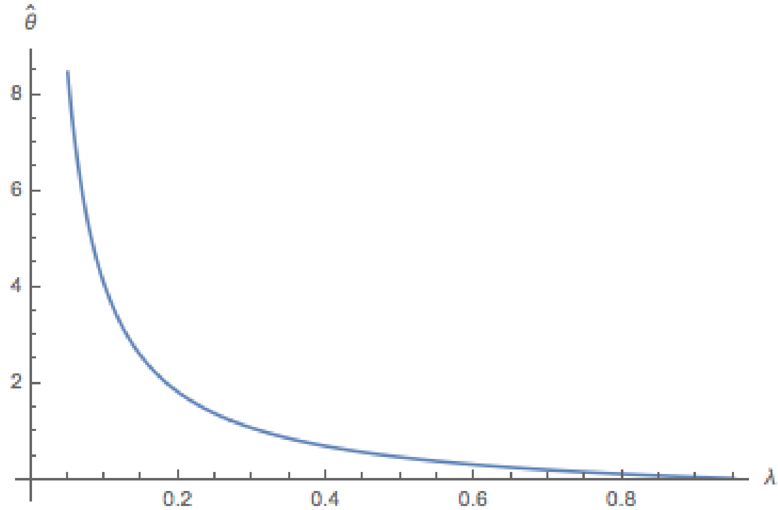


Figure C.1: Values of $\hat{\theta}$ as a function of λ for $0 < \lambda < 1$.

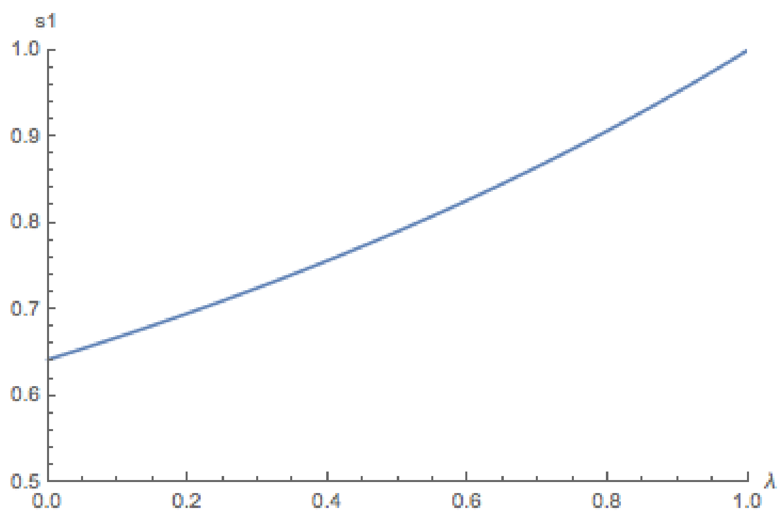


Figure C.2: s_1 for different values of λ, with $0 < \lambda < 1$.

C.2 Simple Asymmetric Quality Upgrade Game

Here we consider a variation of the simple quality upgrade model discussed in Section 2, relaxing the condition that the returns to quality upgrade are the same for the two firms. We do this by considering that each firm privately observes the realization of its own Γ_i, with Γ_1 and Γ_2 normally distributed with mean μ, variance σ_Γ^2 and correlation coefficient

ρ. There is one important aspect that distinguishes this game from the previous example: in the asymmetric case considered here there is no (own) market uncertainty.

		Firm 2	
		$q_2 = 0$	$q_2 = 1$
Firm 1	$q_1 = 0$	1, 1	0, Γ_2
	$q_1 = 1$	Γ_1, 0	$\Gamma_1 - c$, $\Gamma_2 - c$

Figure C.3: Quality upgrade game with asymmetric returns.

As before, we will look for equilibria in cutoff strategies. The expected payoff to Firm 1 from not investing is $Pr\left(\Gamma_2 \leq \overline{\Gamma}_2|\Gamma_1\right)$. When investing, the expected payoff to Firm 1 is

$$\Gamma_1 - c\left[1 - Pr\left(\Gamma_2 \leq \overline{\Gamma}_2|\Gamma_1\right)\right]$$

where

$$Pr\left(\Gamma_2 \leq \overline{\Gamma}_2|\Gamma_1\right) = \Phi\left(\frac{\overline{\Gamma}_2 - \rho\Gamma_1 - (1-\rho)\,\mu}{\sigma_\Gamma\sqrt{\frac{1-\rho^2}{\rho}}}\right).$$

Firm 1 is indifferent between investing or not if

$$\Gamma_1 - c - (1-c)\,\Phi\left(\frac{\overline{\Gamma}_2 - \rho\Gamma_1 - (1-\rho)\,\mu}{\sigma_\Gamma\sqrt{\frac{1-\rho^2}{\rho}}}\right) = 0.$$

The slope of the best response of Firm 1 is always positive (strategic complementarity):

$$b_1'\left(\Gamma_2\right) = \frac{\frac{(1-c)}{\sigma_\Gamma\sqrt{\frac{1-\rho^2}{\rho}}}\phi\left(\frac{\Gamma_2-\rho\Gamma_1-(1-\rho)\mu}{\sigma_\Gamma\sqrt{\frac{1-\rho^2}{\rho}}}\right)}{1+\frac{(1-c)\rho}{\sigma_\Gamma\sqrt{\frac{1-\rho^2}{\rho}}}\phi\left(\frac{\Gamma_2-\rho\Gamma_1-(1-\rho)\mu}{\sigma_\Gamma\sqrt{\frac{1-\rho^2}{\rho}}}\right)} > 0.$$

