THÈSE de DOCTORAT

de l'UNIVERSITÉ TOULOUSE CAPITOLE

Présentée et soutenue par

Monsieur Oscar Diego JARA RAMíREZ

Le 1 juillet 2024

Essais en Organisation Industrielle Empirique

École doctorale : **TSE Toulouse Sciences Economiques**
Spécialité : **Sciences Economiques - Toulouse**
Unité de recherche : **TSE-R - Toulouse School of Economics - Recherche**

Thèse dirigée par *Monsieur Mathias REYNAERT*

Composition du jury
Rapporteure : Mme Claire CHAMBOLLE
Rapporteur : M. Joel STIEBALE
Examinateur : M. Morten SAETHRE
Directeur de thèse : M. Mathias REYNAERT
Co-directrice de thèse : Mme Isis DURRMEYER
Co-directeur de thèse : M. Pierre DUBOIS

Essays in Empirical Industrial Organization

Oscar Jara

May 23, 2024

Contents

Introduction

English version:

The first two chapters focus on understanding the effect of market structure on equilibrium outputs like prices or product offerings in the markets of energy drink distribution and mobile telecommunications. The third chapter explores methodologies for measuring the effects of leniency programs on cartel formation.

In the first chapter, I develop a new structural model of bargaining to evaluate the impact on prices and welfare of a consolidation of upstream firms operating in different geographic markets. The model considers that upstream and downstream firms negotiate for market-specific wholesale prices for many markets simultaneously, instead of market by market independently. I call these two ways of bargaining multi and single market bargaining. A consolidation of upstream firms present in different geographic markets affects the negotiation process and, through multimarket bargaining (MMB), impacts downstream prices. I apply this model to a consolidation of regional distributors in the U.S. energy drinks market, where one of the leading brands transitioned from having two national distributors to only one. National retailers with stores in both affected and unaffected areas began negotiating with one distributor instead of two. The theoretical model predicts that under MMB, retail prices change in every region, not just those affected by the consolidation. Empirical evidence supports this prediction, showing that national retailers reduced their prices by 1.5% in regions directly affected by the consolidation and by 1.6% in indirectly affected regions. The structural model further reveals that when the regional distributors expand into new regions, their bargaining position relative to national retailers weakens. This results in better deals for national retailers, effectively lowering retail prices.

In the second chapter, jointly with Nicolas Martinez, we study the effects of firm entry and technological progress on product offerings and pricing in the Peruvian mobile telecommunications market. The chapter examines how the entry of two new firms and the introduction of 4G technology affected a market previously dominated by two incumbents. After the new firms entered the market, the incumbents began offering more plans with higher variation in data allowances. This change coincided with the introduction of 4G connectivity, which significantly altered consumers' valuation of data and thus demand patterns. Reduced form evidence shows that prices for prepaid tariffs decreased significantly, while postpaid plan prices remained constant or slightly increased. We then use a structural model of supply and demand to analyze the joint effects of competition and technological progress. The demand estimation shows a significant increase in consumer willingness to pay for data after the introduction of 4G, highlighting the crucial role of technological change. The findings indicate that both competition and technological advancements drive changes in market outcomes, with technological progress leading to higher consumer valuations for data and more varied product offerings from incumbents.

The third chapter addresses the challenges of prosecuting cartels due to their secretive nature and examines the effectiveness of leniency programs in detecting and deterring cartels. Leniency programs incentivize cartel members to come forward with evidence, aiming to destabilize existing cartels and deter the formation of new ones. However, measuring the efficacy of these programs is challenging because the full population of cartels is not observed. I propose using Hidden Markov Models (HMM) to estimate the effects of leniency programs, as HMMs can identify unobservable latent states by analyzing observable data that indirectly indicate underlying processes. This chapter emphasizes the importance of measuring the effects of leniency programs and that other policy changes should not disrupt their workings.

French version:

Les deux premiers chapitres étudient l'effet de la structure du marché sur les prix et les offres de produits dans les secteurs des boissons énergétiques et des télécommunications mobiles. Le troisième chapitre examine comment mesurer l'impact des programmes de clémence sur les cartels.

Le premier chapitre développe un modèle de négociation pour évaluer l'impact d'une consolidation des entreprises en amont sur les prix et le bien-être. Le modèle considère que les entreprises en amont et en aval négocient des prix de gros pour plusieurs marchés simultanément, appelés négociation multi-marché (NMM). Une consolidation affecte le processus de négociation et, via la NMM, impacte les prix en aval. En appliquant ce modèle à une consolidation des distributeurs régionaux de boissons énergétiques aux États-Unis, il est prévu que les prix changent dans toutes les régions, pas seulement celles affectées par la consolidation. Les preuves empiriques montrent que les détaillants ont réduit leurs prix de 1,5 % dans les régions directement touchées et de 1,6 % dans les régions indirectement affectées. Le modèle montre que l'expansion des distributeurs affaiblit leur position de négociation, ce qui conduit à des baisses de prix pour les détaillants.

Le deuxième chapitre, en collaboration avec Nicolas Martinez, analyse les effets de l'entrée de nouvelles entreprises et des progrès technologiques sur les offres et les prix dans le marché des télécommunications mobiles au Pérou. L'entrée de deux nouvelles entreprises et l'introduction de la technologie 4G ont bouleversé un marché dominé par deux acteurs. Les entreprises établies ont commencé à offrir plus de plans avec des variations dans les allocations de données. La connectivité 4G a modifié la valeur des données pour les consommateurs. Les tarifs prépayés ont baissé, tandis que les forfaits postpayés sont restés constants ou ont légèrement augmenté. Un modèle structurel de l'offre et de la demande montre une augmentation de la volonté de payer des consommateurs pour les données après l'introduction de la 4G, soulignant le rôle du progrès technologique. Les résultats montrent que la concurrence et les avancées technologiques influencent les résultats du marché, avec des offres plus variées et une valorisation accrue des données.

Le troisième chapitre traite des défis de la poursuite des cartels et de l'efficacité des programmes de clémence. Ces programmes encouragent les membres des cartels à fournir des preuves, déstabilisant ainsi les cartels existants et dissuadant la formation de nouveaux. Cependant, il est difficile de mesurer leur efficacité car la totalité des cartels n'est pas observable. J'utilise des modèles de Markov cachés (MMC) pour estimer les effets des programmes de clémence, en analysant des données qui indiquent des processus sous-jacents non observables. Ce chapitre souligne l'importance de mesurer ces effets et de veiller à ce que d'autres changements de politique ne perturbent pas leur fonctionnement.

Non-technical summary

English version:

The three chapters in this study explore how changes in market structures impact prices, product offerings, and the effectiveness of regulations in different industries. In the first chapter, the focus is on the U.S. energy drinks market. When Monster Energy Drinks switched from two distributors to just one, it affected the negotiation process with retailers. I find that the consolidation made it easier for retailers to negotiate better deals, leading to lower prices for consumers. The second chapter looks at the Peruvian mobile telecommunications market. When two new companies entered the market and 4G technology was introduced, the existing companies started offering more varied plans. This increased competition and new technology led to lower prices for prepaid plans and more choices for consumers. The third chapter examines how leniency programs help catch cartels—groups of companies that secretly fix prices. These programs encourage cartel members to report their illegal activities. The study suggests using advanced models to better understand how effective these programs are. It highlights the need for supportive policies to ensure these programs continue to help uncover and deter cartels.

French version:

Le premier chapitre examine comment la concurrence spatiale et les contraintes de capacité influencent le bien-être dans les marchés du transport. Les résultats mettent en lumière l'interaction complexe entre la disponibilité du service, la qualité et les contraintes réglementaires, avec des implications pour le bien-être et la distribution du service. Dans le deuxième chapitre, nous développons un modèle pour comprendre les conditions de circulation dans une ville. Nous prenons en compte divers facteurs qui influencent le choix du mode de transport et des heures de départ par les personnes. De plus, nous tenons compte de la variation de la congestion routière à travers la ville. Nous appliquons notre modèle aux données de la région métropolitaine de Paris et formulons des recommandations pour différents types de politiques. Le troisième chapitre se concentre sur le rôle de la structure du marché dans la formation des offres de produits dans le secteur des télécommunications mobiles, en particulier en réponse à l'introduction de la connectivité 4G et à l'augmentation de la concurrence au Pérou.

Chapter 1

Multimarket Upstream Consolidation: Evidence from the US Energy Drinks Market

Oscar Jara[1]

Abstract:

Contractual arrangements between upstream and downstream firms can involve different geographic markets, depending on the coverage of each one. In this paper, I show that a consolidation between upstream firms operating in different geographic markets can generate price effects. To observe this outcome, the region specific negotiated price must be bargained simultaneously across all regions and not independently region by region. I refer to these approaches as *multi* and *single* market bargaining, respectively. Under *multimarket* bargaining, the expansion of a distributor into new regions, all else being equal, generates price effects in both new and legacy regions. To empirically explore these effects, I study the consolidation of distributors in the U.S. energy drinks market. Post-consolidation, reduce form evidence shows that national retailers decreased their prices in every region; indicating that negotiations are *multimarket*. Then I build a structural model of *multimarket* bargaining, revealing that distributors rely more on retailers than retailers rely on distributors. This caused the decrease

[1]Toulouse School of Economics, Université Toulouse 1 Capitole.
I thank Pierre Dubois, Isis Durrmeyer and Mathias Reynaert for their valuable advice. I acknowledge financial support from the European Research Council under grant ERC-2019-STG-852815 'PRIDISP'. Researcher(s)' own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

in distributors' bargaining position that lead to reduction in retail prices.

JEL Classification: D40, L11, L81

1.1 Introduction

In markets such as the retail, pay-TV, or healthcare sectors, prices result from negotiations between upstream and downstream firms. Whereas firms' bargaining positions determine whether they secure a favorable deal, variations in the structure of the market, like upstream or downstream mergers, shift it. While some structural changes may be limited to a geographic market, this is not always the case. Industries like retail involve firms operating in multiple regions, meaning that changes in the number of firms affect simultaneously multiple regions. Although the welfare effects of mergers between upstream firms in the same market have been studied, there is comparatively less understanding of mergers between upstream firms present in multiple markets. In this paper, I investigate how a consolidation of regional distributors affect retail prices and welfare and, if so, by which mechanisms.

To evaluate the impact of changes in the upstream market structure on retail prices, I study the consolidation of regional distributors in the U.S. energy drinks market. In this industry, soft drink manufacturers delegate the production and distribution of their products to regional distributors, each with exclusive non-overlapping territories. In March 2015, Monster Energy Drinks (hereafter referred to as Monster) designated The Coca-Cola Company (hereafter referred to as TCCC) as its sole national distributor, terminating its contract with Anheuser-Busch (hereafter referred to as AB) and extending TCCC distribution territories. Regions previously covered by the distribution system of TCCC did not experience changes from the Monster -TCCC agreement. However, in the regions where TCCC expanded, retailers experienced a change in the distributor they were dealing with. Particularly, national retailers with stores in both regions affected and not affected by the consolidation, started negotiating with one distributor instead of two. In this paper, I aim to address whether the expansion of regions supplied shifts the bargaining positions of retailers and the newly consolidated distributor, TCCC, and the mechanism under which this happens.

I study a setting where both retailers and distributors are present in multiple regions. I show that as regional distributors expand into new regions, their bargaining position weakens in comparison to national retailers. This benefits retailers in securing better deals. Then, a consolidation of distributors acts as a downward pressure force on prices. This result holds true when negotiations span multiple markets simultaneously, and not when they are conducted independently for each market. I call these negotiation protocols *multi-market bargaining* (MMB) and *single-market bargaining* (SMB) respectively. The distinction lies in the outside option of a negotiation: SMB stops supply in one market if negotiations fail, while MMB withholds the supply across all markets. Whether MMB or SMB are used depends on the institutional arrangements of each industry and is not always observed by the researcher.

In this paper, I develop a new structural model of bargaining to evaluate the impact on prices and consumer surplus under alternative negotiation protocols. First, I employ a theoretical model of bargaining where after a consolidation of distributors, the MMB protocol predicts a retail price change in every region and not only those directly affected by the consolidation. Then I test this prediction using a reduced form approach. I take advantage of the regional change of distributors and how it affected retailers differently to identify the price effects. National retailers reduced their prices by 1.5% in the regions directly affected by the consolidation, and by 1.6% in the regions indirectly affected. The price change in both groups of regions suggests that the observed data is likely to come from MMB mechanism. However, the stores used as base group may have also changed their prices, as they were competing with firms in the indirectly affected regions. Results coming from the reduced form model might capture the changes done in equilibrium by the firms in the comparison group, disabling obtaining the effects from the consolidation alone.

To further understand these effects, I use a structural model of bargaining, where I can specify if the firms follow a multi or single market negotiation protocol. I find that in the regions affected by a change in distributors, retailers, on average, increased their bargaining position. This is mainly explained by the change in the gains from trade. Although after the consolidation the gains from trade increased for both retailers and TCCC distributor, the increase for the distributor is higher; leading to a weakening of distributor's bargaining position. Using these results, I construct counterfactual scenarios to assess the effects on consumer surplus and prices. I find that compared to a benchmark counterfactual where there is the structure of the market, national retailers decreased their prices by 1.3% in the directly affected regions, and by 2.8% in the indirectly affected regions. For the group of regional retailers, the consolidation of distributors led to decrease their prices in 3.8%. Finally, while under the assumption of multi-market bargaining the model predicted a price decrease, the single-market bargaining model yields a price increase prediction. The contrasting predictions highlight the importance of the assumption on bargaining protocols in vertically-structured markets.

Although the studied setting is not a merger, it resembles the effects of an upstream merger between firms in different geographic markets. This paper shows that the entire vertical structure of the market must be considered when a merger involves firms in different geographic markets and the MMB is the selected bargaining protocol; even if they do not compete across markets. The importance of the previous findings lay on the fact that a consolidation of distributors might improve the bargaining position of the retailers. Even if the retail sector is highly concentrated, a price decrease might still arise from a consolidation of distributors.

The previous results can be useful for competition policy as well. Mergers or acquisitions between retailers that cover many regions change the bargaining position of the new entity against its distributors. While this could have a negative effect on prices by restricting competition, a stronger bargaining position could lead to a lower wholesale price and hence lower retail prices when distribution markets are highly concentrated. So, antitrust authorities must consider the full vertical structure when analyzing mergers between retailers.

This paper contributes to three strands of the literature. First, it complements the previous literature on upstream mergers in vertical relationships, by accounting for the importance of the geographical coverage of the upstream firms. In fact, some papers in the health literature (Gowrisankaran et al., 2015; Ho and Lee, 2017, 2019; Dafny et al., 2019) assume that after a merger, the negotiation for the price between hospitals and insurers is local, disregarding the possibility of multi-market bargaining. In this paper, I show that multi-market contracting is a possibility that must be considered both by academics and policymakers, since it can generate different price predictions. Naturally, whether multi-market contracting or single-market contracting happens will depend on each specific industry.

Similarly, Dafny et al. (2019) and Lewis and Pflum (2015) show that for mergers between hospital that are in the same state but not too close, price effects from the merger still can emerge. They show that this happens because consumers purchase in different markets. In this paper I complement their results by showing that a change in the bargaining positions after a consolidation of upstream firms can lead to a new equilibrium, even when consumer do not purchase in multiple markets. In that same line, other papers show that changes in the bargaining power parameter can be used as a source of price variations (Grennan, 2013; Lewis and Pflum, 2015; Gowrisankaran et al., 2015; Lewis and Pflum, 2017). Since in this paper I work in a setting with many markets, I study the changes in bargaining position and not only the variation in the bargaining power parameter of the firms.

Second, I contribute to the empirical literature on bargaining in markets with vertical structures (Villas-Boas, 2007; Draganska et al., 2010; Bonnet and Dubois, 2010, 2015). I develop a novel yet tractable way of modeling negotiation for market specific wholesale prices. Since I do not have access to wholesale data, I extend the model developed by Draganska et al. (2010) to include region specific wholesale prices and firms' margins and measure of relative bargaining power. In the same line, this paper contributes to the literature on merger simulation (Sheu and Taragin, 2021; Panhans and Taragin, 2022) by showing how bargaining for all the regions at the same time, can lead price effects; such that the inclusion of the entire vertical structure is necessary for estimation.

Lastly, I also contribute to the literature on retail pricing. While previous papers (Adams and Williams, 2019; DellaVigna and Gentzkow, 2019; Butters et al., 2022) focus on retail price variation to local demand or supply shocks, I study the price effects of a local shock to the upstream structure of the market. Other recent papers (Ganapati, 2018; Döpper et al.,

2022) highlight that in the retail sector, the primary source of increase in markups comes from cost savings. I contribute to this literature by studying how changes in bargaining positions affect the retail prices. I show that for the energy drinks industry, part of that increase in retailers' markups comes from a strengthening of their bargaining power after a consolidation of distributors.

In the next section I describe the industry, the context in which this paper is placed, and the data sources used for the estimation. Section 1.3 presents the theoretical model that explains the observed reduced form evidence. Section 1.3.2 develops the structural model employed to get the results in section 1.4. In section 1.5, I compute the counterfactuals. Finally, in the last section, based on the previous results, I conclude.

1.2 The Industry

1.2.1 The Market

In the soft drinks industry, some brands only produce concentrate, the main ingredient of the beverages, and sell it at a linear tariff to local distributors. These last ones are in charge of the finalizing production by adding water, carbon dioxide, and additional sweeteners and flavors. After this process, the distributors pour the beverage into cans and take them to the retailers. In the US, distributors possess exclusive territories in which they are the only producers and distributors. Distribution regions are negotiated with the brand owner such that there are not two distributors for the same market. To get supplies of a beverage for a specific location, the retailer must negotiate with the authorized local distributor.[2][3]

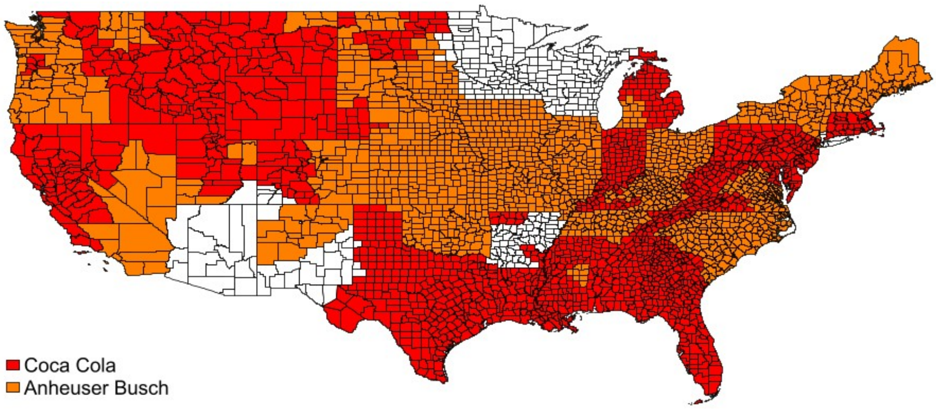
Not all firms are vertically integrated with their distributors. In the energy drinks market, the three leading brands Monster, Rockstar, and Red Bull captured around 70% of market share in 2014. From them, only this last one is vertically integrated with its distributors. Rockstar is distributed by PepsiCo, and Monster had distribution agreements with The Coca-Cola Company (TCCC) and Anheuser-Busch (AB). A common practice in this industry is that a distributor cannot produce or distribute a rival brand, i.e., TCCC cannot distribute products from Red Bull or Rockstar. Figure 1.1 shows the main distribution zones for TCCC (red) and AB (orange) distributors.

In March 2015, TCCC and Monster signed a partnership agreement by which TCCC became

[2] Additionally, distributors engage in some promotional activities (sales, shelf space, among others)

[3] According to their public contracts, Monster representatives can be present at the moment of the negotiation between the distributors and the retailers or even suggest prices

Figure 1.1: Distribution zones before consolidation.
Note: Based on annual presentations to investors of Monster, US Securities and Exchange Commission. The map shows the territories in which TCCC and AB had the right to distribute the products of Monster. The areas in white, not colored, were under the distribution of third parties independent distributors and are not considered in this paper.



Monster's only distributor in the US, TCCC bought a 16.7% equity stake in Monster.[4][5] By having TCCC as its only supplier, there were gains from better coordination on sales, shelf space, and banners, among others.[6] The left panel in Figure 1.2 shows the original arrangement of distributors. The one on the right shows the situation after the agreement between Monster and TCCC.

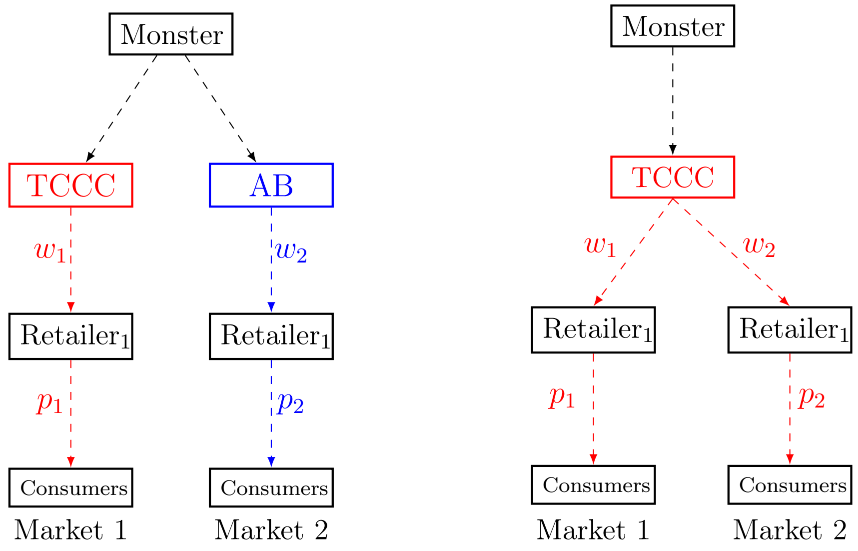
Since energy drinks contain legal stimulants such as taurine, caffeine, group B vitamins, guarana, and L-carnitine, they stand apart from conventional soft drinks. In this paper, I consider that soft drinks are not a close substitute for energy drinks. On the demand side, energy drinks are typically consumed because of its effects on energy enhancement, reduction of fatigue, and mental alertness, characteristics that standard soft drinks do not have. On the supply side, firms' advertisement strategies and internal reports reflect that they consider only other energy drinks as competitors. In supermarkets, they are not typically on the same shelves as the soft drinks and are usually displayed in a separate section. Due to their composition, they are primarily sold in smaller units (8 to 16oz) compared to the popular soft drinks 20, 42.2, and 67.6oz of the traditional sodas. In this sense, since Monster and TCCC's products are not close substitutes, there are no reasons for TCCC to set prices for Monster products strategically.

[4]The agreement also involved other provisions. Mainly, Monster and TCCC also they switched their portfolios of each other's non-core products: Monster gave TCCC its non-energy drinks brands and TCCC gave Monster its energy drinks portfolio. Additionally, Monster must pay to TCCC to use its distributors' network. The exact amount of this linear tariff is confidential.

[5]Strategically, with this agreement TCCC secured its presence in a growing market. In 2015, the US energy drinks sector's sales were growing at 5.5%, while those of the soft drinks were growing at 1.9%

[6]According to the presentations to investors, Monster was getting access to more markets internationally, and by giving its non-energy portfolio to TCCC, it would benefit from focusing only on energy drinks.

Figure 1.2: Market Structure



(a) Before Consolidation (b) After Consolidation

Throughout this paper, I assume two crucial facts about this industry. The first is that distributors and retailers bargain for wholesale prices for different regions of the US. Given the size of the US and high transportation costs, it is reasonable to assume a regional wholesale price instead of a national one. Second, I assume TCCC negotiates on behalf of its franchisees. Since I do not observe any significant difference in prices among regions between TCCC and its franchisees, I consider this a valid assumption.[7]

Both Monster's distributors, TCCC and AB, have its own portfolio of products. TCCC distributes Coca-Cola products, and the Fanta and Sprite brands, among others. On the other hand, AB is the owner of many of the most important beer brands in the US like Budweiser and Busch Beer. Negotiating bundles of products emerges as a realistic possibility in the retail sector. However, it is more likely that being part of the TCCC portfolio affected Monster, but not the other way around. This trait should be captured by the bargaining power of the distributor. I will test if there were price effects on TCCC products.[8]

Finally, it is important to mention other structural changes that occurred in the soft drink market in the US close to the investigated period, 2012-2017. In 2009, TCCC vertically integrated with many of its distributors, becoming the major distributor in the US soft drinks market. However, in 2018 TCCC gave back the property of the production facilities to local owners, who became the only TCCC distributors in the US, each with an exclusive territory.

[7]Additionally, TCCC was the biggest Coca-Cola distributor during the analyzed period, 2012-2017.

[8]During a dispute between TCCC and the wholesaler Costco the Associated Press reported that the beverage brands no longer being sold at Costco include *"Coke Classic, Cherry Coke, Black Cherry Vanilla Coke, Diet Coke, Coke Zero, Sprite and Squirt, Dasani Water and Vitamin Water along with several energy drinks."*

Further away from the analyzed period, in 2020, PepsiCo bought Rockstar for $3.85 billion.

1.2.2 Data Sources and data preparation

I use two primary data sources in this project; one related to sales and prices and another one associated with distribution zones and plant locations. Product characteristics come from the Retail Scanner and Consumer Panel Data of Kilts NielsenIQ for 2012 - 2017. This dataset contains detailed information regarding prices, quantities sold, units, size, among other product characteristics for more than 35 thousand stores in the US each week. I use the period 2012 - 2017 because the change in bottler by Monster was in 2015, and in 2010 and 2018, there were other significant changes in this market, as detailed in the paragraph above.

In the NielsenIQ dataset, a store has a unique code associated with its parent company. Following other papers in the literature (DellaVigna and Gentzkow, 2019), I do not include stores that changed their parent company nor stores that were not in the sample for at least five of the six years. An equal time criterion is applied to retail chains. I exclude liquor stores from the sample since they do not appear during the whole period.

Each product is associated with a universal product code (UPC). Since a new UPC is created for any product variation, I aggregate products at the brand level.[9] A brand is defined as a specific product line of a firm, like Monster Energy Ultra Blue Sugar-Free. Products are generated by combining the brand with one of the possible sizes 8oz, 12oz, and 16oz, like Monster Energy Ultra Blue Sugar-Free 16oz. The price of a product is defined as the revenues over all the units sold (Nevo, 2001; Miller and Weinberg, 2017). Following Nevo (2001), the time scope used is at the month level to avoid the effects of sales and stockpiling. Since meeting to negotiate weekly can be costly to firms, monthly aggregation becomes a reasonable assumption.

I restrict the sample to 29 DMA (Designated Market Area) that have more than 500 household participants yearly in the consumer panel dataset. I combine DMAs with the distribution regions of the bottlers to generate 40 unique regions. For each of these markets, I generate consumer-specific demographic draws from the Consumer Panel Data by sampling with replacement 300 consumers monthly using NielsenIQ's projection weights. From this, I get the average income of the households by retail chain, the age of the household's head, the number of kids, and some additional statistics on the income distribution.

The second dataset I employ is related to distribution areas. I get the information about the distribution zones from the annual presentations to investors that Monster held during the analyzed period, 2012-2017. These presentations must be submitted to the US Securities and

[9] The change of UPC includes not only changes in the size or number of units, but also minor changes like the variations in the packaging.

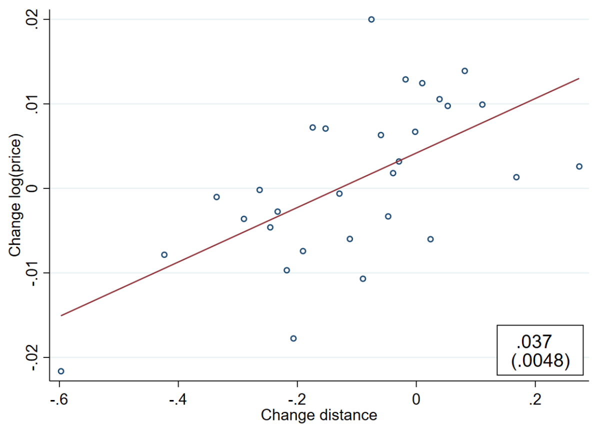
Table 1.1: Distributors market coverage

Distributor	Stores	Chains	Counties	States
TCCC	7,362	51	313	20
AB	5,851	38	233	16

Exchange Commission (henceforth SEC) and are public information. I complement this data with additional documentation submitted to the SEC when TCCC and Monster signed their new distribution agreement, where changes to the distribution areas were introduced.

The change in plan location when the production passed from AB to TCCC changed the driving distances from the plant to the stores. I assume energy drinks are shipped by truck to each store from the nearest brewery. Likewise, I construct a dataset with the plant locations for the bottlers in this industry. I only consider the production facilities that were operating from the year 2012 to 2017. Figure 1.3 shows the distribution of changes in prices and driving distances (in 100 of miles) for Monster 16oz products. For AB, I consider the facilities reported on its webpage. I obtain the latitude and longitude of the plants for both sets of production facilities using Google Maps. Then, using the API from TomTom, I obtain the driving distance between the center of the county where each store is located and the exact location of the production facility. I assume the only time with a change in the production facility is at the time of the agreement between TCCC and Monster. For the other energy drinks in the market, I assume that there are no new production facilities, such that changes in driving distance are equal to zero.

Figure 1.3: Relation between distances and price variation - Monster (16oz)



1.3 The Model

In this section I present a model where wholesale prices are negotiated between distributors and retailers. Distributors have non-overlapping distribution territories and retailers supply multiple markets. I assume the negotiation process follows a Nash bargaining protocol, where firms negotiate for market specific prices. When an upstream distributor supplies multiple market; the new negotiation includes the wholesale price for all the new markets. The question in this section is whether changes in the upstream structure of the market can change equilibrium and under which mechanism.

I assume that firms negotiate either for all the market at the same time or for each one independently.[10] I term these alternative negotiation protocols multi and single market bargaining, respectively. Whether the inclusion of more regions in the negotiation process affects firms' bargaining power, depends on the bargaining protocol followed by them. The model predicts that, everything else constant, only under multi-market bargaining a consolidation of upstream distributors affects prices in all the regions where the retailer has stores. Under the alternative single-market bargaining, a national retailer only changes its prices in the areas directly affected by the consolidation.[11]

First I introduce the theoretical model to be followed in the rest of the paper. Second, using this model I state some hypothesis to be tested in the data. Finally, at the end of this section, I generalize the theoretical model to build an empirical structural model of bargaining.

1.3.1 A model of Multi-Market Bargaining

Consider a setting with two markets $m = \{A, B\}$ and one retailer r that has stores in both markets. There is only one product that is distributed by a different distributor in each market. The retailer, which has no retail costs, negotiates the supply of the product for each store with each local distributor. Upon reaching an agreement with the distributor in market m, retailer's profits are $\pi^r_m = (p_m - w_m)D_m(p_m)$; where p_m, w_m and $D_m(p_m)$ are the retail price, wholesale price and demand for the good in market m, respectively. Distributors' profits when arriving to an agreement with the retailer are $\pi^d_m = (w_m - \mu_m)D_m$, where μ_m is the marginal cost of production.

Following Horn and Wolinsky (1988), I assume that each wholesale price negotiation between

[10] Although contracts are not observed, anecdotal evidence suggests that under a negotiation breakdown, all the stores from a retail chain stop getting supplies from the distributor.

[11] When demand considers goods or services in different locations as substitutes, bargaining over wholesale prices for a higher number of markets can change the outside options of the firms (Dafny et al., 2019; Vistnes and Sarafidis, 2013). In the retail sector, it is not a suitable assumption to think that consumers consider stores in different regions as close substitutes.

the two distributor and the retailer solves a bilateral Nash-in-Nash bargaining, and all the negotiations are carried out simultaneously and independently. The bargaining power weight of retailer r when negotiating with distributor d is, β_{rd}; and its complement, $1-\beta_{rd}$, represents the bargaining power weight of distributor d. Retail price setting takes place at the same time as the wholesale negotiation. In case of a negotiation breakdown for market m, retailer and distributor gain no profits in that market.

First consider, consider the negotiation process of the retailer with each individual distributor.

Single Market Bargaining (SMB): When the retailer negotiates for the wholesale price of each market m separately, w_m maximizes the weighted profits:

$$w_m = \arg\max_{w} \left[(p_m - w_m)D_m\right]^{\beta_{rd}}\left[(w_m - \mu_m)D_m\right]^{1-\beta_{rd}}, \tag{1.1}$$

which yields the equilibrium wholesale price $w^*_m = (1-\beta_{rd})p^*_m + \beta_{rd}\mu_m$.

Now consider a consolidation of the two distributors into one monopolist distributor that covers both regions. Under this new market structure, the retailer needs to negotiate with the consolidated distributor for the supply of the product for both markets. If the retailer and the distributor decide to negotiate for each market independently, *single-market bargaining*, wholesale prices still solves equation 1.1. Everything else constant, the equilibrium prices are the same before and after the consolidation. Instead, if they decide to negotiate for all the markets at the same time, *multi-market bargaining*, wholesale prices solves equation 1.2.

Multi Market Bargaining (MMB): Firms negotiate for both w_A and w_B at the same time,

$$\{w^{**}_A, w^{**}_B\} = \arg\max_{w_A, w_B} \left[\sum_{m\in\{A,B\}} (p_m - w_m)D_m\right]^{\beta_{rd}} \left[\sum_{m\in\{A,B\}} (w_m - \mu_m)D_m\right]^{1-\beta_{rd}} \tag{1.2}$$

The solution to equation 1.2 for market A, alongside retailer's first order conditions, can be expressed as:

$$\begin{aligned} w^{**}_A &= w^*_A(p^{**}_A; \beta_{rd}, \mu_A) + \left(\frac{p^{**}_B}{\varepsilon^{**}_B} - (p^{**}_B - \mu_B)\beta_{rd}\right)\frac{D^{**}_B}{D^{**}_A} \\ &= w^*_A(p^{**}_A; \beta_{rd}, \mu_A) + f_w(p^{**}_B, D^{**}_A, D^{**}_B; \beta_{rd}, \mu_B), \end{aligned} \tag{1.3}$$

where w_A^{**} is the MMB equilibrium wholesale price and $w_A^*(p_A^{**};\beta_{rd},\mu_A)$ denotes the solution under SMB at the MMB equilibrium conditions. The term $f_w(p_B^{**},D_A^{**},D_B^{**};\beta_{rd},\mu_B)$ in the second line of equation 1.3 is an increasing function in p_A, D_B and μ_A; and decreasing in p_B^{**}, D_A and β_{rd}. A similar expression can be found for w_B. The retail price elasticity of demand $\varepsilon_B(p_B) = -(\partial D_B/\partial p_B)(p_B/D_B)$ evaluated at equilibrium values is $\varepsilon_B^{**} = \varepsilon_B(p_B^{**})$; and the equilibrium level of demand is $D_A^{**} = D_A(p_A^{**})$. A similar expression can be found for retail prices[12]:

$$p_A^{**} = p_A^*(w_A^{**};\beta_{rd},\mu_A) + f_p(p_A^{**},p_B^{**},D_A^{**},D_B^{**};\beta_{rd},\mu_B), \tag{1.4}$$

where p_A^{**} is the MMB equilibrium retail price and $p_A^*(w_A^{**};\beta_{rd},\mu_A)$ denotes the price in market A under SMB at the MMB equilibrium conditions. The function $f_p(p_A^{**},p_B^{**},D_A^{**},D_B^{**};\beta_{rd},\mu_B)$ is the net change in bargaining positions that arises only under MMB, that is shaped by other region's market conditions. The arguments of $f_p(\cdot)$ affect it in the same direction as they influence $f_w(\cdot)$. Both $f_w(\cdot)$ and $f_p(\cdot)$ show up only under MMB and represent the net effect of the gains from trade for the retailer and distributor from including region B in the negotiations.

Everything else constant, if MMB is assumed to be the true negotiation process; after a consolidation between the two distributors, it is possible to observe,

$$p_A^{POST} - p_A^{PRE} = f_p(p_A^{**},p_B^{**},D_A^{**},D_B^{**};\beta_{rd},\mu_B), \tag{1.5}$$

where p_A^{POST} and p_A^{PRE} describe the prices after and before the consolidation, respectively.

The model predicts that after a consolidation of distributors, if a retailer has stores in both markets, price effects are expected in both regions, everything else constant, only under MMB. Henceforth, the necessary conditions to observe price changes after a consolidation of distributors are (i) to have a retailer with stores in multiple markets, and (ii) the negotiation protocol followed is MMB. Instead, under SSM, only the market with the change in distributors observes a price variation if β_{rd} or μ_m change and the price in the other market remains unchanged. [13]

[12] In Appendix A4 I solve for the $\{w_A, w_B\}$ that solve Equation 1.3. After rearranging and using retailer's first order conditions I get the expressions for w_A^{**} and p_A^{**}.

[13] It is realistic to think that different distributors have different bargaining power parameters β and costs. These changes not only affect directly the price through the direct negotiation process, $w_A^*(p_A;\beta,\mu_A)$, but also through the indirect effect of the change in bargaining positions terms $f_p(p_A^{**},p_B^{**},D_A,D_B;\beta_{rd},\mu_B)$ and $f_w(p_B^{**},D_A,D_B;\beta_{rd},\mu_B)$. To gain additional insight, a simulation can be found in Appendix A1.

1.3.2 Structural Model

A. The Demand Model

I use a random coefficient logit model as in Berry et al. (1995) to represent the demand side. Each product is defined as the combination of a retail chain and a brand. Consumers buy one of the observed products in market m at time t or selects the outside option (j=0), $j = 1, \ldots, J_{mt}$. I assume can switch stores within the same region. Products are aggregated up to the brand level, such that each retailer sells three products; Monster, Red Bull and Rockstar in market m at time t. The conditional indirect utility that consumer i receives from purchasing good j in region m at month t is:

$$u_{ijmt} = \delta_{b(j)} + \delta_{r(j)} + \alpha_i p_{jmt} + \xi_{jmt} + \varepsilon_{ijmt}, \tag{1.6}$$

where p_{jmt} is the retail price, ε_{ijt} is distributed according to a Type I extreme value distribution, $\delta_{b(j)}$ and $\delta_{r(j)}$ are product j's brand and retailer fixed effects, and ξ_{jmt} is the unobserved demand shock. The indirect utility of the outside option is ε_{i0rt}. I assume that the consumer part of the utility is $[\Sigma\nu_{imt} + \Pi\mathcal{D}_{imt}] * [1, p_{jmt}]'$, where $\mathcal{D}$ contains the age of household's head and the log of income. The matrix Π measures how agent tastes vary with these demographic characteristics. Regarding the unobserved heterogeneity, I let ν_{imt} be independent draws from a standard normal distribution. These draws are scaled by the lower triangular matrix Σ, which denotes the Cholesky root of the covariance matrix.

Regarding the demographics, as income increases, people are expected to consume fewer energy drinks. At the same time, the age of the household's head is also informative about energy drinks consumption. The high content of caffeine, sugar, and taurine in energy drinks leads to a decrease in the consumption of energy drinks as age increases. As explained in detail in Section 1.2.2, demographics are obtained from the NielsenIQ Panel data and are drawn at the region month level using the expansion factors provided by NielsenIQ.

The instruments used shift supply but not demand. I do not use the standard BLP instruments based on product characteristics since there is not much variation in the observed characteristics of the products. Instead, I use three sets of instruments. The first set of instruments I employ is related to idiosyncratic events (Miller and Weinberg, 2017). I consider the number of competitors a particular retail chain faces in the markets affected by the consolidation. The second set of instruments I use are cost shifters (Nevo, 2001). I use the price of fuel interacted with the average reduction in driving distances for each retail chain between the nearest production facility and the center of the region. The advantage of using this instrument is that there is variation after the consolidation. Prices of sugar, coffee, and aluminum are also used, since these are the main components in the production of energy drinks. The drawback is that input prices may vary in time and not by region.

The third group of instruments is composed by the interactions of the idiosyncratic events instruments with moments of the distribution of demographic variables. As in Backus et al. (2021), I use the 10%, 50%, and 90% quantiles of the income distribution and age of the head of the household.

I estimate the parameters of the demand model as in Berry et al. (1995).

B. Supply Side Model : Multi and Single Market Bargaining

As previously stated, each of the J products in the market is a unique combination between brands and retailers. As in section 1.3.1, the negotiation process follows a Nash-bargaining protocol. I keep the assumption of simultaneous determination of wholesale and retail prices. Sheu and Taragin (2021) highlight that simultaneous downstream pricing can be a suitable assumption when the upstream firm does not have a first-mover advantage in pricing. Since there is no evidence about distributors having a first mover advantage, I maintain this assumption.

There are D distributors and R retailers who negotiate for the wholesale price of good j in market m at time t, w_{jmt}. Distributors and retailers send representatives to bargain simultaneously and separately in a Nash-in-Nash fashion. Parts in the negotiation believe that under a disagreement in the negotiation, all the other negotiated wholesale prices remain unchanged. Consumers can always substitute one product for another in the same store or switch to another retailer in the same market.

The profit in market m at time t of retailer $r \in R$ is written as follows:

$$\pi_{rmt}(\mathcal{J}_{mt}, \boldsymbol{p_{mt}}, \boldsymbol{w_{mt}}) = \sum_{j \in \mathcal{J}_{rmt}} (p_{jmt} - c_{jmt} - w_{jmt}) L_{mt} s_{jmt}(\boldsymbol{p_{mt}}; \boldsymbol{\theta}^l), \tag{1.7}$$

where L_{mt} is the potential size of market m at time t and p_{jmt} and c_{jmt} are the retail price and retailer's cost for good j, respectively. The set of products sold by distributor d and retailer r in market m at time t is $\mathcal{J}_{dmt}$ and $\mathcal{J}_{rmt}$. On the other hand, the profit of distributor $d \in D$ in market m at time t:[14]

$$\pi_{dmt}(\mathcal{J}_{mt}, \boldsymbol{p_{mt}}, \boldsymbol{w_{mt}}) = \sum_{j \in \mathcal{J}_{dmt}} (w_{jmt} - \mu_{jmt}) L_{mt} s_{jmt}(\boldsymbol{p_{mt}}; \boldsymbol{\theta}^l), \tag{1.8}$$

and μ_{jmt} and production cost for good j. The set of products sold by distributor d in market m at time t are, respectively, $\mathcal{J}_{dmt}$ and $\mathcal{J}_{rmt}$. The bargaining power parameter, β_{rd}, denotes

[14]See Appendix A at the end of the paper for the complete set of steps taken to obtain the potential market size.

the bargaining weight of retailer r relative to distributor d when negotiating and is between 0 and 1. The closer this parameter gets to one of its limits it means one of the parts is making a take it or leave offer. Setting a value of 0.5 defines a symmetric Nash Bargaining. I assume this parameter remains constant in time.

As previously discussed, it is possible to assume that distributors and retailers bargain for good j for several regions at the same time or, on the contrary, region by region independently. Under the former case, multi-market bargaining, if negotiations fail the retailer stops getting supplies j in all the locations where the distributor distributes the goods.[15] If single market bargaining is assumed, under a negotiation breakdown the retailer stops getting the supplies in one location only, and they continue negotiating for the wholesale price in other regions.[16] Next, I illustrate both negotiation protocols.[17]

Single Market Bargaining (SMB): Under SMB, failing to reach an agreement in one market does not affect the negotiation in other markets. For the wholesale price of product j, firms solve following Nash Product for each market m independently,

$$w_{jmt} = \arg\max_{w} [\pi_{rt}(w_{jmt}, \boldsymbol{w}_{-jmt}) - \pi_{rt}(\infty; \boldsymbol{w}_{-jmt})]^{\beta_{rb}} \times [\pi_{bt}(w_{jmt}, \boldsymbol{w}_{-jmt}) - \pi_{bt}(\infty; \boldsymbol{w}_{-jmt})]^{1-\beta_{rb}}, \tag{1.9}$$

where $\boldsymbol{w}_{-jmt}$ represents the vector of wholesale prices of all the other products but j in market m at time t. The solution to this Nash Product, coming from single-market negotiations can be expressed as:

$$w^*_{jmt}s_{jmt} = \underbrace{\mu_{jmt}s_{jmt}}_{\text{Production Costs}} + \underbrace{\sum_{g \in \mathcal{J}_{bmt} \setminus j} \Gamma_{gmt}\Delta_j s_{gmt}}_{\text{Disagreement payoff of distributor}} + (1-\beta_{rb})\big[GFT^R_t(m) + GFT^B_t(m)\big], \tag{1.10}$$

where $GFT^R(m)$ and $GFT^B(m)$ are the gains from trade for reaching an agreement in market m for the retailer and the distributor, respectively; and $\Gamma_{jmt} = w_{jmt} - \mu_{jmt}$ is the dis-

[15] Under this setting, firms send representatives to negotiate one contract for the wholesale price for many regions.

[16] Under this setting, firms involved in many contracts treat them separately by simultaneously sending different representatives to each negotiation. Once the bargaining process starts, the representatives do not communicate with each other, even if they belong to the same firm. While this assumption can be restrictive, it allows tractability in cases with limited data.

[17] The intermediate case about bargaining for a subset of regions is not studied in this paper. That is similar to disentangling the decision of bargaining by TCCC from the one of its franchisees. As previously discussed, there are no significant price differences between TCCC and its franchisees' regions.

tributor's margin in market m. In equation 1.10, the wholesale price has by three components. The first is the total production cost for distributor d of the goods sold to retailer r, so changes in this factor have a direct impact on the wholesale price. The second component is the profit of distributor d. This is the 'recapture' of consumers that go to other retailers when they do not find product j at their usual retailer. The third term represents the total surplus generated by both parties, from where the distributor takes $(1-\beta_{rd})$ of this quantity. The lower the β_{rd}, the higher the surplus that the distributor captures. Finally, notice that all these expressions only depend on the market conditions of market m.