The condition for uniqueness in this case requires that the slope of the best response $b_1(\Gamma_2)$ is smaller than 1, i.e, the best response $b_1(x_2)$ crosses the 45^o line only once. This is true if

$$\frac{\sigma_\Gamma}{(1-c)} > \sqrt{\frac{\rho}{2\pi}}\sqrt{\frac{1-\rho}{1+\rho}}. \tag{C.2.1}$$

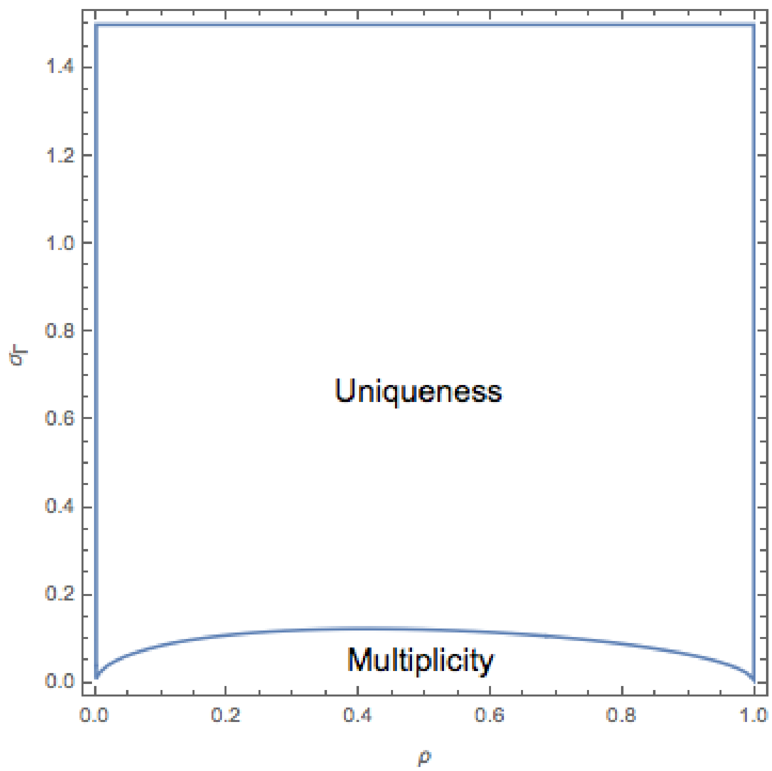


Figure C.4: Combinations of (ρ, σ_Γ) satisfying uniqueness condition for different values of σ_Γ for the asymmetric model. Shaded area represents the uniqueness region for the case of $c = 0.25$.

C.3 Simple Entry Game

Consider the following entry game:

		Firm 2	
		$q_2 = 0$	$q_2 = 1$
Firm 1	$q_1 = 0$	0, 0	0, Γ
	$q_1 = 1$	Γ, 0	$\Gamma - c$, $\Gamma - c$

Figure C.5: Payoff matrix - simple entry game

The individual payoff is normalized to zero when the firm decides not to enter, Γ when the firm enters alone and $\Gamma - c$ when both firms enter. We interpret Γ as the monopoly profit and $c > 0$ as the discount in the profit that comes from competition when both firms upgrade their products. In the complete information game, for $\Gamma > c$, entry is a dominant strategy for both firms. If $\Gamma < 0$, then no entry is dominant. For values of Γ between 0 and c, there will be two pure strategy Nash equilibria and one equilibrium in mixed strategies.

Now assume that the fundamental parameter Γ is normally distributed with mean μ and variance σ_Γ^2. Firms do not observe Γ directly. Instead, each firm receives a signal $x_i = \Gamma + \varepsilon_i$, where $\varepsilon_i \sim N\left(0, \sigma_\varepsilon^2\right)$ with ε_1 and ε_2 independent of each other and of Γ. The signals are jointly normally distributed with correlation coefficient $\rho = \frac{\sigma_\Gamma^2}{\sigma_\Gamma^2 + \sigma_\varepsilon^2}$. This game features both fundamental uncertainty (Γ is not known) and strategic uncertainty (a firm does not know the opponent's action). Similarly to the game of complete information, if the signals are precise about the realization of Γ (i.e., σ_ε^2 is low), high signals will induce firms to invest as a dominant strategy and low signals will lead to no investment. We consider equilibria with cutoff strategies, in which there is a threshold $\overline{x}$ below which a firm will not invest and above which the firm will invest.

The expected payoff to Firm 1 from not investing is zero. When investing, the expected payoff to Firm 1 is

$$\mathbb{E}\left[\Gamma | x_1\right] - c\left[1 - Pr\left(x_2 \leq \overline{x}_2 | x_1\right)\right]$$

where

$$\mathbb{E}\left[\Gamma | x_1\right] = \rho x_1 + (1-\rho)\,\mu$$

and

$$Pr\left(x_2 \leq \overline{x}_2 | x_1\right) = \Phi\left(\frac{\overline{x}_2 - \rho x_1 - (1-\rho)\,\mu}{\sigma_\Gamma \sqrt{\frac{1-\rho^2}{\rho}}}\right).$$

Firm 1 is indifferent between investing or not if

$$\rho \overline{x}_1 + (1-\rho)\,\mu - c + c\Phi\left(\frac{\overline{x}_2 - \rho \overline{x}_1 - (1-\rho)\,\mu}{\sigma_\Gamma \sqrt{\frac{1-\rho^2}{\rho}}}\right) = 0.$$

The slope of the best response of Firm 1 is

$$b_1'\left(x_2\right) = -\frac{\frac{c}{\sigma_\Gamma \sqrt{\frac{1-\rho^2}{\rho}}}\phi\left(\frac{\overline{x}_2 - \rho \overline{x}_1 - (1-\rho)\mu}{\sigma_\Gamma \sqrt{\frac{1-\rho^2}{\rho}}}\right)}{\rho - \frac{c\rho}{\sigma_\Gamma \sqrt{\frac{1-\rho^2}{\rho}}}\phi\left(\frac{\overline{x}_2 - \rho \overline{x}_1 - (1-\rho)\mu}{\sigma_\Gamma \sqrt{\frac{1-\rho^2}{\rho}}}\right)}.$$

Since the numerator is always positive, the best response will be decreasing in x_2 (i.e., strategic substitutes) if

$$\sigma_\Gamma > c\sqrt{\frac{\rho}{2\pi\left(1-\rho^2\right)}}. \tag{C.3.1}$$

If $\overline{x}_2 \to \infty$, this indifference condition becomes $\overline{x}_1 = -\frac{(1-\rho)\mu}{\rho}$ and when $\overline{x}_2 \to -\infty$, $\overline{x}_1 = \frac{c-(1-\rho)\mu}{\rho}$. Focusing on the case of strategic substitutes, the slope of the best response $b_1'\left(x_2\right)$ will be larger than -1 if

$$\sigma_\Gamma > \frac{c}{\sqrt{2\pi\rho}}\sqrt{\frac{1-\rho}{1+\rho}}. \tag{C.3.2}$$

From (C.3.1) we have that in order to have strategic substitutes, ρ cannot be too close to one. If that happens for some range of values of the signals, the slope of the best response will be positive and $b_1\left(.\right)$ will be a correspondence (and not a function). These cases are always associated to multiple equilibria. Provided that the best response $b_1(.)$ has a negative slope, one condition that ensures uniqueness is that it crosses the -45^o line only once. This is guaranteed if $b_1'(x_2)$ is larger than -1, which is true whenever condition (C.3.2) is satisfied. Turns out that this condition will be satisfied more easily the closer ρ is to 1. These two conditions together

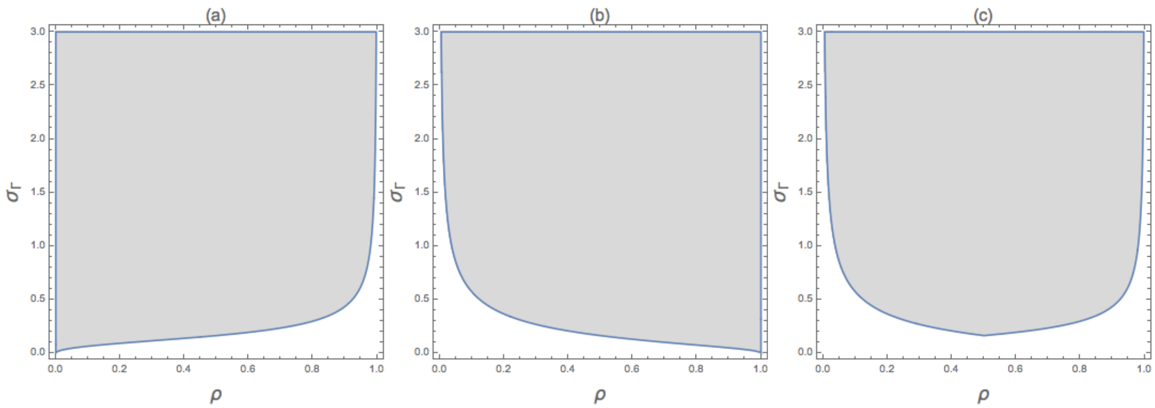


Figure C.6: Combinations of (ρ, σ_Γ) satisfying uniqueness condition for different values of σ_Γ. Shaded area represents the uniqueness region.

C.4 Numerical Analysis

In this appendix we study the slope and concavity of $b_1(x_2)$. Let

$$\begin{aligned}\Delta(x_1;\overline{x}_2) &= \int \delta(\Gamma,\overline{x}_2)\, dG(\Gamma|x_1) \\ &= F_L(x_1;\sigma_\varepsilon,\rho) + F_R(x_1,\overline{x}_2;\sigma_\varepsilon,\rho)\end{aligned}$$

where

$$F_L(x_1;\sigma_\varepsilon,\rho) = \Phi\left(\frac{\rho x_1 - 1}{\sigma_\varepsilon\sqrt{\rho}}\right) + \sigma_\varepsilon\sqrt{\rho}\left(\phi\left(\frac{\sqrt{\rho}x_1}{\sigma_\varepsilon}\right) - \phi\left(\frac{\rho x_1 - 1}{\sigma_\varepsilon\sqrt{\rho}}\right)\right) + \rho x_1\left(\Phi\left(\frac{\rho x_1}{\sigma_\varepsilon\sqrt{\rho}}\right) - \Phi\left(\frac{\rho x_1 - 1}{\sigma_\varepsilon\sqrt{\rho}}\right)\right)$$

and

$$F_R\left(x_1, \overline{x}_2; \sigma_\varepsilon, \rho\right) = \frac{1}{\sigma_\varepsilon\sqrt{\rho}}\left[\int_{-1}^{0}(1+\Gamma)\,\phi\left(\frac{\Gamma-\rho\bar{x}_1}{\sigma_\varepsilon\sqrt{\rho}}\right)\Phi\left(\frac{\Gamma-\bar{x}_2}{\sigma_\varepsilon}\right)d\Gamma + \int_{0}^{1}(1-\Gamma)\,\phi\left(\frac{\Gamma-\sigma_\varepsilon\bar{x}_1}{\sigma_\varepsilon\sqrt{\rho}}\right)\Phi\left(\frac{\Gamma-\bar{x}_2}{\sigma_\varepsilon}\right)d\Gamma\right].$$