Multi-market Bargaining (MMB): Retailer r and distributor d bargain over j for the set of markets $\mathcal{M}_{rdt}$ such that $\boldsymbol{w}_{rdt} = \{w_{jmt}\}_{m \in \mathcal{M}_{rdt}}$ is a vector containing the wholesale prices for product j for all the markets for which r and d negotiate. Consider $\mathcal{J}^R_{mt}$ and $\mathcal{J}^D_{mt}$ as the sets of retailers and distributors, respectively, operating in market m at time t, such that $\mathcal{J}_{mt} = \bigcup_{r \in \mathcal{J}^R_{mt}, d \in \mathcal{J}^D_{mt}} \mathcal{J}_{rdmt}$ is the set of all the products sold in market m at time t. Denote by $\boldsymbol{W}_{-rdt} = \{w_{gmt}\}_{g \in \mathcal{J}_{mt} \setminus \{j\} \text{ and } m \in \mathcal{M}_{rdt}}$ the set of all products except j in markets for which r and d negotiate. Finally, if negotiations for good j breakdown, this gets out of the market and the demand for any other goods k in market m at time t changes by $\Delta_j s_{kmt} = s_{kmt}(\mathcal{J}_{mt}) - s_{kmt}(\mathcal{J}_{mt} \setminus j) > 0$. This happens in every market where j is being sold.[18]

When bargaining, the negotiated wholesale price maximizes the following Nash product,

$$\begin{aligned} \boldsymbol{w}_{rdt} = \arg\max_{\boldsymbol{w}} [\Pi_{rt}(\boldsymbol{w}_{rdt}, \boldsymbol{W}_{-rdt}) - \Pi_{rt}(\infty; \boldsymbol{W}_{-rdt})]^{\beta_{rd}} \\ \times [\Pi_{dt}(\boldsymbol{w}_{rd}, \boldsymbol{W}_{-rdt}) - \Pi_{dt}(\infty; \boldsymbol{W}_{-rd})]^{1-\beta_{rd}}, \end{aligned} \tag{1.11}$$

where $\Pi_{rt}(\boldsymbol{w}_{rdt}, \boldsymbol{W}_{-rdt}) = \sum_m \pi_{rmt}(\boldsymbol{p}_{mt}, \boldsymbol{w}_{mt})$ and $\Pi_{dt}(\boldsymbol{w}_{rdt}, \boldsymbol{W}_{-rdt}) = \sum_m \pi_{bmt}(\boldsymbol{p}_{mt}, \boldsymbol{w}_{mt})$. The profits under a disagreement are $\Pi_{rt}(\infty; \boldsymbol{W}_{-rdt}) = \sum_{\substack{k \in \mathcal{J}_{rmt} \setminus j, \\ m \in \mathcal{M}_{rdt}}} (p_{jmt} - c_{jmt} - w_{jmt}) L_{mt} \Delta_j s_{kmt}$ and $\Pi_{dt}(\infty; \boldsymbol{W}_{-rdt}) = \sum_{\substack{g \in \mathcal{J}_{dmt} \setminus j, \\ m \in \mathcal{M}_{rdt}}} (w_{jmt} - \mu_{jm}) L_{mt} \Delta_j s_{gmt}$, for the retailer and the distributor, respectively. Each term in brackets represents the gains from trade (GFT) of reaching an agreement; the first one being retailer's, whereas the second one to those of the distributor. The higher the GFT for a firm, the higher the reliance on the other firm and the lower his bargaining power. Additionally, since negotiations are carried for all the markets at the same time, the GFT is a function of the number of regions included in the negotiation. Since negotiations are simultaneous and independent, the vector of wholesale prices of all other products $\boldsymbol{W}_{-rdt}$ does not change in the event of a disagreement between r and d. The solution to equation 1.11 for

[18] Where this last term is expressed as:

$$\Delta_j s_{kmt} = \int \frac{\exp(\delta_{kmt} + \mu_{ikmt})}{1 + \sum_{l \in \mathcal{J}_{mt}} \exp(\delta_{lmt} + \mu_{ilmt})} - \frac{\exp(\delta_{kmt} + \mu_{ikmt})}{1 + \sum_{l \in \mathcal{J}_{mt} \setminus \{j\}} \exp(\delta_{lmt} + \mu_{ilmt})} dF(\mu)$$

market m can be expressed as

$$w^{**}_{jmt} = w^{*}_{jmt} + g(s_{jmt}, \mathbf{s}_{-mt}, L_{mt}, \mathbf{L}_{-mt}, \mathbf{p}_{-mt}; \mathbf{c}_{-mt}, \mu_{-mt}, \beta_{rd}), \tag{1.12}$$

where $\boldsymbol{L}_{-mt} = \{L_{nt}\}_{n\in\mathcal{M}_{rdt}\setminus m}$ represents the set of potential market sizes for all markets except m where retailer r and distributor d negotiate. Similarly, $\boldsymbol{s}_{-mt} = \{s_{nt}\}_{n\in\mathcal{M}_{rdt}\setminus m}$, $\boldsymbol{p}_{-mt} = \{p_{nt}\}_{n\in\mathcal{M}_{rdt}\setminus m}$, $\boldsymbol{c}_{-mt} = \{c_{nt}\}_{n\in\mathcal{M}_{rdt}\setminus m}$, and $\boldsymbol{\mu}_{-mt} = \{\mu_{nt}\}_{n\in\mathcal{M}_{rdt}\setminus m}$ are the sets of market shares, retail prices, wholesale prices, retail costs, and distributor's cost for all goods in the markets where retailer r and distributor d negotiate, except m.[19]

The main difference between Equations 1.10 and 1.12 is that in the later wholesale prices not only depend on local market conditions, captured by w^{*}_{jmt}; but also depends on $g(\cdot)$, which captures how the inclusion of other markets in the negotiation changes the bargaining positions of the firms. Precisely, the function $g(\cdot)$ in equation 1.12 resembles function $f_w()$ in equation 1.3 in section 1.3.1, which was a special case for two markets and only one retailer.

From equation 1.7, the first order conditions for all the retailers, in matrix form, can be expressed as,

$$\boldsymbol{\gamma}_{mt} = \boldsymbol{p}_{mt} - \boldsymbol{w}_{mt} - \boldsymbol{c}_{mt} = -\sum_{r=1}^{R}(\boldsymbol{\Omega}_{rmt}\boldsymbol{S}_{p_{mt}}\boldsymbol{\Omega}_{rmt})^{-1}\boldsymbol{\Omega}_{rmt}\boldsymbol{s}_{mt},$$

where $\boldsymbol{\Omega}_{rmt}$ is a $J_{mt} \times J_{mt}$ matrix, with $\boldsymbol{\Omega}_{rmt}[jm, km] = 1$ if products j and k are sold by the same retailer in market m; and $\boldsymbol{S}_{p_{mt}}$ is the $J_{mt} \times J_{mt}$ matrix of substitution effects, with $\boldsymbol{S}_{pmt}[jm, km] = \partial s_{jmt}/\partial p_{kmt}$ and $\boldsymbol{s}_{mt}$ is the $J_{mt} \times 1$ vector of market shares. The vector of retailer's markup for product j in each market m where is sold is denoted by $\boldsymbol{\gamma}_{mt}$.

Equations 1.10 and 1.12 can be expressed in matrix form, for all the distributors at time t as

$$\boldsymbol{\Gamma}_t(\boldsymbol{\beta}, \boldsymbol{p}_t, \boldsymbol{s}_t) = \boldsymbol{w}_t - \boldsymbol{\mu}_t = \sum_{d=1}^{D}\sum_{r=1}^{R}[\boldsymbol{\Omega}_{dt}\tilde{\boldsymbol{S}}\boldsymbol{\Omega}_{dt}]^{+}[\boldsymbol{\beta} \circ (\boldsymbol{\Omega}_{rt}\tilde{\boldsymbol{S}}\boldsymbol{\Omega}_{rt}\boldsymbol{\gamma}_t)], \tag{1.13}$$

where the symbols $+$ and $\circ$ represent, respectively, the generalized Moore-Penrose inverse and the Hadamar product operator for element by element multiplication. The vector of stacked $\boldsymbol{\gamma}_{mt}$ across markets at time t is denoted by $\boldsymbol{\gamma}_t$, and $\boldsymbol{\Omega}_{rt}$ and $\boldsymbol{\Omega}_{dt}$ are $J_t \times J_t$ ownership matrix for the retailers and distributors, respectively; with element $[j, k] = 1$ if products j and k are sold by the same firm and zero otherwise.

The rest of the terms depend on whether the bargaining is at multi or single market level.

[19]The complete expression of equation 1.12 can be found in Appendix A6.

If it is the former, the matrix of market shares and changes in market shares, $\tilde{\boldsymbol{S}} = s_{jmt}L_{mt}$ if j is a product distributed in each market m by distributor d, and $\tilde{\boldsymbol{S}}[j,k] = \Delta_j s_{kmt}L_{mt}$ for good k in market $m \in \mathcal{M}_{rb}$. On the other hand, if the bargaining process takes place at the market level; $\tilde{\boldsymbol{S}}[j,j] = s_{jmt}$ and $\tilde{\boldsymbol{S}}[j,k] = \Delta_j s_{kmt}$ otherwise. The rest of the elements are defined as before, they just accommodate to the type of bargaining taking all the markets by time t or only at the market level.

Assuming MMB is the right model accounts for studying the price effects when $\tilde{\mathcal{M}}_{rdt} = \mathcal{M}_{rd,t<t^*} \bigcup \mathcal{M}_{rd,t>t^*}$, where t^* is the date of the consolidation. To obtain those effects it is necessary to get the prices if $\mathcal{M}_{rdt}$ had remained constant over time. With that aim, in the next section I first introduce a reduced form model to assess whether MMB is the right model. Then, I estimate both demand and supply models to later calculate the counterfactual outcomes.

1.4 Estimation and Results

In this section I test the predictions obtained in the theoretical model by studying the consolidation of distributors in the US energy drinks market. Since wholesale prices are not observable, I rely on changes in retail prices as an indicator of the effects of the consolidation, as shown in equation 1.5. I take advantage of AB and TCCC's non-overlapping distribution territories and use this the regional shift in distributors in the estimation procedure. Although all the stores in affected areas were influenced, stores outside this area that belonged to affected retail chains were also potentially affected. If they do, this is a clear sign that SMB is not the mechanism generating the data.

The first step involves employing a reduced-form approach to assess whether the observed price fluctuations align with the predictions of either the SMB or MMB models. Under MMB, the theoretical model predicts price effects in both affected and not affected regions by the consolidation. Then, I apply the structural model developed in the previous section to assess the origins of the observed price variation.

1.4.1 Reduced form Evidence

As pointed out at the end of section 1.3.1, the consolidation only generates price effects if (i) prices are negotiated through an MMB protocol, and (ii) retailers had stores in both AB and TCCC distribution territories. This last group of retailers is called 'national retailers', because of their presence in both regions. Retailers with stores in only the AB territories are called 'regional retailers'. Regional retailers with stores only in TCCC territories did not change distributors after March 2015. Instead, national retailers were getting supplies from both AB and TCCC, each for a different territory. So, national retailers have stores in regions affected and not affected by consolidation. Taking this into account, in Table 1.2 I classify the stores according to which potential bargaining protocol affected their prices.

Table 1.2: Classification of Stores

Group	Change of Distributor	Chain affected by Consolidation	Store affected by Consolidation	Possible Price Effects
Group 1	Yes	Yes	Directly	Under SMB or MMB
Group 2	No	Yes	Indirectly	Under MMB
Group 3	No	No	Not	In equilibrium and MMB

The first group of stores are those located in the regions affected by the consolidation, so their prices are likely to change regardless the bargaining protocol used by the firms. Following Equation 1.4, the change in prices could have come from changes in the bargaining power parameter of the firms or difference in distributors' production costs. The second group of stores is composed by those that are not in the regions affected by the consolidation, but belong to a retail chain that was. Price effects arise for this group only if MMB is the true model; i.e., through $f_p(\cdot)$. Finally, the third group of stores belongs to regional retailers exclusive to TCCC territories, adjusting prices only in response to Group 2 price changes.

To assess whether the consolidation generated effects, I test if national retailers changed their prices in the regions not affected by the consolidation. As discussed before, this cannot happen under SMB. I use the third group of stores as comparison group, given that they only changed their prices after the consolidation of distributors as response to Group 2. As a consequence, comparisons between Group 1 or Group 2 against Group 3 are a difference between the effects of the consolidation and the effects of adjustment of prices in equilibrium. I estimate separately for each $k \in \{\text{Monster, Red Bull, Coca-Cola}\}$ the logarithm of the price of product j in store s at time t:

$$\begin{aligned}\log(price_{jst}) = \alpha_{1k}\mathbb{1}\{CONS\}_{js} \times \mathbb{1}\{t > t^*\} + &\alpha_{2k}\Delta MILES_s \times \mathbb{1}\{t > t^*\}_t \\ &+ \zeta_{jt} + \zeta_{js} + \boldsymbol{x}_{st}\boldsymbol{\delta} + \varepsilon_{jrt},\end{aligned} \tag{1.14}$$

where $\mathbb{1}\{CONS\}_{js}$ takes the value of one for national retailers whether the store is affected or not by the consolidation and zero otherwise, $\mathbb{1}\{t > t^*\}$ takes the value of 1 after the change in distributors by Monster in $t^* = 2015$; and, ζ_{jt} and ζ_{js} are to product time and product store fixed effects. Store's county variables weather, population, and median income are represented by $\boldsymbol{x}_{st}$. $\Delta MILES$ measures the variation in the number of miles (in thousands of miles) of driving distances from the center of the county to the nearest production facility.[20] The interaction $\Delta MILES \times \mathbb{1}\{t > t^*\}$ is an approximation of the possible cost efficiencies after the consolidation of distributors. This variable can serve as a conservative estimate for the efficiencies resulting from the consolidation of distributors, considering that additional efficiencies may emerge post-consolidation. The overall price effects are captured by α_{1k}, which represents the difference between consolidation effects and equilibrium effects.

For national retailers' stores from the not affected regions, Group 2, the price effects are:

$$\log(p_{jst}^{POST}) - \log(p_{jst}^{PRE}) = \mathbb{1}\{CONS\}_{js} \times \mathbb{1}\{t > t^*\} = \begin{cases} f_p(\cdot), & \text{under MMB} \\ 0, & \text{otherwise,} \end{cases} \tag{1.15}$$

which aims to capture variations in retail prices based on the type of bargaining protocol. If the single-market bargaining is the used bargaining protocol, this difference should be not significant. However, under MMB the difference should capture the change in firms' bargaining positions, or $f_p(p_m, p_{-m}, D_m, D_{-m}; \beta, c_{-m})$. On the other hand, for the group of directly affected stores, Group 1, $\mathbb{1}\{CONS\}_{js} \times \mathbb{1}\{t > t^*\}$ could emerge from either MMB through $f_p(\cdot)$ or from SMB through changes in β and c.

Table 1.3: Effects of Consolidation on retail prices

	$\log(price)$			
	Monster		Red Bull	
	Affected (i)	Not Affected (ii)	Affected (iii)	Not Affected (iv)
$\mathbb{1}\{CONS\} \times \mathbb{1}\{t > t^*\}$	-0.015**	-0.016***	0.004	0.003
	(0.006)	(0.004)	(0.003)	(0.003)
$\Delta MILES \times \mathbb{1}\{t > t^*\}$	0.010	-	-0.004	-
	(0.008)	-	(0.005)	-
Observations	863,567	993,524	1,351,315	1,572,325
R^2	0.983	0.983	0.986	0.984
Prod-Store, Prod-Region FE	Yes	Yes	Yes	Yes
Controls, trend	Yes	Yes	Yes	Yes

Standard errors in parentheses, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Results are shown in Table 1.3. Column (i) shows a significant price effect of -1.5% for Monster products for stores in Group 1. A close price effect, -1.6%, is depicted in Column (ii). This means that in average, for stores in Group 2, the average price of Monster products went down by 1.6%, compared to the stores in Group 3. From Table 1.2, and as discussed before, significant changes in Group 2's prices only arise under multi-market bargaining and not under

[20] Similar results are obtained when using driving time instead.

single-market bargaining. The last result suggests that in the retail sector the negotiation for products is not market by market, but for all the markets where retailers are present, at the same time.

Notice also that the price decrease in national retailers' stores in regions affected and not affected by the consolidation are not significantly different. This goes in line with the theory of uniform pricing, that claims that, everything else constant, the price variation retail chains are homogeneous after a shock. Also, in absolute value these results are close to the one obtained by Luco and Marshall (2020) (1.5%) from a price increase by TCCC distributors when selling a competitor's product. So, these results are in line with previous price variations in the literature. Regarding Monster's main competitor, columns (iii) and (iv) show the effects for Red Bull products. Since they are under equilibrium effects in the all the stores, there are no expected effects for this product. Finally, it seems that the cost efficiencies captured by $\Delta MILES \times \mathbb{1}\{Post\}_t$ do not significantly affect prices. They do not exhibit a significant price effect after the change in distributors. Overall the results show that the distributors' consolidation in the energy drinks market decreased prices. An opposite effect is shown for the quantities. The increase in quantities sold is shown in the next table.

Table 1.4: Changes in retail quantities

	Monster		Red Bull	
	Directly	Indirectly	Directly	Indirectly
$\mathbb{1}\{CONS\} \times \mathbb{1}\{t > t^*\}$	0.210***	0.103***	-0.003	-0.026
	(0.023)	(0.019)	(0.024)	(0.022)
$\Delta MILES \times \mathbb{1}\{t > t^*\}$	-0.018	-	-0.018	-
	(0.049)	-	(0.029)	-
Observations	863,567	993,524	1,351,315	1,572,325
R^2	0.855	0.853	0.856	0.846
Prod-Store, Prod-Region FE	Yes	Yes	Yes	Yes
Controls, trend	Yes	Yes	Yes	Yes

Standard errors in parentheses, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

So far I have assumed that the consolidation only affects prices through a change in the bargaining positions of the firms, which are incorporated into prices through the multi-market bargaining. However, it is important to consider other possible sources for the observed price decrease. First, Butters et al. (2022) show that local cost shocks do not lead to price effects in other regions. As a result, price changes in national retailers' stores in not affected territories cannot be attributed to cross-subsidization of local costs shocks among stores. Second, another possibility is that distributors and retailers re-negotiate contracts nationally based on the new scale of their agreement. While this might be a reasonable explanation for the price changes, this mechanism is implicitly included in the multi-market bargaining mechanism previously described. On the contrary, single-market bargaining will not be able to include the scale of the amount traded, since the bargaining is taken for each region independently and separately.

Third, regarding costs and scale, the regions that changed distributors from AB to TCCC experienced a change in production costs. Results in Table 1.4 show that variations in driving distances did not translate in lower prices. TCCC's production facilities are spread across the US. The decrease in the number of distributors for Monster was not translated into an overload of production for the existing production facilities. Instead, the production for the new territories was carried by the TCCC facilities in those territories. It is possible that some production facilities in the regions not affected by the consolidation increased the production to supply the regions affected by the consolidation, but this variation is expected to be marginal. [21]

The previous results show that a regional expansion of a distributor can have price effects and this is only possible when wholesale prices are negotiated for all the regions at the same time. Nonetheless, since the parameter α_{1k} captures both the consolidation and equilibrium effects, it is not possible to claim that the observed price variation is coming exclusively from the consolidation of distributors. This motivates the additional structure that I put into the model. In the next section, I estimate the structural model that can disentangle the different effects from the consolidation.

1.4.2 Demand

Table 1.5 presents the demand estimation results. The price coefficient is negative and significant for the logit and the RCL specifications. Since there is not much variation in product characteristics like t levels of caffeine or sugar intake, I do not include them in the regression. Instead, brand fixed effects are considered.

I also include demographics interacted with the price and the constant. Households with higher income are less price sensitive, but overall have a lower preference for energy drinks. On the other hand, older individuals are more price sensitive.

Although not significant, σ_1 and σ_{p_t} represent the diagonal terms of square root of the covariance matrix for the unobserved taste heterogeneity for prices.

1.4.3 Supply

Equation 1.13 depicts the margins of the distributors as a function of both the observed data and the unobserved bargaining power parameters and the retailer's margin (Draganska et al.,

[21] In the Appendix A3 I discuss the possible causal interpretation of the results.

Table 1.5: Changes in retail prices

Variable	Logit	RCL
Price	-3.7844^{***}	-3.5930^{***}
	(0.9263)	(0.7359)
Income x Constant		0.2316
		(0.1415)
Age x Constant		0.09530^{***}
		(0.0090)
Income x Price		-0.0213
		(0.0456)
Age x Price		0.00068
		(0.0040)
σ_1		0.0012
		(0.0294)
σ_{p_t}		0.0106
		(0.0595)
Time fixed effects	Yes	Yes
Retailer fixed effects	Yes	Yes
Brand fixed effects	Yes	Yes

Standard errors in parentheses

2010). Retailers' marginal costs, $w_t + c_t$, can be expressed as the sum of $w_t - \mu_t$ and $c_t + \mu_t$, where the first term is the distributors' margins and the second represents the total cost along the vertical chain. The nature of this total cost is industry specific. For example, Grennan (2013) assumes stent costs do not vary by the downstream firm (hospital) or in time, so there are no unobserved cost shocks.

Following other papers in the literature (Draganska et al., 2010; Gowrisankaran et al., 2015), I assume that total costs along the vertical chain, $c_t + \mu_t$, can be modeled as a function of cost shifters, $\boldsymbol{\eta}_t$, and an unobserved cost shock ω_t, such that $c_t + \mu_t = \boldsymbol{\eta}_t \boldsymbol{\kappa} + \omega_t$. I include in $\boldsymbol{\eta}_t$ distributor and retailer fixed effects, prices of sugar and caffeine interacted with the amount of sugar and caffeine, respectively, by brand; the price of aluminum interacted with the number of cans of 16oz; and, an index for the price of gasoline interacted with the change in miles from the production facilities to the center of the region. The structural error term is expressed as:

$$\boldsymbol{\omega}_t(\boldsymbol{\beta}, \boldsymbol{\kappa}) = \boldsymbol{p}_t - \gamma_t(\boldsymbol{\beta}, \boldsymbol{p}_t, \boldsymbol{s}_t) - \boldsymbol{\Gamma}_t(\boldsymbol{\beta}, \boldsymbol{p}_t, \boldsymbol{s}_t) - \boldsymbol{\eta}_t \boldsymbol{\kappa} \tag{1.16}$$

Where $[\boldsymbol{\beta}, \boldsymbol{\kappa}]'$ is the vector of nonlinear parameters to estimate. Given the absence of wholesale data, identification of $\boldsymbol{\beta}$ is based on downstream behavior. Since retailers and distributors' margins are based on retail data, I cannot separately identify the bargaining power from retailers' conduct. Retail prices could be high because of collusive retailers with low bargaining power or because of competitive retailers with high bargaining power. Since there is no evidence of collusive behavior in this industry, I assume that retail data comes from a competitive market.

The markup terms $\boldsymbol{\gamma}_t(\boldsymbol{\beta}, \boldsymbol{p}_t, \boldsymbol{s}_t)$ and $\boldsymbol{\Gamma}_t(\boldsymbol{\beta}, \boldsymbol{p}_t, \boldsymbol{s}_t)$ are endogenous because the unobserved cost shock, $\boldsymbol{\omega}_t$, enters implicitly through price. To tackle this issue, I use a GMM estimator based on the moment condition $\mathbb{E}[\boldsymbol{\omega}_t(\boldsymbol{\beta}, \boldsymbol{\kappa})|\boldsymbol{Z}_t] = 0$. The matrix of instruments $\boldsymbol{Z}$ includes a dummy for Monster products in the regions where there was a change in distributor after 2015. The power of the instrument comes from the reduced form evidence shown in the previous section, where a reduction in the retail price of Monster happened in the affected regions after the change in distributor. The validity of the instrument is based on its orthogonality to the unobserved cost shock. If the unobserved cost shock of producing Monster is not systematically different from those of Red Bull and Rockstar before and after the change in distributors, it is likely that the orthogonality condition holds. Product and time fixed effects should be able to capture the difference in levels between the different brands.