Define[1] $F\left(x_1, \overline{x}_2; \sigma_\varepsilon, \rho, r\right) = F_L\left(x_1, ; \sigma_\varepsilon, \rho\right) + F_R\left(x_1, \overline{x}_2; \sigma_\varepsilon, \rho\right) - r$. The best response function $b_1\left(\overline{x}_2\right)$ is defined implicitly as

$$F\left(b_1\left(\overline{x}_2\right), \overline{x}_2; \sigma_\varepsilon, \rho, r\right) = 0.$$

The derivative $b_1^{'}\left(\overline{x}_2\right)$ is then obtained from the Implicit Function Theorem by

$$b_1^{'}\left(\overline{x}_2\right) = -\frac{\frac{\partial F_R}{\partial \overline{x}_2}}{\frac{\partial F_L}{\partial x_1} + \frac{\partial F_R}{\partial x_1}}, \tag{C.4.1}$$

where

$$\frac{\partial F_L\left(x_1; \sigma_\varepsilon, \rho\right)}{\partial} = \Phi\left(\frac{\rho x_1 - 1}{\sigma_\varepsilon\sqrt{\rho}}\right) + \sigma_\varepsilon\sqrt{\rho}\left(\phi\left(\frac{\sqrt{\rho}x_1}{\sigma_\varepsilon}\right) - \phi\left(\frac{\rho x_1 - 1}{\sigma_\varepsilon\sqrt{\rho}}\right)\right) + \rho\bar{x}_1\left(\Phi\left(\frac{\rho x_1}{\sigma_\varepsilon\sqrt{\rho}}\right) - \Phi\left(\frac{\rho x_1 - 1}{\sigma_\varepsilon\sqrt{\rho}}\right)\right)$$

$$\frac{\partial F_R\left(x_1, \overline{x}_2; \sigma_\varepsilon, \rho\right)}{\partial \overline{x}_2} = -\frac{1}{\sigma_\varepsilon^2\sqrt{\rho}}\left[\int_{-1}^{0}(1+\Gamma)\,\phi\left(\frac{\Gamma-\rho\bar{x}_1}{\sigma_\varepsilon\sqrt{\rho}}\right)\phi\left(\frac{\Gamma-\bar{x}_2}{\sigma_\varepsilon}\right)d\Gamma + \int_{0}^{1}(1-\Gamma)\,\phi\left(\frac{\Gamma-\rho\bar{x}_1}{\sigma_\varepsilon\sqrt{\rho}}\right)\phi\left(\frac{\Gamma-\bar{x}_2}{\sigma_\varepsilon}\right)d\Gamma\right],$$

$$\frac{\partial F_L\left(\overline{x}_1; \sigma, \rho\right)}{\partial \overline{x}_1} = \rho\left(\Phi\left(\frac{\sqrt{\rho}\bar{x}_1}{\sigma_\varepsilon}\right) - \Phi\left(\frac{\rho\bar{x}_1 - 1}{\sigma_\varepsilon\sqrt{\rho}}\right)\right),$$

and $\frac{\partial F_R(x_1, \overline{x}_2; \sigma_\varepsilon, \rho)}{\partial \overline{x}_1}$ is given by

$$\frac{1}{\sigma_\varepsilon\sqrt{\rho}}\left[\int_{-1}^{0}(1+\Gamma)\left(\frac{\Gamma-\rho\bar{x}_1}{\sigma_\varepsilon^2}\right)\phi\left(\frac{\Gamma-\rho\bar{x}_1}{\sigma_\varepsilon\sqrt{\rho}}\right)\Phi\left(\frac{\Gamma-\bar{x}_2}{\sigma_\varepsilon}\right)d\Gamma + \int_{0}^{1}(1-\Gamma)\left(\frac{\Gamma-\rho\bar{x}_1}{\sigma_\varepsilon^2}\right)\phi\left(\frac{\Gamma-\rho\bar{x}_1}{\sigma_\varepsilon\sqrt{\rho}}\right)\Phi\left(\frac{\Gamma-\bar{x}_2}{\sigma_\varepsilon}\right)d\Gamma\right].$$

[1] Here we are developing this section in the space $(\sigma_\varepsilon, \rho)$ instead of (σ_Γ, ρ) because it simplifies the expressions.

Using IFT once more we can obtain $b_1^{''}(\overline{x}_2)$:

$$\frac{\partial^2\Delta}{\partial x_1^2}\frac{\partial x_1}{\partial x_2}+\frac{\partial\Delta}{\partial x_1}\frac{\partial^2 x_1}{\partial x_2^2}+\frac{\partial^2\Delta}{\partial x_2^2}=0$$

which implies that

$$b_1^{''}(x_2)=\frac{\partial^2 x_1}{\partial x_2^2} \quad = \quad \frac{\frac{\partial^2\Delta}{\partial x_1^2}\frac{\partial\Delta}{\partial x_2}-\frac{\partial^2\Delta}{\partial x_2^2}\frac{\partial\Delta}{\partial x_1}}{\left(\frac{\partial\Delta}{\partial x_1}\right)^2}. \tag{C.4.2}$$

Results used in the proof of Proposition 6

The first numerical result showed here is that the best response $b_1(x_2)$ is strictly increasing. We show this by evaluating $b_1^{'}(x_2)$ (derived in (C.4.1)) in 100,000 different combinations of (σ_Γ,ρ,x_2). The Mathematica code is presented in pictures 17 and 18 alongside with the results. We test for violations of $b_1^{'}(x_2)>0$ and none was found in the 100,000 evaluations tested.

The second numerical result showed here is that for $r=1/2$, the best response $b_1(x_2)$ is strictly concave when $x_2>0$ and strictly convex when $x_2<0$. Similarly, we show this by evaluating $b_1^{''}(x_2)$ (derived in (C.4.2)) in 100,000 different combinations of (σ_Γ,ρ,x_2). The Mathematica code is presented in pictures 19 and 19 alongside with the results. We test for violations of $b_1^{''}(x_2)<0$ when $x_2>0$ and none was found in the 100,000 evaluations tested.

```
Clear["Global`*"]
n = 100 000;
SeedRandom[1234]
{capRho, capSigmaGamma, capR, capX2} = {1, 3, .5, 4};

randomRho = RandomReal[{0, capRho}, {n, 1}];
randomSigmaGamma = RandomReal[{0, capSigmaGamma}, {n, 1}];
randomR = RandomReal[{0.5, capR}, {n, 1}];
randomX2 = RandomReal[{0, capX2}, {n, 1}];

parms = Join[randomX2, randomRho, randomSigmaGamma, randomR, 2];

Φ[x_] := CDF[NormalDistribution[0, 1], x];
(* function for standard normal CDF, Φ *)
ϕ[x_] := PDF[NormalDistribution[0, 1], x];
(* function for standard normal PDF, ϕ *)

(* starting values used in the computation of best respose b1(x2) *)
{x1start, x2start} = {0.01, -0.01};

(* Equilibrium *)
FL[x1_, ρ_, σΓ_] := Φ[(ρ * x1 - 1)/(σΓ * Sqrt[1 - ρ])] + σΓ * Sqrt[1 - ρ] * (ϕ[(ρ * x1)/(σΓ * Sqrt[1 - ρ])] - ϕ[(ρ * x1 - 1)/(σΓ * Sqrt[1 - ρ])]) +
    ρ * x1 * (Φ[(ρ * x1)/(σΓ * Sqrt[1 - ρ])] - Φ[(ρ * x1 - 1)/(σΓ * Sqrt[1 - ρ])]);
FR[x1_, x2_, ρ_, σΓ_] := NIntegrate[1/(σΓ * Sqrt[1 - ρ]) * (1 + Γ) *
      ϕ[(Γ - ρ * x1)/(σΓ * Sqrt[1 - ρ])] * Φ[(Γ - x2)/(σΓ * Sqrt[(1-ρ)/ρ])], {Γ, -1, 0}] +
    NIntegrate[1/(σΓ * Sqrt[1 - ρ]) * (1 - Γ) * ϕ[(Γ - ρ * x1)/(σΓ * Sqrt[1 - ρ])] * Φ[(Γ - x2)/(σΓ * Sqrt[(1-ρ)/ρ])], {Γ, 0, 1}];

Δ[x1_, x2_, ρ_, σΓ_] := FL[x1, ρ, σΓ] + FR[x1, x2, ρ, σΓ];

b1[x2_, ρ_, σΓ_, r_] :=
  Quiet[FindRoot[Δ[x1, x2, ρ, σΓ] == r , {x1, x1start}]][[1]][[2]];

randomX1 = ParallelMap[Apply[b1, #] & , parms];

parms = Join[randomX2, randomX2, randomRho, randomSigmaGamma, 2];

For[i = 1, i ≤ n, i++, parms[[i]][[1]] = randomX1[[i]]];

Db1Dx2[x1_, x2_, ρ_, σΓ_] := -(Quiet[Derivative[0, 1, 0, 0][Δ][x1, x2, ρ, σΓ]])/(Quiet[Derivative[1, 0, 0, 0][Δ][x1, x2, ρ, σΓ]]);
```

Figure C.7: Mathematica code for testing if the best response function is increasing (Part 1).

```
Print[""]
Print["Number of simulations" → n]
Print[""]
Print["Timing of simulation"]
time = AbsoluteTiming[results = ParallelMap[Apply[Db1Dx2, #] & , parms]][[1]]
Print[""]
Print["Number of violations"]
Print[""]
Length[Select[results, # <  0 &]]
Print[""]
Print["Magnitude of largest violation"]
Print[""]
min = Min[Select[results, # < 0 &]]
Print[""]
Print["Position of largest violation"]
Print[""]
pos = Position[results, min][[1]][[1]]
Print[""]
Print["Parameters of largest violation"]
Print[""]
Print["{x1,x2,ρ,σT}" → parms[[pos]]]

Number of simulations → 100 000

Timing of simulation

5447.589302

Number of violations

0

Magnitude of largest violation

∞

Position of largest violation

Part::partw : Part 1 of {} does not exist. >>
{}
```