I also use as instruments variables that affect demand but not costs. Temporary feature and display of the products in the store make a good work seizing this effect. They affect both the demand for the featured or displayed product, and the other products, they do not affect production costs. This is a well-founded exclusion restriction when retail prices are set by retailers and display and feature are chosen by the distributor. This guarantees the price variation is not coming from a change in the retailer's costs, which could also affect the prices of competing products. [22]Anecdotal evidence suggests that distributors choose sales periods as a part of their negotiations with retailers. I also use the number of flavors of the competitors' brands in the same retail chain as another instrument. This is similar to the typical BLP instrument and follows the same relevance and exclusion restrictions. Finally, I include the time and region fixed effects.

Since $\boldsymbol{\kappa}$ is a function of $\boldsymbol{\beta}$, I concentrate it out before the minimizing the following GMM objective function,

$$\widehat{\boldsymbol{\beta}} = \arg\min_{\beta}(\boldsymbol{Z}\boldsymbol{\omega}(\boldsymbol{\beta}))'\boldsymbol{A}^{-1}\boldsymbol{Z}\boldsymbol{\omega}(\boldsymbol{\beta}), \tag{1.17}$$

where $\boldsymbol{A}$ is a weighting matrix and $\boldsymbol{Z}$ is a matrix of included and excluded instruments.[23]

[22]This is not going to be the case if retailers set feature and display because of spacing issues or close expiration dates of the products.

[23]In the first step this matrix is the initial one $\boldsymbol{A} = \boldsymbol{Z}'\boldsymbol{Z}$. Then, I compute the optimal weighting matrix using the residuals from the first stage.

Regarding the estimation of β, I choose a grid of initial values for the vector of bargaining power parameters. Since having bargaining power parameters outside the 0-1 interval does not have economic meaning, the search for possible solutions is constrained to the previously mentioned limits. With the results from the first step I get the residuals for the second step of the estimation, where I use the same grid of initial values as in the first stage. I perform the previous two steps for different set of initial values, getting estimates that are close to each other each time. The fact that the results are numerically close to each other for all the range of possible initial values indicates that the observed results might be a global optimum rather than a local solution. Nonetheless, to guarantee these results are in fact global solutions, I also solve the GMM objective function by using the global optimizer dual annealing. The results using the global optimizer not only lead to results qualitatively identical to those found under the local optimizers, but also quantitatively closed. Although it is computationally more demanding than using a local optimizer, a dual annealing does not require initial values; just boundaries for the values of the parameters.

Before turning to the estimation results, recall that distributors keep other products in their portfolio. I assume that the importance of distributors' portfolio is partially captured by the bargaining power parameters and not by the number of products negotiated. As shown by Crawford and Yurukoglu (2012), negotiating for a bundle of products can affect the bargaining outputs. However, since this paper aims to account for the price effects of changes in market structure, I do not include other non-energy drinks products. Instead, I rely on the low substitution patterns between energy drinks and traditional soft drinks to justify the absence of additional price effects from working with just the category of energy drinks.[24]

Table 1.6: Supply Side results - Multi Market Bargaining

	Joint model		Cost model
β_{TCCC}	1.0000	Aluminium	-0.0363
	(0.0353)		(0.0083)
$\beta_{RedBull}$	0.8441	Coffee	-1.9498
	(0.1035)		(1.8259)
β_{Pepsi}	1.0000	Sugar	0.0097
	(0.1026)		(0.0131)
β_{AB}	0.3938	$\Delta MILES \times P_{gas}$	-0.0043
	(0.0297)		(0.0087)
Obs			66344
Retailer Fixed Effects			Yes
Bottler Fixed Effects			Yes

Standard errors shown in parentheses.

[24] In Crawford and Yurukoglu (2012) negotiating for bundles affect the bargaining output to the extent that consumers demand some products negotiated as bundles. In this paper, consumers just get one energy drink and do not purchase other products together.

Results: Following Gowrisankaran et al. (2015), I parametrize the bargaining power parameters for the estimation by considering distributor-specific bargaining weights, instead of distributor-retailer specific. The reason lies in the need for using at least one instrument for each parameter. I employ the four instruments described in the previous section to estimate the four bargaining power parameters. With the results from the demand side, I recover retailer and distributor margins that solve the Nash-in-Nash bargaining process.

Recall that Rockstar only had Pepsi as unique national distributor and Red Bull had his own network of distributors. Instead, Monster had TCCC and AB as distributors before the consolidation and only TCCC after. Results are shown in Table 1.6.[25, 26]

Note that $\beta_b \approx 1.0$ for $b = \{\text{TCCC}, \text{Pepsi}\}$. Retailers' bargaining power when negotiating with the upstream distributor is a take-it-or-leave-it offer. The reason retailers' bargaining power is large might come from the large size they represent in the US or their importance due to their multi-product nature. On the other side, the fact that $\beta_{RedBull} = 0.84$ makes more likely to thinking about a powerful retailer bargaining for wholesale prices for the supply of one product. Recall that Red Bull is integrated with its distributors, and they only distribute Red Bull products, unlike their main competitors Monster and Rockstar. For the regions where there was a change in distributors, retailers passed from having a bargaining power of $\beta_{AB} = 0.39$ to one equal to $\beta_{TCCC} \approx 1.0$.

Finally, Table 1.7 shows the estimates under single market bargaining. The estimation procedure is the same that under multi-market bargaining. Although these results for β are different across bargaining models, the ranking of values is close in both estimations with the parameter for Red Bull being the exception. Under both models, however, retailers increased their bargaining power parameter vis-à-vis their distributor. The increase in retailers' bargaining power allowed them to get better deals, partially explaining the observed reduction in retail prices depicted in Table 1.3.

The reduced form evidence allowed to conclude that the best model generating the data was the MMB model. In this section I estimated both models with the aim of testing which one fits the data in the best way. In the next section I show the procedure and result to do this.

[25]The details in the computation of the outside option can be found in Appendix A.1.

[26]To ensure finding a global optimum, I use the Dual Annealing algorithm to solve the GMM objective function. Dual Annealing is an optimization algorithm that combines the principles of simulated annealing and local search to efficiently search for the global optimum in complex optimization problems. By iteratively exploring the solution space and adapting the exploration rate, it can effectively navigate through potential local optima and converge to the best solution. This makes Dual Annealing a valuable tool for estimating models and finding optimal parameter values, especially in scenarios where traditional optimization methods may struggle to find the global optimum.

Table 1.7: Supply Side results - Single Market Bargaining

	Joint model		Cost model
β_{TCCC}	1.0000	Aluminium	-0.0335
	(0.2279)		(0.2279)
$\beta_{RedBull}$	1.0000	Coffee	-0.4489
	(0.2801)		(0.2801)
β_{Pepsi}	0.3920	Sugar	0.0106
	(0.2716)		(0.2716)
β_{AB}	0.1108	$\Delta MILES \times P_{gas}$	-0.0094
	(0.3339)		(0.3339)
Obs			66344
Retailer Fixed Effects			Yes
Bottler Fixed Effects			Yes

Standard errors shown in parentheses. Bootstrapped standard errors reported with data resampled at the month-market level. In this version I am using 57 samples.

1.4.4 Model Fitness

Using the supply side estimates from the SMB and MMB model, I test which model fits best the data. To do this, I follow the methodology developed by Rivers and Vuong (2002), where they compare among different models to assess which one satisfies best the moment restrictions. The benefit of using the Rivers-Vuong test is that it does not require any of the candidate models to be the true one, unlike the Cox test alternative. In that sense, when comparing models the test only tells which is one is preferred, but not which one is the true one.

Bonnet and Dubois (2010) started using the Rivers-Vuong test to compare different supply specifications. Recently, other papers like Backus et al. (2021) started using the Rivers-Vuong test for conduct testing. Starc and Wollmann (2022) used the Rivers-Vuong test to compare models of competition against collusion in the generic drug manufacturing market. Duarte et al. (2023) have shown that the Rivers-Vuong approach exhibits a superior performance compared to the model assessment alternatives like the Cox or Anderson-Rubin test. Following these previous papers, I use the non-nested algorithm employed in Backus et al. (2021) where for each candidate model h, with $C^h_{jmt} = \mu^h_{jmt} + c^h_{jmt}$, the following regression is applied,

$$C^h_{jmt} = g_V(V'_{jmt}; \gamma^h_j, \lambda^h) + \eta^h_{jmt} \quad \text{for } h \in \{\text{MMB, SMB}\},$$

where η^h_{jmt} is the error of a regression of marginal costs on the function $g_V(V'_{jmt}; \gamma^h_j, \lambda^h)$, V'_{jmt} is a matrix containing the prices of aluminum, coffee, sugar, and the price of gas times the change in distances from the center of the region to the plan location, γ^h_j is a product specific parameter specific to each model h, and λ^h is the vector of costs coefficients. Note that

$g_V(V'_{jmt};\gamma_j^h,\lambda^h)$ can be either linear or non-linear. Specifically, for the nonlinear case, I employ a random forest algorithm to perform a non-linear regression.

I follow the steps described on Backus et al. (2021). First, using the results of both models I get $\Delta\Gamma_{jt} = \Gamma^1_{jt} - \Gamma^2_{jt}$. Then, I regress $\Delta\Gamma_{jt}$ on $q_I(\mathbf{z}_t)$, such that $\Delta\hat{\Gamma}_{jt} = \hat{q}_I(\mathbf{z}_t)$. Third, regress C^h_{jmt} on $g_V(V'_{jmt};\gamma_j^h,\lambda^h)$ to get $\hat{\eta}^h_{jmt}$. Fourth, I get the values of $\hat{Q}(\Gamma^h)$, where $\hat{Q}(\Gamma^h) = (n^{-1}\sum_{j,t}\hat{\eta}^h_{j,t}\cdot\hat{g}_I(\mathbf{z}_t))^2$for each model $h \in \{MMB, SMB\}$. Finally, I construct the test static $T = \sqrt{n}(\hat{Q}(\Gamma^1) - \hat{Q}(\Gamma^2))/\hat{\sigma}$ that I get from getting the values of $\hat{Q}(\Gamma^1)$ and $\hat{Q}(\Gamma^2)$ over 500 bootstrap samples. The T statistic follows a standard normal distribution. Finally, identification is based on the assumption that $\mathbb{E}[\eta^h_{jmt}A(\mathbf{z}_t)] = 0$ and $A(\mathbf{z}_t) = \mathbb{E}[\Delta\Gamma_{jt}|z^h_{jmt}] = 0$. The vector of instruments, $\mathbf{z}_t$, includes the number of stores affected by the consolidation in the region a particular firm faces and its interactions with demographics specifics to the region. Results can be found in Table 1.8.

Table 1.8: Non-nested Model Test

Specification	T	p-value
Costs - linear specification	346.15	0.0000
Log(Costs) - linear specification	553.60	0.0000
Cost - linear quadratic	318.58	0.0000
Log(Costs) - linear quadratic	514.14	0.0000
Costs - non linear specification	-442.66	0.0000
Log(Costs) - non linear specification	103.25	0.0000

Note: The test T statistic is distributed standard normal. The standard error of the difference between $\hat{Q}_1$ and $\hat{Q}_2$ is obtained via 500 bootstrap samples.

The T test compares the results for SMB versus MMB. A positive T favors MMB over SMB. Table 1.8 shows six specifications for the model, depending on whether the costs are in levels or in logs and if the functions, $g_V(V'_{jmt};\gamma_j^h,\lambda^h)$ and $\hat{g}_I(\mathbf{z}_t)$, are linear, quadratic or non-linear. I assume $g_I(\cdot)$ follows the linearity or non-linearity assumption picked for $g_V(\cdot)$. When the costs are logarithmic, the residual η^h also takes a logarithmic form.

Table 1.8 shows the results. For almost all the specifications ,the values of T are positive, favoring multi-market bargaining over single-market bargaining. Only the non-linear specification with costs in levels rejects MMB in favor of SMB. The previous evidence suggests multi-market bargaining is the right model. The specification tests reject that SMB is the true model generating the observed data.

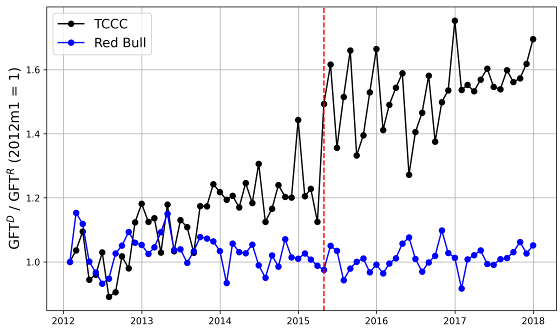
1.4.5 Evolution of the Relative Bargaining Positions

The reduced form evidence showed that national retailers decreased their prices everywhere, compared to the regional retailers that did not change distributors. The model developed in Section 1.3.1 predicts that, everything else constant, price variations after a consolidation of distributors comes from multi-market bargaining. Specifically, when both parts maximize the Nash Product, they maximize the weighted gains from trade. The gains from trade represent the relative bargaining power of one part against its trading partner. Higher gains from trade imply higher losses from not trading, increasing the bargaining power of the other part in the negotiation. After including more markets in the negotiation process as a result of the consolidation of distributors, the gains from trade changed for both retailers and distributors. In the following, using the previous result on demand and supply, I evaluate the evolution of the relative gains from trade, φ_{rdt}.

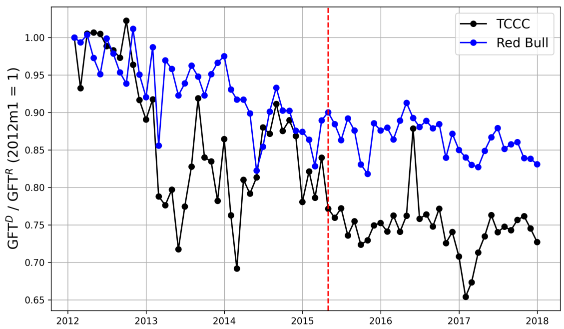
$$\varphi_{rdt} = \frac{GFT_t^d}{GFT_t^r} = \frac{\Pi_{dt}(\boldsymbol{w}_{rd}, \boldsymbol{W}_{-rdt}) - \Pi_{dt}(\infty; \boldsymbol{W}_{-rd})}{\Pi_{rt}(\boldsymbol{w}_{rdt}, \boldsymbol{W}_{-rdt}) - \Pi_{rt}(\infty; \boldsymbol{W}_{-rdt})} \tag{1.18}$$

The ratio φ_{rdt} shows the evolution in time of the gains from trade of the distributors over those of the retailers. A measure higher than one implies that the distributors have higher gains from trade compared to the retailers, and so a weaker bargaining power. The results from this ratio are shown in Figure 1.4. For comparing the results of TCCC and Red Bull, the ratios are shown using January 2012 as the base period.

Figure 1.4: Evolution of Ratio of Gains from Trade



(a) National Retailers



(b) Regional Retailers

The blue line panel (a) shows the results for Red Bull distributor against national retailers. The ratio tends to be stable during the period 2012 -2017, indicating no significant changes in the gains from trade for Red Bull. In the same panel, the black solid line shows the evolution of the relative gains from trade of TCCC with national retailers. In the period from January to

March 2015, the ratio evolves steadily with a slight positive slope. Nonetheless, the consolidation of distributors happened in April 2015, and more regions were included in the negotiation problem between TCCC and the national retailers. The ratio makes a jump from 1.13 in March 2015 to 1.49 in April 2015, an increase in 33%. This means that the gains from trade increased more for TCCC than for the national retailers.

After an increase in the number of regions included in the negotiation, it is expected that the gains from trade increase. However, the increase was higher for TCCC, compared to the national retailers, by about 33%. The increase of 33% in the relative gains from trade was translated into a weaker bargaining power for TCCC relative to the national retailers. These last ones, after the consolidation, were in a better bargaining position and hence able to negotiate for a lower wholesale price, which was passed through as lower retail prices.

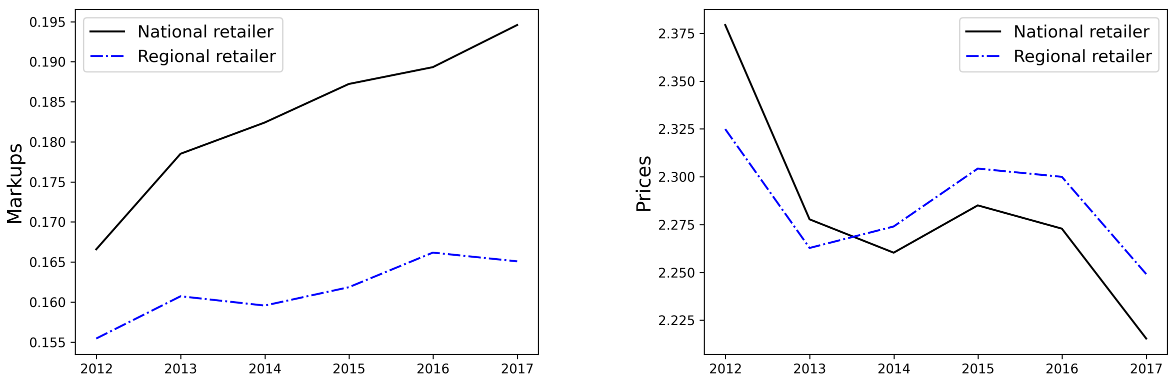
Finally, panel (b) also shows the evolution of the ratio of relative gains from trade for both distributors TCCC and Red Bull against regional retailers. The fall in the relative gains from trade for both distributors indicates that their bargaining power against smaller retailers was actually increasing. But the consolidation of distributors stopped the fall of the relative gains from trade, stopping at the same time the improvement in their bargaining position against regional retailers.

Overall, it seemed that the consolidation of regional distributors weakened the bargaining position of the distributors. Including more regions increased their dependency to trade with the retailers. While it is also true that retailers also increased their bargaining position, the relative increased of theirs was lower in comparison to the one of distributors.

1.4.6 Markups

In this section I show how the consolidation of distributors had differentiated effects according to the type of retailer, focusing on prices and markups. The evolution of markups has been studied for other industries in the US, like the automobile sector, cement industry, retail sector, etc. I start by showing the evolution of the markup and the retail price. The markup is defined as the inverse of the Lerner Index, $(p_{jmt} - mc_{jmt})/p_{jmt}$. Figure 1.5 shows the monthly average markup by type of retailer. For regional retailers, which did not change distributors, the markups depict an almost constant pattern. For the national retailers a small positive trend can be noticed. This increase becomes more pronounced after the consolidation, whereas for the regional retailers the change in the markups happens one year after the consolidation, in 2016. On the other hand, while prices were increasing from one to two years before the consolidation, they start falling around the consolidation date, with the fall being more pronounced for the national retailers.

Figure 1.5: Evolution of Markups and prices



(a) Yearly evolution of markups

(b) Yearly evolution of prices

Comparing the results for both groups, national retailers have a clear increase in their markups in the year of the consolidation. This is explained by the further decrease in marginal costs, relative to the price. As pointed out by Döpper et al. (2022), in the retail sector firms tend to keep most of the cost savings. On the contrary, the regional retailers showed a constant decrease in markups, that only stop decreasing at the end of the analyzed period.

Results from the reduced form model and the conduct testing model allow to conclude that multi-market bargaining seems to be a more reasonable data generating process than the single market bargaining protocol. Assuming that prices come from a negotiation process where distributors and retailers negotiate market by market could lead to biased results. To analyze this, in the next section I introduce counterfactual scenarios to disentangle the sources behind the observed variation in prices and compare the different outputs under different bargaining models.

1.5 Counterfactuals

Based on the estimation results, I conduct several counterfactual scenarios to uncover the individual impacts of cost efficiencies, bargaining power, and multi-market bargaining. To do so, I used the first-order condition equation $\mathbf{p}_t - \mathbf{\Gamma}_t - \gamma_t = \mathbf{C}_t$, where $\mathbf{C}_t = \mathbf{c}_t + \mu_t$ is the total cost. By varying the total costs, the producers' ownership matrix, or both, I am able to create different counterfactuals. In every scenario, the wholesale prices are the output of a Nash bargaining negotiation process as stated before. The different counterfactuals to be evaluated are summarized in Table 1.9. The baseline counterfactual where there is no change in distributor is denoted by (0). This scenario is obtained by keeping both TCCC and AB as

Monster's distributors.

Next, to isolate the effects from cost changes, I simulate a scenario where there is only a change in the production costs but not a change in distributors. In this scenario, there is no change in the retailer's bargaining power. This counterfactual is denoted by (1). Counterfactual (2) is the observed situation, where there are both costs and bargaining power changes. Then, I compare the average monthly consumer surplus and prices for the two previous scenarios against the benchmark counterfactual (0).

Table 1.9: Counterfactuals

Counterfactual	Description
(0) No Changes	No change in distributor (nor in costs)
(1) Distributor only	TCCC gets the distribution but without cost savings
(2) Observed	TCCC becomes the only distributor of Monster with cost savings

Next I compute the change in prices as $\Delta p_t(x) = \sum_{t \in 2015}(p_t(x) - p_t(0))/p_t(0)$, where x represents the counterfactual.[27] Using the estimated results from Table 1.6, in Table 1.10 I show the results comparing the different scenarios with the baseline one for multi and single-market bargaining for the year 2015.

Table 1.10: Counterfactual Analyses

	No Changes	Distributor Only	Observed
	Prices - Multi Market Bargaining		
National Retailer - Affected areas	2.29	2.29	2.25
National Retailer - Not affected areas	2.24	2.30	2.26
Regional Retailer - Not affected areas	2.41	2.39	2.20
	Prices - Single Market Bargaining		
National Retailer - Affected areas	2.09	2.09	2.25
National Retailer - Not affected areas	-	-	2.2625
Regional Retailer - Not affected areas	-	-	2.2031
	Welfare Statistics		
ΔCS - Affected areas	-	0.02%	3.14%
ΔCS - Not affected areas	-	0.02%	2.82%
CCR Profit with National Retailers	-	1.81%	2.04%
CCR Profit with Regional Retailers	-	-0.17%	-0.58%

Averages shown for the year 2015.

In the first row of Table 1.10 prices are depicted for the all the scenarios for national retailers' stores in the affected areas. A scenario where there is only a change in distributors in the

[27] The algorithm I follow can be found in Appendix A.2.

market, there is a reduction of 0.07%. Finally, when comparing the observed price variation to the baseline scenario, there is a reduction of 1.3%. This last result represent a monthly average reduction of 2¢. On the other hand, the price variation for the national retailers in the areas not affected are shown in the second row. Comparing the observed price versus that one of the scenario (0) give a variation in prices of 0.8%, around 2¢.

For the regional retailers the results are shown in the third row. The most striking result is the price reduction when comparing scenarios (2) and (0), which is of around 8¢ (8.71%). This group of retailers reduce their prices further compared to the national retailers in the areas not affected by the consolidation of distributors. As a response in equilibrium, they have decrease their prices more to remain competitive. In data, the observed prices are 19¢ lower than those of the national retailer.

The results under single-market bargaining are show in the fourth row. Notice that under the assumption of single-market bargaining there are no indirect price effects. This is a direct consequence of the model. When the negotiations are done market by market the areas for which there were no changes were not affected at all. In every scenario, under SMB prices are higher than in the baseline scenario (0). The model predicts a price increase of 6.45%, almost 15¢. This represents a significant difference between the predictions of both models.

The last four rows describe the changes in surplus for consumers and producers. For consumers, there are positive effects from the average reduction in prices when comparing the observed prices to the baseline scenario. While small, these effects go in line with the prices variations described above.

Given that the price of the product moves between $2.0 and $2.5 in the analyzed period, the nature of these effects is small. However, the main lesson from the counterfactual exercise is that assuming multi-market bargaining leads to the price effects in the regions not directly affected by the consolidation. A similar conclusion cannot be supported under single market bargaining. This is relevant when evaluating mergers between upstream providers in different markets, like the case of mergers or acquisitions between hospitals in different regions. The impact of market structure change can lead to changes in the acquirer's market of origin.

Finally, a word of caution is in order here. Notice that the main efficiency gain employed in this paper are the decrease in transportation distances. Negotiating with fewer firms might entail other efficiencies that are not being captured by the model, because I do not have access to more detailed data. However, the decrease in transportation distances taken in this paper can be taken as a lower bound on the efficiency gains that occurred in the US energy drinks market.

1.6 Conclusion

This paper relies on a reduced form and a structural model of bargaining to show that the upstream structure of the market influences retail prices. When a retail chain is in multiple geographical markets, a consolidation of regional distributors leads to a reduction in the number of firms the retailer needs to negotiate over wholesale prices with. I show that the inclusion of new regions into the bargaining process shifts the bargaining positions of the firms. Precisely, bargaining positions shift only if firms engage in *multi-market bargaining*, i.e. they negotiate for all the wholesale prices for multiple regions at once; and not *single-market bargaining*, i.e. negotiating for each market independently. To test this, I evaluate the welfare effects of the consolidation of distributors by one of the leading brands in the US energy drinks market.

I show that retail prices went down after the consolidation of regional distributors. Using a reduced form approach, I find that prices went down, in average by 1.5% in the regions under consolidation in distributors. Reduced form model shows that contracts between retailers and distributors in the Energy drinks market are negotiated for all the regions at the same time and not region by region. To further understand the price decrease I build a structural model of bargaining. The results reveal that a consolidation of distributors weakened distributor's bargaining position against *national retailers* and stop the strengthening of it against *regional retailers*. While these results are related to the geographic market, the conclusions drawn in this paper can be easily extended to conclude about product market. In the retail sector, concentration at the upstream level either in the regions covered or in the products offered potentially lessens the bargaining position of distributors.

Assessing the importance of bargaining for multiple markets is not exclusive to the retail sector. In other industries, like the health sector, where there was a recent wave of mergers, is also important to consider the role of both types of alternatives ways of bargaining. The results from the counterfactual exercises highlight that using single or multi-market bargaining can lead to opposite predictions. Antitrust authorities need to consider firms' shifts in bargaining positions when adding more regions to the bargaining process as a result of upstream changes in the market structure. In particular, the results in this paper can be used to analyze mergers between retailers in different geographic markets. While a merger between competitors can decrease competition, the overall effects depend on the structure of the upstream market as well. If distributors are just regional firms, classical antitrust analyzes should be applied. However, if distributors are national firms or if the distribution market is highly concentrated, the bargaining position for the new merged retailer could increase, representing a potential source of downward pressure on prices.

Finally, although in this paper I study the energy drinks sector in the US, future research could consider expanding the analysis to other products. While the task seems daunting, it can improve the way policymakers understand vertical structures. Additionally, some assumptions used in this paper could be relaxed. Among them, I assume simultaneous determination of retail and wholesale prices. Considering sequential pricing, as in Bonnet et al. (2021), could be incorporated to the current analyses of market specific wholesale prices. However, multi-market bargaining in a sequential pricing context can be computationally cumbersome and is left as a potential area of research to improve.

Appendix

A1 Simulation of the Model

Simulation. I start from a setting where the retailer negotiates separately with each distributor and then only with one after the distributor expands from one market to the next one. I show results under two scenarios. In the first one, the distributor from the small market expands to the large one. Scenario number two portrays the distributor from the large market expanding to the small market. Figure A1 illustrates the new equilibrium effects, comparing the results under *multi market bargaining* versus those under *single market bargaining.* The first row shows scenario one. The second one shows the effects when the distributor expands from the small to the large market. Solid lines show retail prices, while the dotted ones do the same for wholesale prices. The red lines depict prices under joint bargaining, while the blue ones do it for region by region bargaining.

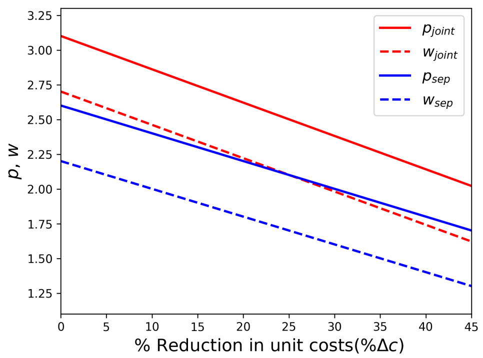
Panels (a) and (b) show that when the consolidation involves switching distributors in a large market, under joint bargaining there is an increase in retail prices in all the markets involved in the negotiation, even under cost deficiencies. Most notably, assuming that firms negotiate separately for each region predicts no effect in the market not directly affected by the consolidation. On the other hand, as shown in panels (c) and (d), when the consolidation affects directly a small market, prices are smaller under region by region bargaining. However, under *joint bargaining* prices in the large market increase. These effects are related to equation 1.3, such that the whole price in the large market will be a decreasing function of the price elasticity in the small market.

Although these results are shown for a fixed β, similar results are obtained under *small* changes in the bargaining power of the retailers, β. When the retailer is in both locations, price variations coming from changes in this parameter are not distinguishable from changes due to *multi market bargaining.* Finally, note that for a *single location* retailer the bargaining protocol has no effect, and he will always be negotiating using a *single market bargaining* protocol. If the retailer is in a large market, the blue line in figure A1 - panel (a) describes the price effects.

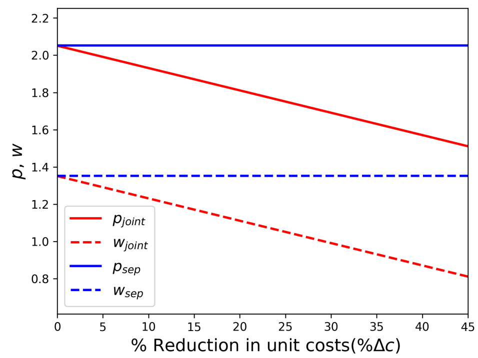
Figure A1: Price effects under different bargaining protocols and Cost Efficiencies

Scenario 1: Consolidation affects Large Market

(a) Large Market: Affected by consolidation

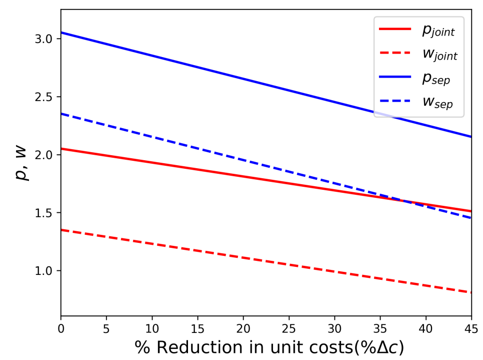


(b) Small Market: Not directly affected

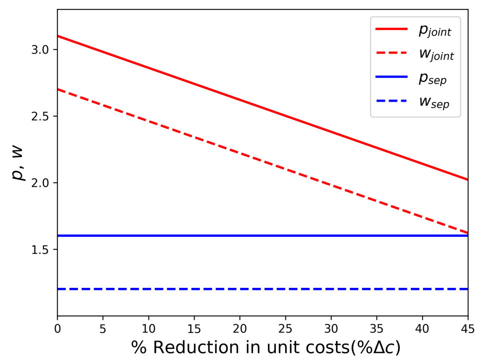


Scenario 2: Consolidation affects Small Market

(c) Small Market: Affected by consolidation



(d) Large Market: Not directly affected



Similarly, with panel (c) if the retailer has only one store, but it is located in the small market instead.

Overall, cost efficiencies alongside assumptions on the type of bargaining protocol followed by the firms have different effects depending on which market is directly affected by the consolidation of distributors:

1. When there is a *common retailer* in both affected and not affected markets by the consolidation, price effects will arise even in the region not directly affected by the consolidation.
2. Cost efficiencies can be distributed towards the regions indirectly affected by the consolidation.
3. Changes in the bargaining ability and changes in the redistribution are not distinguishable in the observed retail prices.

Although this simulation has been performed under no competition at the upstream or downstream level, similar conclusions hold when introducing competition to either market segment. When there is a consolidation of upstream firms such that it ends up covering more markets, cross-market effects will arise. As previously discussed, the level of cost efficiencies as well as the size of the market directly affected by the consolidation will determine the price effects.

A2 Small retailers

The following table describes the results for those retail chains with stores only in the regions affected by the consolidation of bottlers.

A3 Causality in reduced form

In this section I address the possibility of giving a causal interpretation to the results shown in section 1.4.1. First, Group 3 in Table 1.2 can be considered as a control group and groups 1 and 2 can be taken as two different treatment groups. From these two groups, on the first it is possible to test the effects of the consolidation, while on the second one spillover effects are tested. Naturally the spillovers are possible to test only under the assumption that the data is generated by multi-market bargaining. However, there is endogeneity for the treatment. Retail chains with more stores across the US are more likely to receive the treatment, i.e., change in distributor. It is possible that the treatment is related to the observable firm's size. To

Table A1: Changes in retail prices

	log(*price*)	
	Monster	Red Bull
$\mathbb{1}\{Treat\} \times \mathbb{1}\{Post\}$	0.009	0.002
	(0.007)	(0.005)
$\Delta MILES \times \mathbb{1}\{Post\}$	-0.014	-0.013
	(0.017)	(0.014)
Observations	115,236	180,182
R^2	0.981	0.991
Prod-Store, Prod-Region FE	Yes	Yes
Controls, trend	Yes	Yes

Standard errors in parentheses, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

incorporate this possible source of endogeneity, I follow De Chaisemartin and d'Haultfoeuille (2018) and run a fuzzy difference in difference regression. I take into account that the treatment is a function of the number of stores per region for each retail chain. I set 30 as threshold value for this ratio for possibly receiving the treatment. Using this alternative methodology, I find that prices also went down after the consolidation of distributors in the treated regions. This result is close to the one obtained using above, which could be considered as 'sharp difference in difference'.

A4 Solution to Equation 1.2

$$\{w_A, w_B\} = \underset{w_A, w_B}{\arg\max} \left[\sum_{m \in \{A,B\}} (p_m - w_m) D_m \right]^{\beta_{rb}} \left[\sum_{m \in \{A,B\}} (w_m - \mu_m) D_m \right]^{1-\beta_{rb}} \tag{1.19}$$

Where the solution to the previous equation is:

$$\sum_{m \in \{A,B\}} w_m^{**} D_m = (1 - \beta) \sum_{m \in \{A,B\}} p_m^{**} D_m + \beta \sum_{m \in \{A,B\}} \mu_m D_m \tag{1.20}$$

Re-arranging the last equation in terms for market A:

$$w_A^{**} D_A = [(1 - \beta) p_A^{**} D_A + \beta \mu_A D_A] + [(1 - \beta) p_B^{**} D_B + \beta \mu_B D_B - w_B^{**} D_B] \tag{1.21}$$

Note the first term in brackets is the solution to single-market bargaining for market A times D_A. The previous equation can be re-expressed as:

$$w_A^{**} = w_A^*(p_A^{**};\beta,\mu_A) + \Big[(p_B^{**} - w_B^{**}) - (p_B^{**} - \mu_B)\beta\Big]\frac{D_B(p_B^{**})}{D_A(p_A^{**})} \tag{1.22}$$

The retailer that maximizes the profit function for each market: $\pi_m = (p_m - w_m)D_m$. From his first order condition it is possible to express the retail price as $p_m = -D_m\frac{\partial p_m}{\partial D_m}\frac{p_m}{p_m} + w_m$. This last expression can be re-arranged as $p_m = \frac{1}{\epsilon_m}p_m + w_m$, where $\epsilon_m = -\frac{\partial D_m}{\partial p_m}\frac{p_m}{D_m}$. Finally, the wholesale price can be expressed as $w_m = p_m - \frac{1}{\epsilon_m}p_m$, or equivalently, $p_m - w_m = \frac{p_m}{\epsilon_m}$. Replacing this expression in equation 1.22,

$$w_A^{**} = w_A^*(p_A^{**};\beta,\mu_A) + \Big[\frac{p_B^{**}}{\epsilon_B(p_B^{**})} - (p_B^{**} - \mu_B)\beta\Big]\frac{D_B(p_B^{**})}{D_A(p_A^{**})} \tag{1.23}$$

Using the notation employed in the main text,

$$w_A^{**} = w_A^*(p_A^{**};\beta,\mu_A) + f_w(p_B^{**}, D_A^{**}, D_B^{**};\beta,\mu_B) \tag{1.24}$$

On the other side, the retail prices in market A are:

$$p_A^{**} = p_A^* + \left[\frac{p_B^{**}}{\varepsilon_B^{**}} - (p_B^{**} - \mu_B)\beta\right]\frac{D_B^{**}}{D_A^{**}}\frac{\varepsilon_A^{**}}{\varepsilon_A^{**}\beta - 1} = p_A^* + (w_A^{**} - w_A^*)\frac{\varepsilon_A^{**}}{\varepsilon_A^{**}\beta - 1}$$

Or equivalently,

$$p_A^{**} = p_A^* + f_w(p_A^{**}, p_B^{**}, D_A^{**}, D_B^{**};\beta,\mu_B)\frac{\varepsilon_A^{**}}{\varepsilon_A^{**}\beta - 1} = p_A^* + f_p(p_B^{**}, D_A^{**}, D_B^{**};\beta,\mu_B),$$

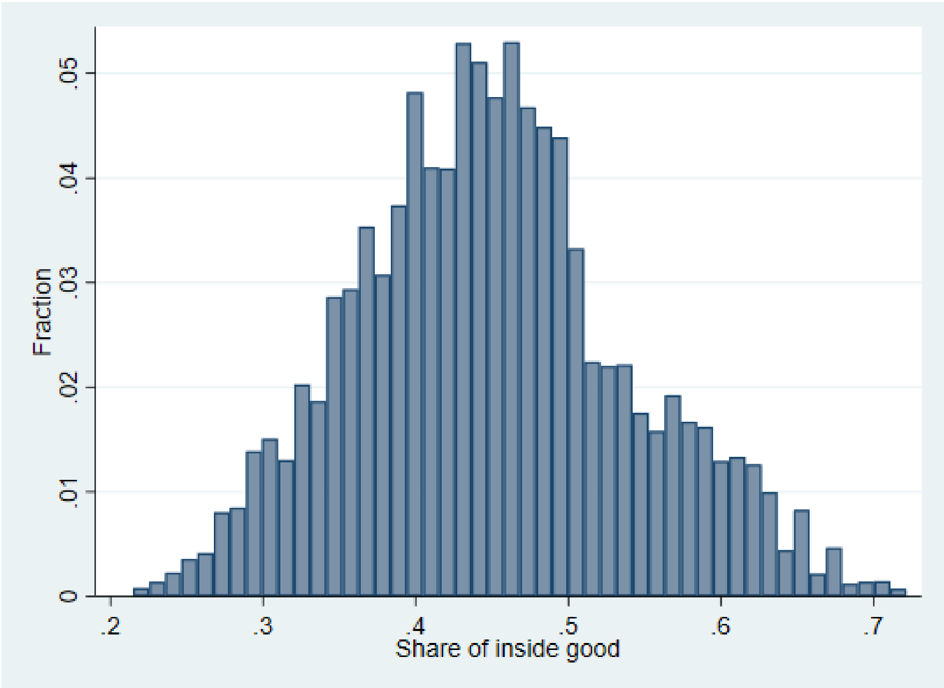
where $f_p(p_A^{**}, p_B^{**}, D_A^{**}, D_B^{**};\beta,\mu_B) = f_w(p_B^{**}, D_A^{**}, D_B^{**};\beta,\mu_B)(\varepsilon_A^{**}/(\varepsilon_A^{**}\beta - 1))$. A similar expression follows for the retail price in market B

A5 Outside Option

In this section I describe the procedure employed in this paper to compute the outside option, based on Döpper et al. (2022). However, I adjust their procedure by computing the potential market size at the region level rather than at the retail chain level. The outside option is computed using the following steps,

1. Take the population of the region between 16 and 60 years as potential population at period t and market m, POP_{mt}.

Figure A1: Distribution of Inside good



2. Obtain the total quantities sold at market m at time t, $Q_{mt} = \sum_r q_{rmt}$. This represents the total size of the inside good.

3. Compute $\gamma_m = mean_{mt}(Q_{mt}/POP_{mt})$. This ratio will be used as the average ratio of quantity to population by market in time.

4. Finally, the market size is obtained by scaling the population size to have an average share of the inside good around 0.45.

$$M_{mt} = \left(\frac{1}{0.45}\right)\gamma_m POP_{mt} \tag{1.25}$$

The distribution of the distribution of inside goods can be seen in figure A1

A6 Multi Market Bargaining equation

Equation 1.12 can be fully written in the following way

$$\sum_{m\in\mathcal{M}_{rbt}} w_{jmt}s_{jmt}L_{mt} = \sum_{m\in\mathcal{M}_{rbt}} \mu_{jmt}s_{jmt}L_{mt} + \sum_{\substack{g\in\mathcal{J}_{bmt}\backslash j,\\ m\in\mathcal{M}_{rbt}}} \Gamma_{gmt}\Delta_j s_{gmt}$$

$$+(1-\beta_{rb})\Bigg[\sum_{m\in\mathcal{M}_{rbt}}(\Gamma_{jmt}+\gamma_{jmt})s_{jmt}L_{mt}-\Bigg[\sum_{\substack{k\in\mathcal{J}_{rmt}\backslash j,\\ m\in\mathcal{M}_{rbt}}}\gamma_{kmt}\Delta_j s_{kmt}L_{mt}+\sum_{\substack{g\in\mathcal{J}_{bmt}\backslash j,\\ m\in\mathcal{M}_{rbt}}}\Gamma_{gmt}\Delta_j s_{gmt}L_{mt}\Bigg]\Bigg]$$

where $\gamma_{jmt} = p_{jmt} - w_{jmt} - c_{jmt}$ is the retailer's margin for product j in market m at time t.

This last expression can also be re-expressed as:

$$\begin{aligned} w^{**}_{jmt}s_{jmt}L_{mt} = w^{*}_{jmt}s_{jmt}L_{mt} + \sum_{\tilde{m}\in\mathcal{M}_{rbt}\backslash\{m\}} \mu_{j\tilde{m}t}s_{j\tilde{m}t}L_{\tilde{m}t} + \sum_{\substack{g\in\mathcal{J}_{b\tilde{m}t}\backslash j,\\ \tilde{m}\in\mathcal{M}_{rbt}\backslash\{m\}}} \Gamma_{g\tilde{m}t}\Delta_j s_{g\tilde{m}t}L_{\tilde{m}t} \\ +(1-\beta_{rb})[GFT^R_t(n\in\mathcal{M}_{rbt}\backslash m) + GFT^B_t(n\in\mathcal{M}_{rbt}\backslash m)], \end{aligned} \tag{1.26}$$

where $GFT^R(n \in \mathcal{M}_{rbt}\backslash m)$ and $GFT^B(n \in \mathcal{M}_{rbt}\backslash m)$ are the gains from trade for the retailer and the distributor, respectively, for reaching an agreement in every market $n \in \mathcal{M}_{rbt}\backslash m$ where both parties trade.

A7 Estimation

When implementing the search procedure for $\boldsymbol{\beta}$ that minimizes equation 1.17, I start by searching $\tilde{\beta}$ with a normalization of the parameter to search $\beta = \exp(\tilde{\beta})/1 + \exp(\tilde{\beta})$. In the second step, I perform a search process without limitations on $\boldsymbol{\beta}$.

A8 Counterfactual Algorithm

1. Get the initial conditions $\boldsymbol{p_t}^*$, $\boldsymbol{\gamma_t}^*$, $\boldsymbol{\Gamma_t}^*$. Fix initial values for the iteration at $\boldsymbol{p_t}^* \times 1.05$
2. From the expression $p_t - \gamma_t = w_t + c_t$, and knowing that $w_t + c_t = c_t + \mu_t + w_t - \mu_t$ it is possible to express: $c_t + \mu_t = p_t - \gamma_t - \Gamma_t$. With this, I compute the solution to the problem:

$$(\boldsymbol{p_t}^{POST,i} - \boldsymbol{\gamma_t}^{POST,i} - \boldsymbol{\Gamma_t}^{POST,i}) - (\boldsymbol{p_t}^* - \boldsymbol{\gamma_t}^* - \boldsymbol{\Gamma_t}^*) = \mathbf{0}$$

   Before getting into the calculation of $\boldsymbol{p_t}^{POST,i}$, I take $\boldsymbol{p_t}^{POST,i-1}$ as the starting point.

3. The process continues until $\| \boldsymbol{p_t}^{POST,i} - \boldsymbol{p_t}^{POST,i-1} \| < 0$.

The following counterfactuals are calculated separately:

1. Scenario 1: Counterfactual 1 - base scenario - $\Delta C = \Delta\beta = 0$
2. Scenario 2: Counterfactual 2 - No Change in identity - $\Delta C \neq 0$, $\Delta\beta = 0$
3. Scenario 3: Counterfactual 3 - Change in identity - $\Delta C = 0$, $\Delta\beta \neq 0$
4. Scenario 4: Observed - Change in all - $\Delta C \neq 0$, $\Delta\beta \neq 0$

Where ΔC and $\Delta\beta$ are, respectively, change in costs and change in bargaining power. I consider that the change in bargaining power comes from a change in the bottler.

Bibliography

Adams, B. and Williams, K. R. (2019). Zone pricing in retail oligopoly. *American Economic Journal: Microeconomics*, 11(1):124–56.

Backus, M., Conlon, C., and Sinkinson, M. (2021). Common ownership and competition in the ready-to-eat cereal industry. NBER Working Papers 28350, National Bureau of Economic Research, Inc.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 63(4):841–90.

Bonnet, C., Bouamra-Mechemache, Z., and Molina, H. (2021). An empirical model of bargaining with equilibrium of fear: Application to retail mergers in the french soft drink industry. Working papers, Toulouse School of Economics (TSE).

Bonnet, C. and Dubois, P. (2010). Inference on vertical contracts between manufacturers and retailers allowing for nonlinear pricing and resale price maintenance. *The RAND Journal of Economics*, 41(1):139–164.

Bonnet, C. and Dubois, P. (2015). Identifying Two Part Tariff Contracts with Buyer Power: Empirical Estimation on Food Retailing. TSE Working Papers 15-575, Toulouse School of Economics (TSE).

Butters, R. A., Sacks, D. W., and Seo, B. (2022). How do national firms respond to local cost shocks? *American Economic Review*, 112(5):1737–72.

Crawford, G. S. and Yurukoglu, A. (2012). The welfare effects of bundling in multichannel television markets. *American Economic Review*, 102(2):643–85.

Dafny, L., Ho, K., and Lee, R. S. (2019). The price effects of cross-market mergers: theory and evidence from the hospital industry. *The RAND Journal of Economics*, 50(2):286–325.

De Chaisemartin, C. and d'Haultfoeuille, X. (2018). Fuzzy differences-in-differences. *The Review of Economic Studies*, 85(2):999–1028.

DellaVigna, S. and Gentzkow, M. (2019). Uniform Pricing in U.S. Retail Chains*. *The Quarterly Journal of Economics*, 134(4):2011–2084.