Figure C.8: Mathematica code for testing if the best response function is increasing (Part 2 with results).

```
Clear["Global`*"]
n = 100 000;
SeedRandom[1234]
{capRho, capSigmaGamma, capR, capX2} = {1, 3, .5, 4};

randomRho = RandomReal[{0, capRho}, {n, 1}];
randomSigmaGamma = RandomReal[{0, capSigmaGamma}, {n, 1}];
randomR = RandomReal[{0.5, capR}, {n, 1}];
randomX2 = RandomReal[{0, capX2}, {n, 1}];

parms = Join[randomX2, randomRho, randomSigmaGamma, randomR, 2];

Φ[x_] := CDF[NormalDistribution[0, 1], x]; (* Standard normal CDF *)
ϕ[x_] := PDF[NormalDistribution[0, 1], x]; (* Standard normal PDF *)

(* starting values used in the computation of best respose b1(x2) *)
{x1start, x2start} = {0.01, -0.01};

(* Equilibrium *)
FL[x1_, ρ_, σΓ_] := Φ[(ρ * x1 - 1)/(σΓ * Sqrt[1 - ρ])] + σΓ * Sqrt[1 - ρ] * (ϕ[(ρ * x1)/(σΓ * Sqrt[1 - ρ])] - ϕ[(ρ * x1 - 1)/(σΓ * Sqrt[1 - ρ])]) +
    ρ * x1 * (Φ[(ρ * x1)/(σΓ * Sqrt[1 - ρ])] - Φ[(ρ * x1 - 1)/(σΓ * Sqrt[1 - ρ])]);

FR[x1_, x2_, ρ_, σΓ_] := NIntegrate[1/(σΓ * Sqrt[1 - ρ]) * (1 + Γ) *
      ϕ[(Γ - ρ * x1)/(σΓ * Sqrt[1 - ρ])] * Φ[(Γ - x2)/(σΓ * Sqrt[(1 - ρ)/ρ])], {Γ, -1, 0}] +
    NIntegrate[1/(σΓ * Sqrt[1 - ρ]) * (1 - Γ) * ϕ[(Γ - ρ * x1)/(σΓ * Sqrt[1 - ρ])] * Φ[(Γ - x2)/(σΓ * Sqrt[(1 - ρ)/ρ])], {Γ, 0, 1}];

Δ[x1_, x2_, ρ_, σΓ_] := FL[x1, ρ, σΓ] + FR[x1, x2, ρ, σΓ];

b1[x2_, ρ_, σΓ_, r_] :=
  Quiet[FindRoot[Δ[x1, x2, ρ, σΓ] == r , {x1, x1start}]][[1]][[2]];

randomX1 = ParallelMap[Apply[b1, #] & , parms];

parms = Join[randomX2, randomX2, randomRho, randomSigmaGamma, 2];

For[i = 1, i ≤ n, i++, parms[[i]][[1]] = randomX1[[i]]];


(* First derivative of the BR b1() wrt x2 *)
Db1Dx2[x1_, x2_, ρ_, σΓ_] := - Quiet[Derivative[0, 1, 0, 0][Δ][x1, x2, ρ, σΓ]] / Quiet[Derivative[1, 0, 0, 0][Δ][x1, x2, ρ, σΓ]];
```

Figure C.9: Mathematica code for testing concavity of the best response function (Part 1).

```
(* Second derivative of the BR b1() wrt x2 *)
D2b1Dx22[x1_, x2_, ρ_, σΓ_] := - (Quiet[Derivative[0, 2, 0, 0][Δ][x1, x2, ρ, σΓ]] +
      Quiet[Derivative[2, 0, 0, 0][Δ][x1, x2, ρ, σΓ]] * Db1Dx2[x1, x2, ρ, σΓ]) /
   Quiet[Derivative[1, 0, 0, 0][Δ][x1, x2, ρ, σΓ]];

Print[""]
Print["Number of simulations" → n]
Print[""]
Print["Timing of simulation"]
time = AbsoluteTiming[results = ParallelMap[Apply[D2b1Dx22, #] & , parms]][[1]]
Print[""]
Print["Number of violations"]
Length[Select[results, # >  0 &]]
Print[""]
Print["Magnitude of largest violation"]
min = Max[Select[results, # > 0 &]]
Print[""]
Print["Position of largest violation"]
pos = Position[results, min][[1]][[1]]
Print[""]
Print["Parameters of largest violation"]
Print["{x1,x2,ρ,σΓ}" → parms[[pos]]]


Number of simulations → 100 000


Timing of simulation

28 627.264610


Number of violations

0


Magnitude of largest violation

-∞


Position of largest violation

Part::partw : Part 1 of {} does not exist. >>

{}


Parameters of largest violation

{x1,x2,ρ,σΓ} → {}
```

Figure C.10: Mathematica code for testing concavity of the best response function (Part 2 with results).

The next picture shows that for any $\sigma_\Gamma > 0$, $b_1'(0)$ is concave in ρ and converges to 1 as $\rho \to 1$.

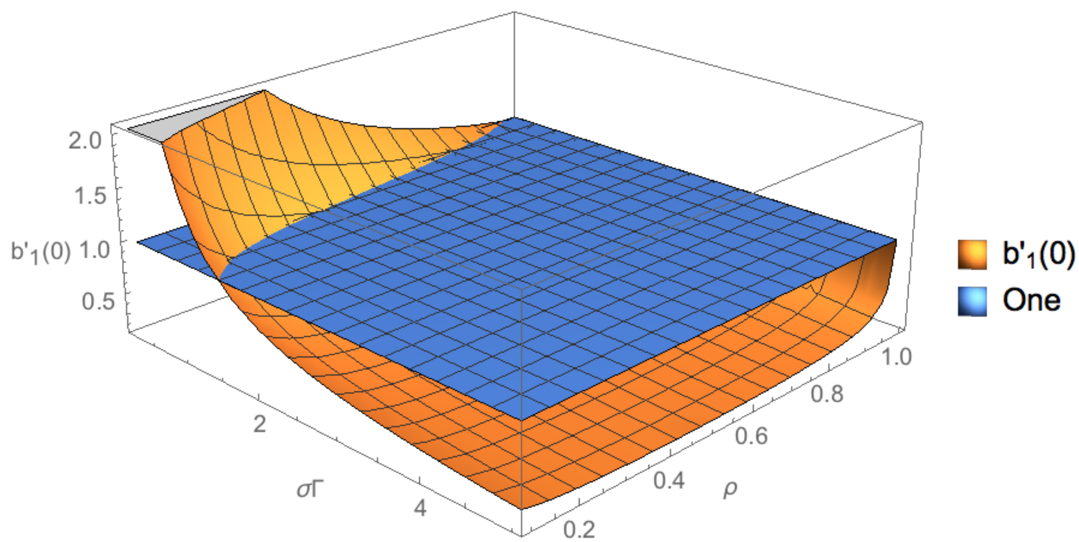


Figure C.11: Function $b_1'(0)$ evaluated at different combinations of (ρ, σ_Γ) in orange and the plane at one (blue).

C.5 Detailed computation

C.5.1 Useful Integrals

$$\int_{L_1}^{L_2} x\phi\left(\frac{x-a}{b}\right)dx = \frac{1}{2}b\left(b\left(e^{-\frac{(a-L_1)^2}{2b^2}} - e^{-\frac{(a-L_2)^2}{2b^2}}\right)\sqrt{\frac{2}{\pi}} + a\left(Erf\left(\frac{a-L_1}{b\sqrt{2}}\right) - Erf\left(\frac{a-L_2}{b\sqrt{2}}\right)\right)\right)$$

where

$$Erf(z) = \frac{2}{\sqrt{\pi}}\int_0^z e^{-t^2}dt = 2\Phi\left(z\sqrt{2}\right) - 1.$$

Hence,

$$\int_{L_1}^{L_2} x\phi\left(\frac{x-a}{b}\right)dx = b^2\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{(a-L_1)^2}{2b^2}} - \frac{1}{\sqrt{2\pi}}e^{-\frac{(a-L_2)^2}{2b^2}}\right) + ab\left(\Phi\left(\frac{a-L_1}{b}\right) - \Phi\left(\frac{a-L_2}{b}\right)\right)$$

$$= b^2\left(\phi\left(\frac{a-L_1}{b}\right)-\phi\left(\frac{a-L_2}{b}\right)\right)+ab\left(\Phi\left(\frac{a-L_1}{b}\right)-\Phi\left(\frac{a-L_2}{b}\right)\right) \quad \text{(C.5.1)}$$

We can always write $max\{X,a\} = X \times \mathbf{1}(X \geq a) + a \times \mathbf{1}(X \leq a)$, where $\mathbf{1}(.)$ is an indicator function. Then,

$$\begin{aligned} E[max\{X,a\}] &= E[X \times \mathbf{1}(X > a) + a \times \mathbf{1}(X \leq a)] \\ &= \int_{-\infty}^{+\infty} x\mathbf{1}(x > a) f_X(x)\,dx + a\int_{-\infty}^{+\infty} \mathbf{1}(x \leq a) f_X(x)\,dx \\ &= \int_{a}^{+\infty} x f_X(x)\,dx + aPr(X \leq a) \end{aligned}$$

When $X \sim N(\mu_X, \sigma_X^2)$, $f_X(x) = \frac{1}{\sigma_X}\phi\left(\frac{x-\mu_X}{\sigma_X}\right)$ and then

$$\begin{aligned} E[max\{X,a\}] &= \frac{1}{\sigma_X}\int_{a}^{+\infty} x\phi\left(\frac{x-\mu_X}{\sigma_X}\right)dx + aPr(X \leq a) \\ &= \sigma_X\phi\left(\frac{\mu_X-a}{\sigma_X}\right)+\mu_X\Phi\left(\frac{\mu_X-a}{\sigma_X}\right)+a\Phi\left(\frac{a-\mu_X}{\sigma_X}\right). \end{aligned}$$

Hence, for the particular case of $a = 0$:

$$E[max\{X,0\}] = \sigma_X\phi\left(\frac{\mu_X}{\sigma_X}\right)+\mu_X\Phi\left(\frac{\mu_X}{\sigma_X}\right).$$

In the conditional case:

$$\begin{aligned} E[max\{X,0\}|Y] &= E[X \times \mathbf{1}(X > 0)|Y] \\ &= \int_{-\infty}^{+\infty} x\mathbf{1}(x > a) f_{X|Y}(x|y)\,dx \\ &= \int_{a}^{+\infty} x f_{X|Y}(x|y)\,dx \end{aligned}$$

C.5.2 Conditional Expectation Normal Random Variable

Let $Y \sim N(\mu, \sigma^2)$. We want to compute $E[Y|Y > \tau]$.