Draganska, M., Klapper, D., and Villas-Boas, S. (2010). A larger slice or a larger pie? an empirical investigation of bargaining power in the distribution channel. *Marketing Science*, 29(1):57–74.

Duarte, M., Magnolfi, L., Sølvsten, M., and Sullivan, C. (2023). Testing firm conduct. *arXiv preprint arXiv:2301.06720*.

Döpper, H., MacKay, A., Miller, N., and Stiebale, J. (2022). Rising markups and the role of consumer preferences. Working Paper 22-025, Harvard Business School.

Ganapati, S. (2018). The Modern Wholesaler: Global Sourcing, Domestic Distribution, and Scale Economies. Working Papers 18-49, Center for Economic Studies, U.S. Census Bureau.

Gowrisankaran, G., Nevo, A., and Town, R. (2015). Mergers when prices are negotiated: Evidence from the hospital industry. *American Economic Review*, 105(1):172–203.

Grennan, M. (2013). Price discrimination and bargaining: Empirical evidence from medical devices. *American Economic Review*, 103(1):145–77.

Ho, K. and Lee, R. S. (2017). Insurer competition in health care markets. *Econometrica*, 85(2):379–417.

Ho, K. and Lee, R. S. (2019). Equilibrium provider networks: Bargaining and exclusion in health care markets. *American Economic Review*, 109(2):473–522.

Horn, H. and Wolinsky, A. (1988). Bilateral monopolies and incentives for merger. *RAND Journal of Economics*, 19(3):408–419.

Lewis, M. S. and Pflum, K. E. (2015). Diagnosing hospital system bargaining power in managed care networks. *American Economic Journal: Economic Policy*, 7(1):243–74.

Lewis, M. S. and Pflum, K. E. (2017). Hospital systems and bargaining power: evidence from out-of-market acquisitions. *The RAND Journal of Economics*, 48(3):579–610.

Luco, F. and Marshall, G. (2020). The competitive impact of vertical integration by multiproduct firms. *American Economic Review*, 110(7):2041–64.

Miller, N. H. and Weinberg, M. C. (2017). Understanding the price effects of the millercoors joint venture. *Econometrica*, 85(6):1763–1791.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 69(2):307–342.

Panhans, M. and Taragin, C. (2022). Consequences of model choice in predicting horizontal merger effects. Working Paper 348, Federal Trade Commision.

Rivers, D. and Vuong, Q. (2002). Model selection tests for nonlinear dynamic models. *The Econometrics Journal*, 5(1):1–39.

Sheu, G. and Taragin, C. (2021). Simulating mergers in a vertical supply chain with bargaining. *The RAND Journal of Economics*, 52(3):596–632.

Starc, A. and Wollmann, T. G. (2022). Does entry remedy collusion? evidence from the generic prescription drug cartel. Technical report, National Bureau of Economic Research.

Villas-Boas, S. B. (2007). Vertical relationships between manufacturers and retailers: Inference with limited data. *The Review of Economic Studies*, 74(2):625–652.

Vistnes, G. S. and Sarafidis, Y. (2013). Cross-market hospital mergers: A holistic approach. *Antitrust Law Journal*, 79(1):253–293.

Chapter 2

Market Structure, Entry and Product Characteristics: Evidence from the Peruvian Mobile Telecommunications Market

Oscar Jara[1] and Nicolás Martínez[2]

Abstract:

Using a unique dataset covering the full Peruvian mobile telecommunications market from 2010 to 2019, we investigate how much of the variation in product offerings comes from changes in market structure and how much is due to technological change. This period contains both the introduction of 4G connectivity in the country and the arrival of two low cost providers. Motivating evidence shows a sharp increase in data consumption and a larger dispersion in the characteristics of mobile plan offerings after the occurrence of these two events. We propose a structural model representing a two stage game were firms chose both which mobile plans to offer and their prices. Demand estimates indicate an important increase in the willingness to pay for larger data allowances after the introduction of 4G. Using moment inequalities, we partially identify a firm specific fixed cost of offering a mobile plan. Finally, we outline counterfactual scenarios for future work that analyze the role of competition and technological progress.

JEL Classification: D22, L13, L96

[1]Toulouse School of Economics, Université Toulouse 1 Capitole.

[2]Toulouse School of Economics, Université Toulouse 1 Capitole.

We are grateful for the support and data received by the Peruvian Telecommunications Regulator, OSIPTEL.

2.1 Introduction

Industries such as the markets for cars, computers, smartphones, or mobile telecommunications plans, are characterized by a stable number of multiproduct firms. When introducing new products, firms face a trade-off between business stealing and cannibalization, which can lead to restrictions in the available product varieties. However, changes in market structure could affect those incentives. For instance, an increase in competition could lead incumbents to offer more and more varied products. Moreover, technological progress can improve the quality of some product characteristics, affecting demand patterns and in turn affecting firms' offerings. In industries experiencing rapid technological change, such as telecommunications, disentangling the effects from competition and technological progress can have important consequences when informing policymakers.

In this paper, we build a structural model to understand how firm entry affects the choice of product offerings and prices. We apply our model to study the Peruvian mobile telecommunications market where 2 firms entered a market previously dominated by two incumbents. After entry, incumbents started offering more plans, with higher variation in the number of minutes and gigabytes offered. Moreover, the entry of the new firms in the market coincided with the introduction of a new technology in the market, 4G connectivity. This technological change affects consumers' valuation of data and thus demand patterns, inducing firms to offer more plans with data allowances. However, as both entry and 4G occur at the same time, it is not clear how much of the changes in product offerings come from each source. The effects of firm exit (Crawford et al., 2018) or firms consolidation (Fan and Yang, 2020) on product variety have already been studied in the literature. Our goal is to complement this body of research by exploring the effects of firm entry on incumbents' decisions to release new products, when interacted with demand changes due to technological progress. By determining the equilibrium product offerings absent of the additional competitive pressure imposed by the entry of new firms, we could identify the role of technological changes.

We use a novel data set from the Peruvian Telecommunications Regulator. Given that firms are required to upload all the features included in their plans to the Tariff Consultation System, we web-scrapped all these data for the period 2010 - 2019. Additionally, we got directly from the regulator information on firms' investments, revenues, infrastructure, number of lines in services, among others. In particular, we received detailed information on the position of the antennas of 2G, 3G and 4G and are able to track its evolution in time. Finally, we also got access to information on other investments done by the firms.

Motivating evidence indicates that prices for the prepaid tariffs (where consumers pre-purchase a bundle of minutes/data) decreased significantly, while the prices for the postpaid plans (contract with a monthly amount of minutes/data) remained, on average, constant or even increased slightly. However, when looking at product offerings, evidence suggests that incumbents were not exploiting the full range of plans they could have released. After entry of the low cost operators, the incumbents offered plans containing a wider range of gigabytes, giving more options to consumers. While technological progress represents a challenge to the use of the previous evidence to establish a direct mechanism, it also serves to motivate the modeling of the joint decision of product choice and prices.

First, we estimate a nested logit demand for products, where these are defined by the ranges of minutes and gigabytes they include. We find, as expected, a negative price sensitivity, implying an average own price elasticity of -1.87. This low price elasticity could partly be explained by consumer inertia or unobserved switching costs when changing operator. Using these estimates, we recover the willingness to pay (W.T.P.) before and after the introduction of 4G. We find that the average W.T.P. for one gigabyte of data increased by a factor of 14 after the introduction of 4G, highlighting the important role of technological change on demand patterns.

We model supply as a 2 stages game. In the first stage, firms decide their plan offerings based on expected profits. In the second stage, firms compete on prices given the set of products chosen in the previous stage. Using moment inequalities, we partially identify the fixed costs of each firm. To deal with the selection bias coming from the observed product offerings, we use the same correction as Eizenberg (2014). Results from the estimation show very similar fixed cost for both incumbents. However, for one of the entrants find much larger bounds, due to the fewer plans they introduced across time.

In the next section, we describe the industry background and the data. Next, we introduce reduce form evidence to motivate the introduction of our structural model. In section 2.4 we describe both the demand and supply side of the two stage model. Section 2.5 presents the results from the estimating the structural model. Finally, we discuss counterfactual scenarios that could be analyzed with the model.

2.1.1 Literature Review

This paper contributes to the literature on endogenous product characteristics. Berry and Waldfogel (2001) find that market consolidation in the radio industry leads to a decrease in the total number of stations but to larger differentiation between stations' formats. McManus (2007) shows that firms with a product line distort less the characteristics of the products with the lowest margins. Fan (2013) develops a structural model of newspaper competition and shows that in the evaluation of any change in market structure, the equilibrium changes in product characteristics play an important role in the computation of welfare effects. Wollmann (2018) develops a structural model of equilibrium product offerings for the US truck market. He finds an important role of product entry and exit on the welfare effects of the hypothetical exit of a firm. Fan and Yang (2020) study the role of competition on the number and characteristics of smartphones in the US, finding that a reduction in the number of competing firms would lead to a decrease in both the number of products offered and in the dispersion of characteristics across products. Chatterjee et al. (2022) study how the rollout of 4G in India affects the firms' decisions on which smartphones to offer, indicating the existence of spillover effects affecting product offerings across markets.

The role of market structure on prices has been studied previously (Miller and Weinberg, 2017; Miller et al., 2021; Starc and Wollmann, 2022). The first two papers show a price increase after a joint venture between the second and third-biggest brewers in the US beer industry. They find that the industry follows a setting that resembles a repeated game of price leadership, in which a leader proposes super-markups over Bertrand prices to a coalition of rivals. A consolidation in the market structure relaxes the leader's incentive constraint and increases prices. On the other hand, Starc and Wollmann (2022), shows how entry of new firms to the market can destabilize a cartel, reducing prices. Finally, in a series of papers applied to the cable television market, Crawford (2012), Crawford and Yurukoglu (2012), and Crawford et al. (2018) study the role of competition in product quality. These papers show that in equilibrium, the presence of high-end competitors motivates quality over provision by traditional cable TV providers.

Our work complements the previous literature analyzing the role of market power on entry and product offerings in telecommunications. Economides et al. (2008) study the welfare effects of entry into local telephone markets, finding that most consumer surplus gains come from the introduction of new plans and not from price effects. Fan and Xiao (2015) use a dynamic entry game to study the role of subsidies aiming at incentivizing entry of the local telephone providers and reduce market power. Seim and Viard (2011) study the role of entry of cellular services providers product offerings, finding that it leads to a more uniform distribution of characteristics across plans and price decreases across plans. Bourreau et al. (2021) study the market structure on product lines offered in France. They find that entry unravels a tacit collusion equilibrium, leading to an increase in offered variety and a decrease in prices.

We incorporate the role of 4G antennas on service quality and product offerings, bridging the product offering and infrastructure investment stages existing in this market. Previous literature has studied the role of regulations and competition on infrastructure investment in telecommunications. Genakos et al. (2018) study the role of mergers on prices and investment decisions, finding that they lead to price increases and do not seem to have an impact on industry-wide investment. Marcoux (2022) uses a static model to study the role of entry on cellphone operators investments in new cell antennas, finding an important role for economies of density. Lin et al. (2023) study the deployment of 4G antennas in the United States, finding that a decrease in competition would lead lower investment, specially in rural areas. Granja (2022) uses a dynamic model to analyze the influence of coverage requirements on investment decisions, showing they help to speed up the introduction of new technology. Elliott et al. (2023) use engineering principles within a price competition model to study a trade-off between market consolidation and economies of scale in mobile telecommunications. Bourreau and Sun (2022) use a dynamic model to study the role of competition on quality and investment.

We follow previous literature on product offerings and market entry, modeling market dynamics as a two stage game (Draganska et al., 2009; Nosko, 2010; Wollmann, 2018; Bontemps et al., 2023). In the first stage, firms choose which products to offer, while in the second stage they compete on prices conditional on the products chosen before. To estimate the cost linked to the offering of different plans, we use moment inequalities as in Eizenberg (2014), Pakes et al. (2015), and Fan and Yang (2022).

2.2 Context and Data

2.2.1 The cellular network market in Peru

Until 2011, Peru's cellular network market was composed by two large multinational operators, Claro and Telefonica, which offer the service to consumers. A third operator, Nextel Perú S.A., offered the service exclusively to business. In 2011, the available spectrum for the 1900 MHz band and the allocation of the AWS band (1.7/2.1 GHz) for mobile communication services were auctioned. This allowed the entrance of two new providers, Viettel Telecom and the Entel Chile Group[3][4].

In 2011, Viettel was awarded the spectrum concession for Block C of the 1900 MHz Band to be able to provide mobile communications service in Peru. Later, in August 2012, additional spectrum was awarded in the 900 MHz band. Entry was accompanied by investments in infrastructure, as the company deployed its own infrastructure in the main cities of the country. Viettel's business operations started on July 26, 2014, under the trademark 'Bitel'. The almost three years delay between the spectrum auctions and the start of business operations arose

[3]The allocation of the bands was assigned by a public contest, where the firm with the highest score gets the tender. The scores are assigned not only on the money the firm is willing to bid to exploit the band, but also in investments commitments as well as other deeds, like giving free internet to rural or suburban schools.

[4]The Peruvian Ministry of Transport and Communications aimed fostering the use of broad band to close the gaps in coverage in Peru and start adopting the 4G technologies.

from two main reasons. First, to start building their infrastructure like the base stations or the antennas around the country, they required municipal permits that were not given on time or which approval was left pending. Second, the recently deployed fiber optics were constantly stolen, delaying the advancement of the installation of their network.[5]

On August 20, 2013, the Entel Chile Group acquired the assets of Nextel Perú S.A. for US$400 million. This acquisition represented a significant milestone in Peru's telecommunications sector, consolidating the market position of the Entel Chile Group and expanding its presence in the Peruvian market. Subsequently, in June 2014, the Entel Chile Group successfully secured Block B of the AWS band, which was specifically allocated for advanced high-speed mobile services (4G) utilizing LTE technology. This strategic move not only demonstrated Entel's commitment to technological innovation but also positioned the company as a key player in delivering cutting-edge mobile services to consumers in Peru.[6]

Despite acquiring Nextel Perú S.A.'s asset, Entel Chile Group continued to operate under the 'Nextel' brand until October 10, 2014. This interim period allowed for a seamless transition of operations and customer services while preparations were made for a rebranding initiative. On October 10, 2014, the company officially underwent a rebranding process and adopted the name 'Entel Perú' (henceforth, Entel). The new company embarked on focused on delivering telecommunications services to consumers rather than business. With a robust infrastructure and a dedicated team of professionals, Entel Perú aimed to provide reliable connectivity and advanced services to its customers across the country. As part of the promotional strategy, Entel Perú offered exclusive deals and incentives to subscribers from other providers to switch to Entel, including discounted plans, free device upgrades, and bonus data packages.

Parallel to the entry of new firms to this market, there was an important regulatory change. A Mobile Number Portability (MNP) law was enacted in Peru in 2010. Under this legislation, consumers gained the ability to retain their phone numbers when switching between different mobile service providers. This process involved the timeline of 7 days for porting phone numbers between providers. By allowing consumers to keep their existing phone numbers regardless of the service provider, the 2010 legislation aimed to reduce barriers to switching providers and stimulate competition within the telecommunications market. Nonetheless, its application had a limited success, with almost no consumers porting. Building upon the initial success of the 2010 legislation, further enhancements to MNP laws were introduced in Peru in December 2013. The most important amendment sought to improve administrative processes, by reduce

[5]See: https://www.gob.pe/institucion/osiptel/noticias/.

[6]The acquisition and integration process was facilitated by Americatel Perú S.A., a subsidiary of the Entel Chile Group. Formerly focused solely on long-distance calling services, Americatel Perú S.A. expanded its operations in Peru by acquiring the necessary permits to utilize the AWS band in July 2013. By June 2014, they successfully transferred the rights to Nextel Peru S.A., facilitating the integration of advanced 4G services into the Peruvian market.

to timeline to port the number to 1 day. This new version of the MNP was introduced to the market from July 16, 2014.

Finally, notice that in this paper we analyze the mobile telecommunications market during the period 2010 - 2019. We investigate this period because the entry of the two new firms happens in the middle of it, and also because of the COVID pandemic in the early 2020. Although we study the effects of the entry of Bitel and Entel to the market, in the analyzed period, other firms entered the market but none of them got a significant market share. Virgin Mobile was a Virtual Mobile Operator (VMO) entered the Peruvian market in July 2016. [7] However, this company was involved in a media scandal upon their entrance to the Peruvian market, which had a significant impact on their early operations. With the goal of achieving a 2% market share in three years, they left the market in August 2017, with a 0.15% of the market. Their assets were bought by a new VMO, Incacel Peru S.A., who by the end of 2019 had a market share below 1%.[8] The incumbent Telefonica introduced the brand 'Tuenti' to fight for the same 'young adults' segment as Virgin Mobile. However, after the exit of Virgin Mobile they close the brand in 2019. [9][10] Given their low market share and short stay in the market, we do not include them in the analyses.

The incumbents, Claro and Telefonica, and the entrants, Bitel and Entel; offered a significant number of plans according to the commercial strategies. These can be grouped according to the requirement to have a contract or not. Contracts indicate that consumers commit to a monthly payment in exchange for the traits offered by the plan. Plans without contracts are called *prepaid*, whereas the ones that have one are called *postpaid*. For the prepaid plans, consumers buy a bundle of minutes and/or data for a certain amount of money. Since prepaid does not tie consumers to a monthly payment, consumers engage in 'recharges' of their minutes and/or data every time they run out of them. The nature of these 'recharges' is sporadic and depends on the consumption pattern of each consumer.

[7]Virtual Mobile Operators (VMOs)are companies that operate without a spectrum license but are perceived by consumers as equivalent to traditional mobile operators (OMRs). VMOs typically rent the telecommunications infrastructure from the OMRs. VMOs have emerged in various countries to overcome the main barrier to entry into the mobile market, namely the scarcity of radio spectrum. In Peru, on September 22, 2013, Law No. 30083 established measures to strengthen competition in the market for Mobile Public Services by introducing the figure of Virtual Mobile Operators.

[8]See the information here.

[9]Although Tuenti it was introduced as a 'fighting brand', the stop in the commercial growth of Virgin Mobile can be traced back to their failed marketing strategy. See the information here.

[10]See the information here.

2.2.2 Data

By combining public data and data requested to the Peruvian Telecommunications Regulator (OSIPTEL), we construct a novel dataset that allows to study in detail the Peruvian cellular network market. Due to the regulations in place, firms operating in the market are required to declare all the features included in their plans. To comply with the requirement, they submit the information to the regulator, either directly to the web using specific forms. Therefore, we work with three main data sources.

First, firms are required to upload all cellphone plan information to the Tariff Consultation System (SIRT, by its acronym in Spanish). We scraped data from the SIRT for the years 2010 to 2019 for all the firms in the market. This contains information such as the dates in which a plan was available to be acquired in the market, the number of included gigabytes, minutes for calls (within provider's own networks and for calls to other networks), SMS messages, type of technology (4G/5G), if the offer is for a specific demographic group, restrictions, among others. The SIRT also registers if the plans gave free access to certain applications such as WhatsApp, Facebook, Instagram, Twitter or Waze.

Second, we use annual survey data from the Residential Survey of Telecommunications Services (ERESTEL, by its acronym in Spanish). This survey has information on both household demographics and on the tariff category of the mobile plans acquired. It also gives information the characteristics of the purchased plans, as minutes gigabytes or social networks as well as the range of consumption of these services they consume. Finally, the ERESTEL also gives information on the range of monthly expenditure of each of the telecommunications services the surveyed households get. The third dataset we used is called PUNKU, and also comes from the regulator. It depicts aggregated information that is available at monthly level and shows information on the total active number of phones (lines) in the market per tariff category, region, time and firm.

Finally, we use non-publicly available data provided by OSIPTEL. The Periodic Information Requirements Standard (NRIP, by its acronym in Spanish) mandates that firms in the Peruvian telecommunications market submit information regarding their operations, investments, revenues, among others. These data allow us to observe the evolution of the number of antennas and coverage of 2G, 3G and 4G both across time and space.

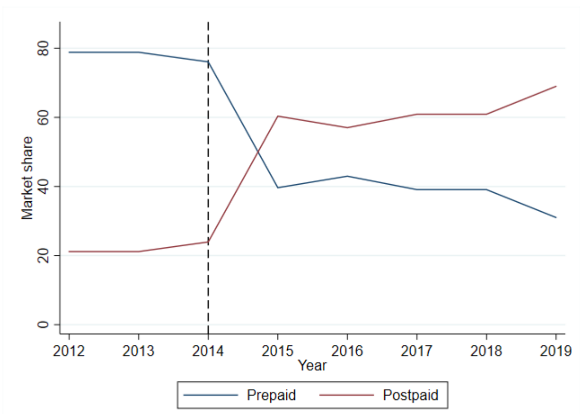
2.3 Motivating Evidence

We examine the two entry events, mentioned in subsection 2.2.1, that occurred in July and October 2014. These events allow us to document how market structure and technological change impact product offerings and prices. In Subsection 2.3.1, we find that the change in market structure was linked to a change in the type of plans chosen by consumers with a large increase in demand for postpaid plans, which offer better terms for data usage. However, as Subsection 2.3.2 shows, the increased competition only lead to price decreases in the prepaid category and not in the postpaid category. In Subsection 2.3.3 we show how the entrants' strategy was to differentiate with respect to the characteristics of the postpaid plans offer. We find that over the years, incumbents react by changing as well the type of plans offered, considerably increasing the data offered. However, this motivating evidence cannot differentiate how much of these changes is coming from competitive pressure and how much comes from changes in consumer preferences, highlighting the need for a structural model to fully understand the market's evolution.

2.3.1 Market evolution

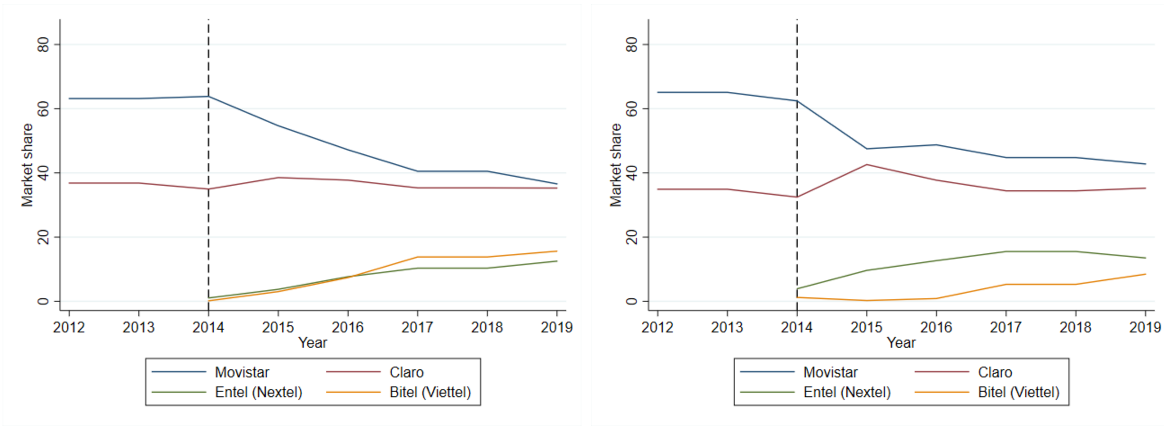
Figure 2.3.1 describes the evolution across time of the demand for different types of plans (prepaid and postpaid). After the entry of the two low-cost providers in 2014, we observe a sharp drop in the demand for prepaid plans and a large increase for postpaid plans. This trend is consistent with an increased demand for data usage. Also, considering this last point, both incumbents and entrants started offering plans which minimum price was lower than before, motivating the switch of many consumers from prepaid to postpaid. Finally, another reason for the drop in the shares of the prepaid was Entel, one of the entrants, focused its commercial strategies to attract consumers to the postpaid sector by offering special deals. Incumbents reacted to the entrant with their own new deals, which fostered the switch from prepaid consumers to postpaid.