$$E[Y|Y > \tau] = \int_{\tau}^{+\infty} y f(y|y > \tau)\, dy$$

where

$$f(y|y > \tau) = \frac{\frac{1}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)}{1 - \Phi\left(\frac{\tau-\mu}{\sigma}\right)} = \frac{\frac{1}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)}{Pr(Y > \tau)}.$$

Then

$$E[Y|Y > \tau]\, Pr(Y > \tau) = \frac{1}{\sigma}\int_{\tau}^{+\infty} y\phi\left(\frac{y-\mu}{\sigma}\right) dy$$

and using the result from last section:

$$\begin{aligned}\int_{\tau}^{+\infty} y\phi\left(\frac{y-\mu}{\sigma}\right) dy &= \sigma^2\phi\left(\frac{\mu-\tau}{\sigma}\right) + \mu\sigma\Phi\left(\frac{\mu-\tau}{\sigma}\right) \\ &= \sigma^2\phi\left(\frac{\mu-\tau}{\sigma}\right) + \mu\sigma\left(1 - \Phi\left(\frac{\tau-\mu}{\sigma}\right)\right)\end{aligned}$$

hence

$$E[Y|Y > \tau]\, Pr(Y > \tau) = \sigma\phi\left(\frac{\mu-\tau}{\sigma}\right) + \mu\left(1 - \Phi\left(\frac{\tau-\mu}{\sigma}\right)\right)$$

or

$$E[Y|Y > \tau] = \mu + \sigma\frac{\phi\left(\frac{\mu-\tau}{\sigma}\right)}{\left(1 - \Phi\left(\frac{\tau-\mu}{\sigma}\right)\right)}.$$

C.5.3 Conditional Distribution (Multivariate Normal)

Theorem. *Let*

$$\begin{pmatrix} Y \\ X \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} Y \\ X \end{pmatrix}, \begin{pmatrix} \Sigma_{YY} & \Sigma_{YX} \\ \Sigma_{XY} & \Sigma_{XX} \end{pmatrix}\right],$$

where $Y \in \mathbb{R}$, $X \in \mathbb{R}^k$ and Σ_{XX} is nonsingular. Then, conditionally on X, Y is normally distributed

$$Y|X \sim \mathcal{N}\left(\mu_Y + \Sigma_{YX}\Sigma_{XX}^{-1}\left(X - \mu_X\right), \Sigma_{YY} - \Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY}\right).$$

Let $x_1 = \Gamma + \epsilon_1$ and $x_2 = \Gamma + \epsilon_2$, where

$$\Gamma \sim N\left(\mu, \sigma_\Gamma^2\right)$$

and

$$\epsilon_1, \epsilon_2 \overset{iid}{\sim} N\left(0, \sigma_\epsilon^2\right).$$

Hence, $cov\left(x_1, x_2\right) = \sigma_\Gamma^2 > 0$. Since

$$\left(\begin{array}{c} x_2 \\ \\ x_1 \end{array}\right) \sim N\left(\left(\begin{array}{c} \mu \\ \\ \mu \end{array}\right), \left(\begin{array}{cc} \sigma_\Gamma^2 + \sigma_\epsilon^2 & \sigma_\Gamma^2 \\ \\ \sigma_\Gamma^2 & \sigma_\Gamma^2 + \sigma_\epsilon^2 \end{array}\right)\right),$$

the distribution of x_i conditional on x_j is given by

$$x_i|x_j \quad \sim \quad N\left(\rho x_j + \left(1 - \rho\right)\mu, \sigma_\Gamma^2\left(\frac{1 - \rho^2}{\rho}\right)\right)$$

Moreover,

$$\left(\begin{array}{c} \Gamma \\ \\ x_1 \end{array}\right) \sim N\left(\left(\begin{array}{c} \mu \\ \\ \mu \end{array}\right), \left(\begin{array}{cc} \sigma_\Gamma^2 & \sigma_\Gamma^2 \\ \\ \sigma_\Gamma^2 & \sigma_\Gamma^2 + \sigma_\epsilon^2 \end{array}\right)\right),$$

the distribution of Γ conditional on x_j is given by

$$\Gamma|x_j \quad \sim \quad N\left(\rho x_j + \left(1 - \rho\right)\mu, \left(1 - \rho\right)\sigma_\Gamma^2\right)$$

C.5.4 Computation of $\mathbb{E}\left[max\left\{q+\Gamma,0\right\}\right]$

$$\begin{aligned}
\mathbb{E}\left[max\left\{q+\Gamma,0\right\}\right] &= \frac{1}{\sigma_\Gamma}\int_{-q}^{+\infty}\left(q+\Gamma\right)\phi\left(\frac{\Gamma}{\sigma_\Gamma}\right)d\Gamma \\
&= q\int_{-q}^{+\infty}\frac{1}{\sigma_\Gamma}\phi\left(\frac{\Gamma}{\sigma_\Gamma}\right)d\Gamma+\frac{1}{\sigma_\Gamma}\int_{-q}^{+\infty}\Gamma\phi\left(\frac{\Gamma}{\sigma_\Gamma}\right)d\Gamma \\
&= q\Phi\left(\frac{\Gamma}{\sigma_\Gamma}\right)+\sigma_\Gamma\phi\left(\frac{q}{\sigma_\Gamma}\right)
\end{aligned}$$

where the last row uses (C.5.1).