Figure 2.3.1: Market shares per type of plan



Note: Computed using ERESTEL survey data.

Figure 2.3.2: Market share evolution



(a) Market shares in prepaid plans

(b) Market shares in postpaid plans

Note: Both graphs are computed using ERESTEL survey data.

Figure 2.3.2 shows the evolution of market shares for the different firms in the prepaid (Figure 2.3.2a) and postpaid plans (Figure 2.3.2b) segments. In prepaid, we observe a large decrease in the market share of the leading firm (Movistar) with gains from both low-cost providers and an almost unchanged share for the second incumbent (Claro). Meanwhile, in the postpaid market we observe large substitution from the market leader (Movistar) towards the other incumbent and one of the low-cost providers (Entel). Interestingly, the other low-cost provider (Bitel) keeps a low market share in postpaid until 2016, indicating that it was focusing on the prepaid market during its first years of operation. Given Peru's nature of being a developing country, Bitel oriented their commercial strategies towards low income consumers, who are the ones that typically purchase prepaid plans. A common strategy for them was to give mobile phone chips for free with an amount of minutes.

2.3.2 Evolution of prices

To check if the changes in market structure lead to changes in prices, we study the effect of the entry of each one of the new operators on the price of minutes in cellphone plans. Table 2.3.1 conducts event study regressions to measure the impact of entry times ($t > t^*$) on the average price. The results show that although the dummy that accounts for prepaid prices was positive, the interaction with the entry date t^* was negative for the month of October 2014, which was the month of the second entrant. While we find a correlation between the entry of new firms and a price drop for the prepaid plans, causality cannot be established from this exercise.

Table 2.3.1: Price effects under different thresholds

	Average Price			
	May-14	June-14	July-14	October-14
Constant	0.165***	0.163***	0.162***	0.159***
	(0.023)	(0.021)	(0.020)	(0.017)
$\mathbb{1}\{t > t^*\}$	0.008	0.010	0.011	0.018
	(0.029)	(0.028)	(0.027)	(0.027)
Prepaid	0.073*	0.075**	0.075**	0.079***
	(0.040)	(0.036)	(0.034)	(0.028)
Prepaid $\times \mathbb{1}\{t > t^*\}$	-0.050	-0.054	-0.056	-0.067**
	(0.043)	(0.040)	(0.037)	(0.033)
Trend	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	432	432	432	432
R^2	0.424	0.425	0.426	0.43

Notes: Includes firm fixed effects. Standard errors in parentheses. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Furthermore, these results reveal that there are no noticeable effects on the prices of the postpaid plans. Nonetheless, it is important to note that each tariff category, such as postpaid and restricted postpaid, contains dozens of plans at any point in time, each with specific characteristics such as gigabytes, minutes, SMS, unlimited data for certain applications, among others.

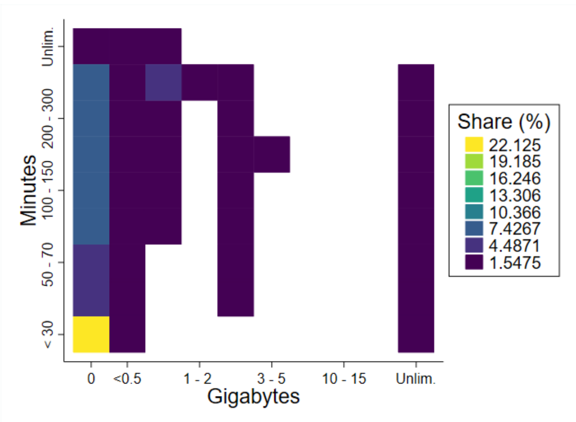
2.3.3 Product offerings and demand for different characteristics

The previous section showed how entry was accompanied by a reduction in prices in the prepaid offerings but had no effect on postpaid plans. While this result is surprising given the increasing demand for postpaid plans showed in 2.3.1, in this section we show evidence of non-price responses to the increase in competition. While competing for postpaid clients, firms offer several plans with different characteristics such as number of gigabytes, minutes, SMS, or unlimited data for certain applications. Figure 2.3.3 shows the evolution of the distribution of plans offered by incumbents and entrants across their two main characteristics, gigabytes and minutes offered.

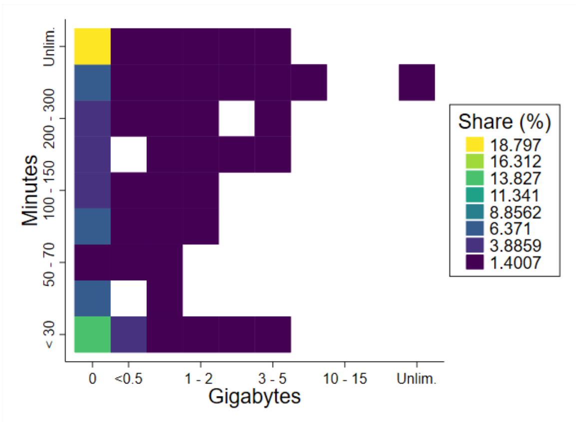
As it can be seen in Figure 2.3.3a, before entry, the incumbent firms focused the majority of their offer in categories with no gigabyte allowances. As well, while they offered several levels of minutes, most offerings are focused on low amounts of minutes. Instead, the offers from the entrants for the same year, showed in Figure 2.3.3b, include a larger variety of gigabyte allowances and are also particularly concentrated on plans offering an unlimited amount of minutes. These differences in products offered evidence a strategy of differentiation from the entrants, aiming at grabbing market share by increasing considerably the quantity of minutes and data offered to consumers.

We can compare the distribution prior to the change in market structure with the distribution of plans five years later, in 2019, and study how firms have reacted to the change and to the new demand patterns from consumers. Figure 2.3.3c shows the distribution of plans for the incumbent firms, while Figure 2.3.3d shows it for the entrant firms. First, we observe both types of firms have presence over a larger part of the gigabytes-minutes space. In particular, incumbents have focused their offers on the unlimited minutes and several levels of gigabyte allowances. While the entrants offer the same type of plans, they have a large set of plans with either low or unlimited minutes and no data.

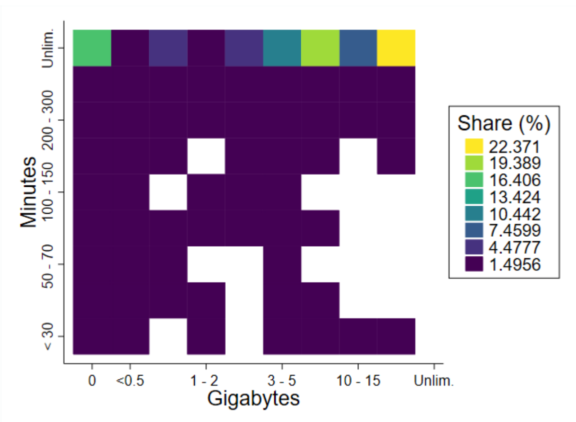
Figure 2.3.3: Evolution of minutes and gigabytes offered in postpaid plans



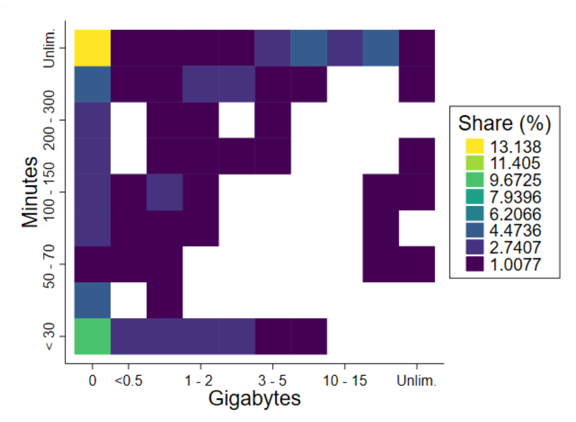
(a) 2014 - Incumbents



(b) 2014 - Entrants



(c) 2019 - Incumbents



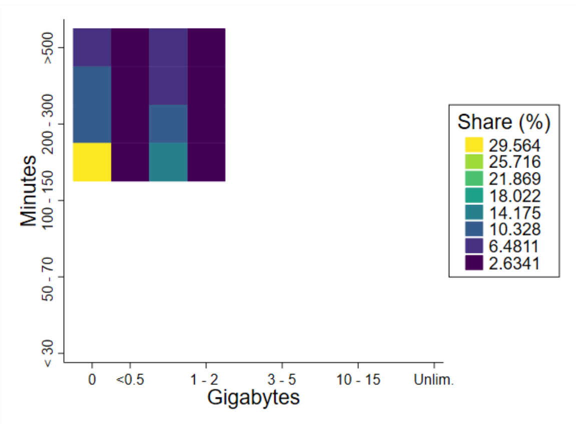
(d) 2019 - Entrants

Note: Computed using SIRT data.

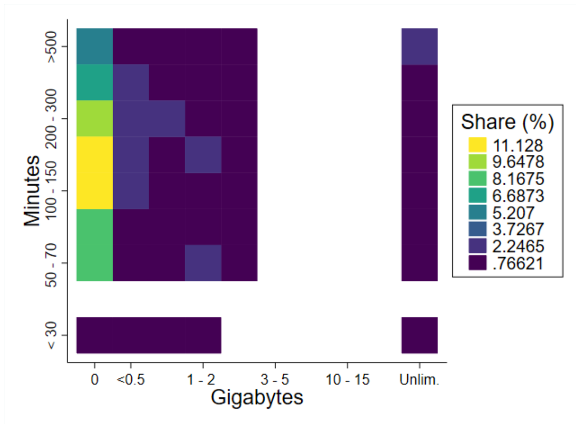
Variation in consumption patterns

Figure 2.3.4 shows the evolution in the consumption of minutes and gigabytes. Before the entry year, Figure 2.3.4a, most consumers have large minutes usage and low data usage, consistent with the previously documented product offerings at the time. During the entry year, Figure 2.3.4b, we see a more evenly distributed consumption of minutes, reflecting a change in consumption, where individuals use fewer minutes but demand more data. Interestingly, such pattern does not hold across time, as consumers increase their usage of both minutes and data, as evidenced in Figures 2.3.4c and 2.3.4d.

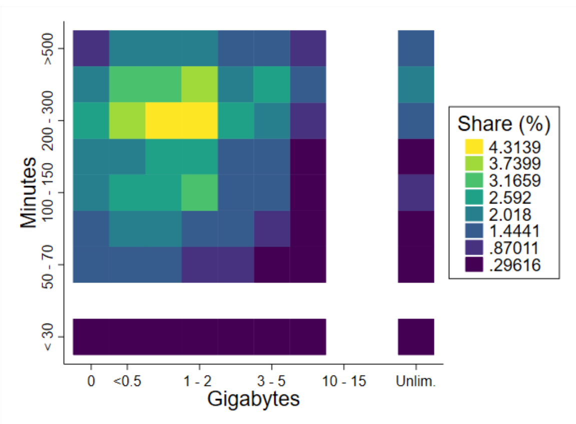
Figure 2.3.4: Evolution of consumption for minutes and gigabytes



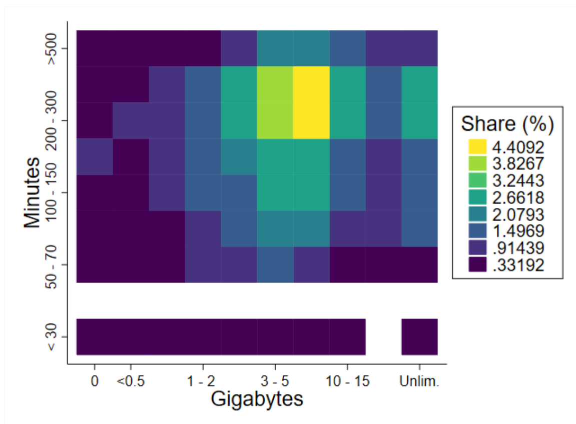
(a) 2013 - 1 year before entry



(b) 2014 - Entry year (Oct)



(c) 2016 - 2 years after entry



(d) 2019 - 5 years after entry

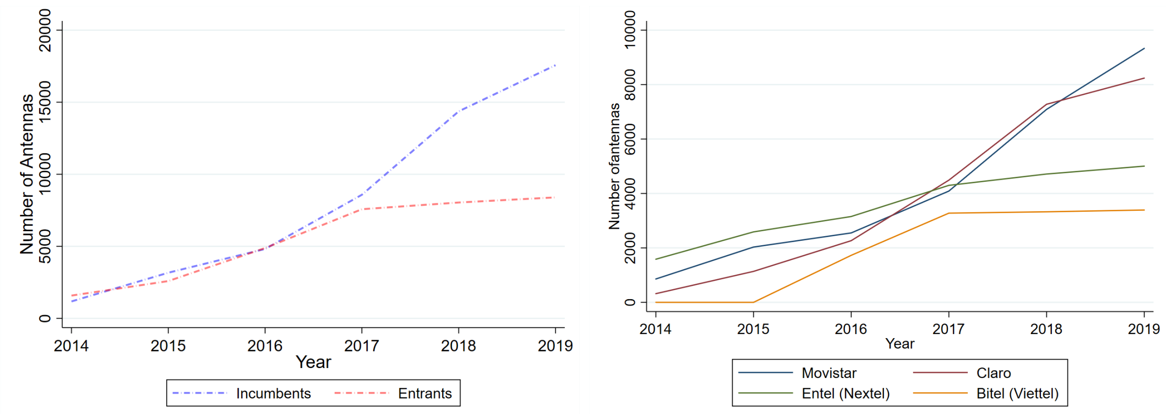
Note: Computed using ERESTEL survey data.

These results evidence how across the period of interest consumer start demanding more data and have a more spread demand for minutes. These changes in demand depend both on changes in consumer preferences, prices, and product offerings, highlighting the importance of using a structural model in order to properly disentangle the role of the new technology, 4G connectivity, and increased competition.

2.3.4 Investment dynamics

Finally,in Figure 2.3.5 we study the evolution of antennas at the nationwide level. This variable allows us to study the long term investment decision of firms, as well as the quality of the service, as more antennas should lead to better quality in calls and data download. Figure 2.3.5a shows the evolution of the 4G antennas by firm type. Interestingly, until 2017, both types of firms have similar numbers of 4G antennas. However, from 2017 onward, entrants stop investing in additional antennas while the incumbents increase considerably their investment. Figure 2.3.5b shows the evolution by brand, Bitel is a late entrant to the 4G sector, having antennas only from 2016 onward. As it was mentioned in section 2.3.1, Bitel's commercial strategy was oriented to postpaid, with an intense offer on lower prices per minute of voice calls. So, installing 4G antennas was not a priority for them.

Figure 2.3.5: Nationwide number of 4G antennas



(a) By firm type

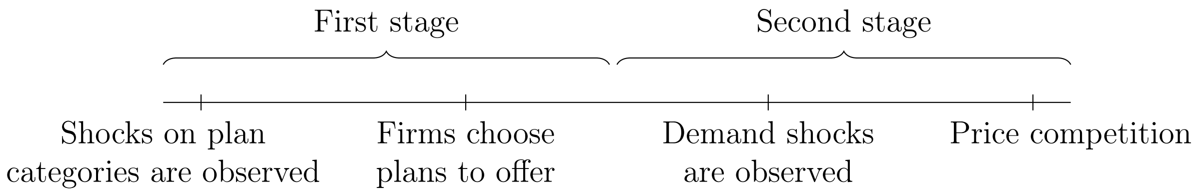
(b) By brand

Importantly, both incumbents increase investment in new antennas from 2017 onward, showing that the trend in Figure 2.3.5a was not driven by a single firm. These trends could signal an important role of competition on investment on the first years of the service, while coverage is built on high demand urban areas. The decrease in investment from the entrants after 2017 was related to a lack of investment in rural areas. Such locations have a lower demand, making them less attractive for the entrants. As mentioned before, Bitel's strategy was heavily oriented to prepaid plans, so their lack of investment in 4G antennas. Entel, on the other hand mostly used the infrastructure they acquired from Entel, making lower additional investments. In future versions of the paper we plan to cover the role of competition on investment and its geographic distribution.

2.4 Model

We model the market, as a static game with two stages, represented in Figure 2.4.1. In the first stage, firms observe shocks affecting their valuation of the different product offering categories, based on this information they form expected profits and choose which products to offer. For postpaid plans, a product is defined as a bundle of minutes and gigabytes allowed. Notice that firms consider the expected profits, before the realization of demand and cost shocks that will affect the second stage of the game. In this second stage, firms compete on prices given the set of products chosen in the previous stage.

Figure 2.4.1: Game timing



2.4.1 Second stage: Demand, price competition and marginal costs

Demand for cellphone plans

To estimate consumer demand, we use a nested logit model as in Berry (1994). We divide plans into two groups prepaid and postpaid. Within the postpaid category, firms can virtually offer any combination of minutes and gigabytes. To simplify our analysis, we divide the space of possible postpaid plan offerings into a two-dimensional grid of minutes ($]0, 50]$, $]50, 100]$, $]100, 250]$, ≥ 250) and gigabytes (0, $]0, 2]$, $]2, 5]$, $]5, 10]$, ≥ 10) categories. Thus, a plan is defined as $j \in \mathcal{J}$, where $\mathcal{J}$ includes all the possible combinations of characteristics from the grid and a prepaid offering. Each firm f offers the set of plans $\mathcal{J}_{ft} \subseteq \mathcal{J}$, always including a prepaid offering. Finally, each consumer purchases one of the available plans $\mathcal{D}_t = \{\mathcal{J}_{1t} \cup \mathcal{J}_{2t} \cup \mathcal{J}_{3t} \cup \mathcal{J}_{4t} \cup 0\}$ where 0 represents the outside option of not having a cellphone plan.

The conditional indirect utility that consumer i receives from purchasing plan j from firm f in market m and period t is:

$$u_{ijft} = \delta_{jfmt} + \epsilon_{ijfmt} \tag{2.1}$$

Where $\epsilon_{ijfmt} = \zeta_{igmt} + (1 - \sigma)\,\tilde{\epsilon}_{ijfmt}$ is iid following a type one extreme value distribution and g represents all plans withing one of the two groups (prepaid and postpaid). When $\sigma = 1$ substitution happens within the plan's type nest. If $\sigma = 0$ all products are substitutes as in a standard logit model.

$\delta_{jfmt} = \alpha_m \left(p_{jft}/\bar{y}_{mt}\right) + \boldsymbol{x}_{jt}\boldsymbol{\beta}_{jt} + \xi_{jfmt} + \tau_{jf} + \tau_t$ corresponds to the utility provided by the different characteristics of the product. p_{jft} corresponds to the average price of the plans in bin j offered by firm f and $\bar{y}_{mt}$ is the average income in market m at period t. $\boldsymbol{x}_{jmt}$ is a vector with the average number of gigabytes and minutes offered in bin j, while ξ_{jfmt} is a plan specific unobservable. Each product is available in every geographical market at time t. Therefore, we assume all consumers have the same choice set. We use as price the average among all the plans available in the bin. Finally, τ_{jf} and τ_t are product and market at time t fixed effects. We normalize the utility of the outside option to 0 and given the utility specification, we have the following aggregate market share for plan j offered by firm f:

$$s_{jfmt} = \frac{\exp\left(\frac{\delta_{jfmt}}{1-\sigma}\right)}{D_g^{\sigma} \sum_{g' \in \mathcal{G}} D_{g'}^{1-\sigma}}, \tag{2.2}$$

where $D_g = \sum_{j \in g} \exp\left(\frac{\delta_{jfmt}}{1-\sigma}\right)$ and g represents the group (prepaid, postpaid, outside option). Finally, using the inversion from Berry (1994), we obtain the linear equation:

$$\log\left(s_{jfmt}\right) - \log\left(s_{0mt}\right) = \alpha_m \left(p_{jft}/\bar{y}_{mt}\right) + \boldsymbol{x}_{jt}\boldsymbol{\beta}_{jt} + \sigma \log\left(s_{jf|gmt}\right) + \xi_{jfmt} \tag{2.3}$$

Where $s_{jf|gmt} = s_{jfmt}/s_{gmt}$, and s_{gmt} is the share of group g in market m at period t.

We consider instruments that shift supply but not demand. The first set of instruments we use are cost shifters. These account for the different costs between products which secure enough variation to differentiate them. This involves using the costs that firms need to pay for when there is a phone call finishes at other operator's network. By law providers must pay a 'termination rate' if a phone call ends in the network of other providers. So, using information on the minutes that end up in other networks, we calculate the total costs for termination rates. The second set of instruments considers the distance in the characteristics space between a product and others products. Define $h_{jk,t}^{(l)} = h_{k,t}^{(l)} - h_{j,t}^{(l)} \quad \forall \quad l \in \{Minutes, Gigabytes\}$. With this, we compute the following instruments for plan j:

$$z_{jt}^{*} = \left\{\sum_k {d_{jk,t}^{(l)}}^2, coverage_{-f(j),t}, share\_prods_{f(j),t}\right\} \quad \forall \quad l \in \{Minutes, Gigas\}$$

$$z_{jt} = z_{jt}^{*} \otimes \{1, inc_t^{10\%}, inc_t^{50\%}, inc_t^{90\%}\}$$

Given that all the products are offered nationally, we interact the distances in the characteristics space among products with demographics. As in Backus et al. (2021), we use the 10%, 50% and 90% quantiles of the regional income distribution. This helps to capture the regional heterogeneity in consumer income levels. Additionally, using interactions with moments of the income distribution allows controlling for the potential correlation between income and unobserved factors that may affect demand. Finally, for the nesting parameter, we also build an additional instrument based on the number of products per nest. This instrument should be negatively correlated with the conditional share $s_{jf|gmt}$.

Pricing and marginal cost

The firm maximizes its profit function, by setting prices across all the markets $m \in \mathcal{M}$.[11] Defining the vector of prices as $\boldsymbol{P}_{ft}$ Its maximization problem is given by:

$$\max_{\boldsymbol{P}_{ft}} \left\{ \Pi_{ft} = \sum_{j \in \mathcal{J}_{ft}} (p_{jft} - c_{jft}) \sum_{m \in \mathcal{M}} s_{jfmt}(p_{jft}; \theta, x_{jt}, x_{-jt}, \xi_t) N_{mt} \right\} \tag{2.4}$$

Where c_{jft} are the marginal costs from plan j for firm f and $\boldsymbol{X}_{-jt}$ is the matrix of products characteristics of plans different from plan j. N_{mt} denotes the potential market size in region m at period t.

During the studied period, the number of phones in the market exceed the number of inhabitants in Peru. Nonetheless, the reason behind this observation is that a part of the population has two or even three phones. We take a conservative approach, and assume that N_{mt} is the total number of lines observed in the market.

Finally, the marginal cost c_{jft} is given by:

$$\log(c_{jft}) = \gamma_0 + \boldsymbol{W}_{jft}\gamma_1 + \gamma_2 wages_t + \tau_t + \omega_{jt},$$

where $\boldsymbol{W}_{ft}$ is matrix containing an indicator for incumbents' products in post-entry periods and an indicator for postpaid plans. The variable $wages_t$ captures the average wage in market t and τ_t is a market fixed effect. Finally, ω_{jt} is the structural error term capturing unobserved costs specific to the product, region, and period.

[11] A given plan has the same price across all regional markets.

2.4.2 First stage: Choice of product offerings

At the beginning of the first stage, firms know the distribution of the second stage shocks but not their realizations. Given their expectation of second stage profits, each firm f at period t chooses the set of plans offerings $\mathcal{J}_{ft} \subseteq \mathcal{J}$, where $\mathcal{J}$ corresponds to the set of all possible plan offerings given the minutes and data bins defined in Subsection 2.4.1. To offer a given plan p_{jf}, the firm incurs a fixed cost given by:

$$FC_{jf}(\theta_f) = \kappa_f + \eta_{jf} \tag{2.5}$$

Where k_f is a firm specific fixed cost to be estimated, while η_{jf} is a unobserved cost shock with mean zero.

The first stage profit depends on the expected profit from the chosen plans in the second stage minus the fixed cost of offering them. Defining as $|min|$ the cardinality of the bins for the number of minutes offered and $|gb|$ the cardinality of the set of minutes for the number of gigabytes offered, the decision of the firm is given by the vector $D_f = (d_{f,0}, d_{f,1}, ..., d_{f,J}) \in \{0,1\}^{|min|\times|gb|+1}$. Where $d_{f,0} = 1$ corresponds to the decision of firm f of offering prepaid plans, and $d_{f,j} = 0$ corresponds to the decision of not offering the postpaid plan in space j of the grid of product characteristics. Defining the information state of firm f in the first stage as I_f and using the decision vector d_f we get the first stage expected profit:

$$\mathbb{E}\left[\Pi_{ft}(D_f)|I_f\right] = \mathbb{E}\left[\sum_{j\in\mathcal{J}} \mathbb{1}_{\{d_{f,j}=1\}}\left((p_{jft}-c_{jft})\sum_{m\in\mathcal{M}} s_{jfmt}(p_{jft};\theta,x_{jt},x_{-jt})N_{mt} - FC_{jf}(\kappa_f)\right)|I_f\right]$$

Identification of the fixed cost

The firm chooses which elements to offer within the product characteristics grid. As it can been seen in Equation 2.4, entry of a given $j \in \mathcal{J}$ affects the profits of the firm across all other plans $j \in \mathcal{J}_f$ offered by the firm, through the cannibalization of sales of other plans. Since firms internalize this effect, the maximization problem of the firm becomes a combinatorial problem. Given that there are four possible levels of minutes offered and five levels of gigabytes offered, the action space for each firm has a cardinality of 2^{21}. To overcome this dimensionality issue and the likely existence of multiple equilibria, we employ moment inequalities ((Eizenberg, 2014), (Pakes et al., 2015)).

If the observed plan offerings of each firm, $D^*_{ft} \quad \forall f \in \mathcal{F}$, are an equilibrium of the game, they constitute a best response and thus give to the firm a higher or equal profit than any other action they could have made. Formally, this implies:

$$\mathbb{E}\left[\Pi_{ft}(D^*_{ft}; D^*_{-ft})|I_{ft}\right] - \mathbb{E}\left[\Pi_{ft}(D_{ft}; D^*_{-ft})|I_{ft}\right] \geq 0 \quad \forall \quad D_{ft} \in \{0,1\}^{|min|\times|gb|+1}$$

By comparing the observed plan offerings to appropriate hypothetical offerings, we can create inequalities giving upper and lower bounds to the mean fixed cost parameter θ_f. In particular, applying the above inequality to alternative decisions where offered plan k was not included, D^-_{ft}, would give an upper bound, while comparing the observed actions to an alternative offering where an additional plan k' is offered, D^+_{ft}, would give a lower bound. Defining $VP_{ft}(D_{ft})$ as the variable profit under action D_{ft}, we can write these bounds as:

$$\begin{aligned} &VP_{ft}\left(D^*_{ft}\right) - VP_{ft}\left(D^-_{ft}\right) - \eta_{kft} \geq \kappa_f \\ &\kappa_f \geq VP_{ft}\left(D^+_{ft}\right) - VP_{ft}\left(D^*_{ft}\right) - \eta_{k'ft} \end{aligned} \tag{2.6}$$

However, the bounds in Equation 2.6 depend on the unobservable part of the fixed cost and thus cannot be directly used to infer the cost of the firm. We follow Eizenberg (2014) and build moment inequalities by taking the expectation of the inequalities in Equation 2.6 over two subsets of the firms action spaces. Defining the sets $A^+_{ft} = \left\{k' : d^*_{kft} = 0\right\}$ and $A^-_f = \left\{k : d^*_{kft} = 1\right\}$ which indicate which plans were, or were not, offered by firm f at period t leads to the following moment inequalities:

$$\begin{aligned} &\mathbb{E}\left[VP_{ft}\left(D^*_{ft}\right) - VP_{ft}\left(D^-_{ft}\right) - \eta_{kft}|k \in A^-_f\right] \geq \kappa_f \\ &\kappa_f \geq \mathbb{E}\left[VP_{ft}\left(D^+_{ft}\right) - VP_{ft}\left(D^*_{ft}\right) - \eta_{k'ft}|k' \in A^+_f\right] \end{aligned} \tag{2.7}$$

Which give unbiased bounds as long as $\mathbb{E}\left[\eta_{k'ft}|k' \in A^+_f\right] = 0$ and $\mathbb{E}\left[\eta_{kft}|k \in A^-_f\right] = 0$. However, the definition A^+_{ft} and A^-_{ft} imply that these conditional expectations are affected by a selection bias and thus they are not equal to 0. To overcome this endogeneity issue, we follow Eizenberg (2014) and use as bounds:

$$\mathbb{E}\left[B^L_{jf}(\theta_0)\right] \leq \kappa_f \leq \mathbb{E}\left[B^U_{jf}(\theta_0)\right] \quad j \in \mathcal{J} \tag{2.8}$$

Where $B^L_{jf}(\theta_0)$ and $B^U_{jf}(\theta_0)$ are defined as:

$$B^L_{jf}(\theta_0) = \begin{cases} V^L_f(\theta_0) & j \in A^-_f \\ \tilde{\kappa}^L_{jf}(\theta_0) & j \in A^+_f \end{cases} \qquad B^U_{jf}(\theta_0) = \begin{cases} \tilde{\kappa}^U_{jf}(\theta_0) & j \in A^-_f \\ V^U_f(\theta_0) & j \in A^+_f \end{cases}$$

With $V_f^L(\theta_0) = \min_{j\in\mathcal{J}} \{V_{jf}(\theta_0)\}$ and $V_f^U(\theta_0) = \max_{j\in\mathcal{J}} \{V_{jf}(\theta_0)\}$, where:

$$V_{jf}(\theta_0) = \begin{cases} \tilde{\kappa}_{jf}^U(\theta_0) & j \in A_f^+ \\ \tilde{\kappa}_{jf}^L(\theta_0) & j \in A_f^- \end{cases}$$

With $\tilde{\kappa}_{jf}^U(\theta_0) = VP_{ft}\left(D_{ft}^*\right) - VP_{ft}\left(D_{ft}^-\right)$ and $\tilde{\kappa}_{jf}^L(\theta_0) = VP_{ft}\left(D_{ft}^+\right) - VP_{ft}\left(D_{ft}^*\right)$.

Finally, the empirical counterpart of 2.8 is given by:

$$b_f^L(\hat{\theta}) = \frac{1}{J}\sum_{j\in\mathcal{J}} B_{jf}^L(\hat{\theta}) \qquad b_f^U(\hat{\theta}) = \frac{1}{J}\sum_{j\in\mathcal{J}} B_{jf}^U(\hat{\theta})$$

Where J is the cardinality of $\mathcal{J}$ and $\hat{\theta}$ are the parameters estimated for the second stage of the model. Following Imbens and Manski (2004), the confidence region is given by:

$$\left[b_f^L(\hat{\theta}) - \sqrt{\frac{S_L(\hat{\theta})}{J}} z_{1-\alpha}, \quad b_f^U(\hat{\theta}) + \sqrt{\frac{S_U(\hat{\theta})}{J}} z_{1-\alpha} \right] \tag{2.9}$$

Where $S_L(\hat{\theta})$ and $S_U(\hat{\theta})$ are the empirical variance of $b_f^L(\hat{\theta})$ and $b_f^U(\hat{\theta})$.

2.5 Results

2.5.1 Demand

As discussed in Subsection 2.4.1, our main specification is a nested logit. We also estimate demand under a logit specification and in both cases we compare the results with and without using instruments to control for endogeneity. Table 2.5.1 presents the estimation results for the different specifications discussed above. For both the logit and nested logit models, the usage of instruments heavily impacts estimates. Without instruments price sensitivity is close to zero and many other parameters don't have the expected sign. As well, we can see an increase on the nest parameter, showing costumers are more likely to substitute within their plan category, postpaid and prepaid, rather than across categories.

In the demand specifications we consider the price of the plan divided by the average income in the region in order to account for the heterogeneity across the country in terms of purchasing power and thus accessibility to cellphone services. As expected the price sensitivity is negative and implies an average own price elasticity of -1.87. The low price elasticity could partly be explained by consumer inertia or unobserved switching costs when changing operator.[12] We will try to address such concerns in future versions. We also find a positive impact of the number of minutes and gigabytes on the demand for a plan. Interestingly, the sensitivity to the amount of data included increases quite considerably after the introduction of 4G technology in the country, indicating the importance of the increased internet speeds on consumers' valuation for the service. Similarly, we include the logarithm of the number of 4G antennas installed by each operator in each region. This measure gives of a proxy for the firm-region specific quality of the data service, as expected, it has a positive impact on the demand for a firm's plans.

Table 2.5.1: Demand estimates

	Logit	IV- Logit	Nested Logit	IV-Nested Logit
Price / avg. income	-0.03***	-34.78***	-0.01***	-3.15***
	(0.001)	(0.743)	(0.000)	(0.171)
log(minutes)	0.14***	1.08***	0.02***	0.08***
	(0.007)	(0.023)	(0.002)	(0.005)
Gigabytes (before 4G)	-0.85***	0.55***	-0.21***	0.02**
	(0.025)	(0.060)	(0.004)	(0.008)
Gigabytes (after 4G)	0.04***	2.41***	0.03***	0.24***
	(0.007)	(0.052)	(0.002)	(0.011)
Log(4G antennas)	0.23***	-0.04**	0.06***	0.02***
	(0.008)	(0.020)	(0.002)	(0.002)
σ			0.73***	0.85***
			(0.003)	(0.004)
Mean own price elasticity	-0.0	-3.33	-0.0	-1.87
Observations	63 378	63 378	63 378	63 378
Market Fixed Effects	✓	✓	✓	✓
Year Fixed Effects	✓	✓	✓	✓
Firm × plan type Fixed Effects	✓	✓	✓	✓

*Notes: Standard errors in parentheses. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$*

To further understand the valuation that consumers give to the different characteristics of the offered plans, Table 2.5.2 presents the summary statistics of the distribution of willingness to pay (W.T.P.) for minutes and data across regions and years. Given that the willingness to pay is affected by the average income in a region, we observe large variation between the minimum and maximum willingness to pay for all characteristics. However, we do observe a much smaller dispersion between the 25 and 75 percentiles of the W.T.P. distributions. The dispersion highlights the role of economic disparities between regions on the adoption of the service. Given the nationwide uniform pricing strategy that firms use, larger usage of data is

[12]The Telecommunications Regulator, OSIPTEL, highlights that consumers might prefer to stay with their current plans to avoid searching for additional ones: '*... the subscriber may have a strong inertia or status quo bias to continue in the current contracting dynamic in order to avoid all the costs that were once incurred*'. See https://rb.gy/9liqg1, p. 20.

harder to achieve in lower income regions. This leads to geographic differences in investment for the service as for example in the investment of 4G antennas discussed in subsection 2.3.4. Results indicate that the average W.T.P. for an increase of one gigabyte of data increased by a factor of 14 after the introduction of 4G. Such an importance change in the valuation of a characteristic affects the incentives of the firms and highlights the utility of our structural model both to understand the observed dynamics and for future counterfactual analysis.

Table 2.5.2: Willingness to pay for the different plan's characteristics

	Min.	Q25	Median	Mean	Q75	Max.
W.T.P. for 10 minutes	1.6	2.2	2.6	2.7	2.9	5.9
W.T.P. for 1 GB before 4G	0.4	0.5	0.6	0.6	0.6	1.4
W.T.P. for 1 GB after 4G	5.5	7.2	8.2	8.4	9.1	12.2

Notes: Summary statistics computed across years and regions. Amounts in Peruvian soles. One Peruvian sol corresponds to around 0.27 U.S. dollars.

2.5.2 Fixed cost from product offerings

To estimate the fixed cost of entry for a given plan, we follow the methodology outlined in Subsection 2.4.2. Table 2.5.3 presents the implied fixed cost bounds for each firm under three different methodologies. The first one corresponds to the estimated set using the correction from Eizenberg (2014) presented in equation 2.8. Then, we report its corresponding confidence interval, as shown in equation 2.9. Finally, as a robustness check, we compare our results with the implied bounds when no selection correction is applied to equation 2.7.

Table 2.5.3: Estimated bounds for fixed cost parameter

Firm	Estimated set	95% Confidence interval	No selection correction	
			Lower bound	Upper bound
Telefonica	[6.46, 33.14]	[6.40, 33.31]	6.36	4.28
Claro	[6.36, 34.56]	[6.30, 34.74]	6.23	2.90
Entel	[19.39, 487.05]	[19.33, 487.23]	19.07	93.03
Bitel	[6.28, 48.46]	[6.22, 48.64]	6.23	1.28

Notes: Fixed costs in millions of Peruvian soles. One Peruvian sol corresponds to around 0.27 U.S. dollars. The implied bounds when not correcting for selection bias are not presented as an interval since they cross for most firms.

We find few differences between the estimated set and its corresponding confidence interval. However, we do see a large difference between these values and the bounds obtained when no correction for the selection bias is applied. While the lower bound remains similar, the upper bound is considerably smaller and for most firms it is even smaller than the lower bound, indicating we cannot properly identify the costs under that approach. These results highlight the importance of the selection bias in our sample and why it is important to account for it in order to get the proper fixed cost estimates.

Results from the confidence interval show very similar fixed cost for both incumbents (Movistar and Claro). For Bitel, we find a similar lower bound but an upper bound that is more than ten million soles larger than for the incumbent firms. Surprisingly, for the other entrant, Entel, we find much larger bounds. By looking at the observed plans offered, we see that across years Entel offers fewer plans than any of its rivals and thus the larger fixed cost reflects such behavior. This larger cost also reflects Entel's strategy of large investments in advertising their offerings. [13]

2.6 Conclusion and future steps

Counterfactual analysis

While the current version of the paper does not include any counterfactual analysis using the estimated structural model, we would like to discuss the possible hypothetical scenarios we could analyze and the type of policy recommendations we could get from them.

First, both from the motivating evidence and the structural model we see a clear change in consumer demand patterns after the introduction of 4G connectivity. These changes translate into a higher valuation for data in mobile plans. Meanwhile, the motivating evidence also documents a large increase in the variety of plans and their amount of data after the arrival of the two entrant firms. However, both entry and 4G introduction happened simultaneously, and thus we cannot disentangle their effects on the incumbent's choices. To better understand the separate effect of these two mechanisms, the first counterfactual we could run corresponds to solving the two stage game for all the periods after the 4G introduction (in 2014) without the entry of new firms. The predicted actions of the incumbents in that scenario will help us to understand the role of competitive pressure on product offerings.

The second counterfactual we are interested in pursuing is also related to the structure of the market. Entel and Bitel entered the market focusing on prepaid and postpaid, respectively. The complementarity between these strategies generated externalities among them that are not internalized. We could analyze the entry of a single firm that competes in both segments of the market. The third counterfactual we would investigate is related to the timing of the entrants and municipal permits. Recall that entrants, specially Bitel, had constant delays for obtaining their permits to deploy their infrastructure. We would like to analyze the effects of a one year

[13] Entel's expenses for attracting consumers in their first 2 years increased in 37.7%. Nonetheless, according to Entel's legal manager: *'[Entel] designed a launch strategy, but it cannot be maintained throughout our entire presence in the market'*. See https://rb.gy/b3ta6z.

earlier entry by the Entel and Bitel. This could help understand the role of regulation as a possible barrier to entry and competition in oligopolistic markets. Additionally, notice that a one-year early entry to the market implies entry before the 4G. Indirectly, this counterfactual could help us understand the role of competition before the technological change.

Finally, the descriptive evidence shows diverging patterns in the investment on 4G antennas across firms while the demand estimates show the importance of good coverage within a region (i.e. number of antennas) on consumer preferences. In relation to the third counterfactual discussed above, an interesting extension for the paper would be the inclusion of a dynamic yearly stage where firms choose their investment in each region of the country, impacting the number of antennas available and thus the quality of the service. Such extension would also allow studying in more detail the spatial differences in access to the new technology and to understand if the competitive pressure from the entrants has a different impact across the rural and urban areas.

Conclusion

This paper uses a unique dataset covering the Peruvian mobile telecommunications market during the period of two major events: the introduction of 4G connectivity in the country and the arrival of two low-cost providers. Our motivating evidence shows a large change in consumption patterns and product offerings after the realization of these two events. In particular, we find larger consumption of data after the introduction of 4G and a larger number of plan offerings, covering a much larger set of minutes-gigabytes allowances. We propose a structural model representing a two stage game were firms chose both which mobile plans to offer and their prices. We estimate the model and find that the willingness to pay for gigabytes in mobile plans increased significantly after the introduction of 4G. Furthermore, we also estimate the fixed cost of offering a product, finding reasonable values and heterogeneity in these costs between incumbents and entrants. In future work we plan to use the estimated structural model to analyze counterfactual scenarios that will help us to better understand the role of competition and technological progress in this market.

Bibliography

Backus, M., Conlon, C., and Sinkinson, M. (2021). Common ownership and competition in the ready-to-eat cereal industry. NBER Working Papers 28350, National Bureau of Economic Research, Inc.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, pages 242–262.

Berry, S. T. and Waldfogel, J. (2001). Do mergers increase product variety? evidence from radio broadcasting. *The Quarterly Journal of Economics*, 116(3):1009–1025.

Bontemps, C., Gualdani, C., and Remmy, K. (2023). Price competition and endogenous product choice in networks: evidence from the us airline industry. Technical report, TSE Working Paper.

Bourreau, M. and Sun, Y. (2022). Competition and quality: Evidence from the entry of mobile network service. Technical report.

Bourreau, M., Sun, Y., and Verboven, F. (2021). Market entry, fighting brands, and tacit collusion: Evidence from the french mobile telecommunications market. *American Economic Review*, 111(11):3459–99.

Chatterjee, C., Fan, Y., and Mohapatra, D. P. (2022). Spillover effects in complementary markets: A study of the indian cellphone and wireless service markets. Working papers.

Crawford, G. S. (2012). Endogenous product choice: A progress report. *International Journal of Industrial Organization*, 30(3):315–320.

Crawford, G. S., Lee, R. S., Whinston, M. D., and Yurukoglu, A. (2018). The welfare effects of vertical integration in multichannel television markets. *Econometrica*, 86(3):891–954.

Crawford, G. S. and Yurukoglu, A. (2012). The welfare effects of bundling in multichannel television markets. *American Economic Review*, 102(2):643–85.

Draganska, M., Mazzeo, M., and Seim, K. (2009). Beyond plain vanilla: Modeling joint product assortment and pricing decisions. *QME*, 7:105–146.

Economides, N., Seim, K., and Viard, V. B. (2008). Quantifying the benefits of entry into local phone service. *the RAND Journal of Economics*, 39(3):699–730.

Eizenberg, A. (2014). Upstream innovation and product variety in the us home pc market. *Review of Economic Studies*, 81(3):1003–1045.

Elliott, J., Houngbonon, G. V., Ivaldi, M., and Scott, P. (2023). Market structure, investment and technical efficiencies in mobile telecommunications.

Fan, Y. (2013). Ownership consolidation and product characteristics: A study of the us daily newspaper market. *American Economic Review*, 103(5):1598–1628.

Fan, Y. and Xiao, M. (2015). Competition and subsidies in the deregulated us local telephone industry. *the RAND Journal of Economics*, 46(4):751–776.

Fan, Y. and Yang, C. (2020). Competition, product proliferation, and welfare: A study of the us smartphone market. *American Economic Journal: Microeconomics*, 12(2):99–134.

Fan, Y. and Yang, C. (2022). Estimating discrete games with many firms and many decisions: An application to merger and product variety. Technical report, National Bureau of Economic Research.

Genakos, C., Valletti, T., and Verboven, F. (2018). Evaluating market consolidation in mobile communications. *Economic Policy*, 33(93):45–100.

Granja, J. (2022). Regulation and service provision in dynamic oligopoly: Evidence from mobile telecommunications. *Working paper.*

Imbens, G. W. and Manski, C. F. (2004). Confidence intervals for partially identified parameters. *Econometrica*, 72(6):1845–1857.

Lin, Z., Tang, X., and Xiao, M. (2023). Endogeneity in games with incomplete information: Us cellphone service deployment. *Available at SSRN 3741805.*

Marcoux, M. (2022). Strategic interactions in mobile network investment with a new entrant and unobserved heterogeneity. *International Journal of Industrial Organization*, 82:102829.

McManus, B. (2007). Nonlinear pricing in an oligopoly market: The case of specialty coffee. *The RAND Journal of Economics*, 38(2):512–532.

Miller, N. H., Sheu, G., and Weinberg, M. C. (2021). Oligopolistic price leadership and mergers: The united states beer industry. *American Economic Review*, 111(10):3123–3159.

Miller, N. H. and Weinberg, M. C. (2017). Understanding the price effects of the millercoors joint venture. *Econometrica*, 85(6):1763–1791.

Nosko, C. (2010). Competition and quality choice in the cpu market. *Manuscript, Harvard University.*

Pakes, A., Porter, J., Ho, K., and Ishii, J. (2015). Moment inequalities and their application. *Econometrica*, 83(1):315–334.

Seim, K. and Viard, V. B. (2011). The effect of market structure on cellular technology adoption and pricing. *American economic journal: Microeconomics*, 3(2):221–251.

Starc, A. and Wollmann, T. G. (2022). Does entry remedy collusion? evidence from the generic prescription drug cartel. Technical report, National Bureau of Economic Research.

Wollmann, T. G. (2018). Trucks without bailouts: Equilibrium product characteristics for commercial vehicles. *American Economic Review*, 108(6):1364–1406.

Chapter 3

Measuring the effects of leniency programs under undetected cartels

Oscar Jara[1]

Abstract:

This paper reviews the different approaches used in the literature to measure the efficacy of leniency programs in detecting and deterring cartels. Focusing on the incomplete data typical of the hidden cartel activities, I propose using a Hidden Markov Model to model the cartel formation and detection and asses the effects this program. Finally, I propose using this approach to measure the effects of the introduction of leniency programs in Brazil, Chile, Colombia, Mexico, and Peru.

JEL Classification: L13, L41, K21, K42

[1]Toulouse School of Economics, Université Toulouse 1 Capitole.

3.1 Introduction

Due to their secretive nature, prosecuting cartels continues to be a significant challenge for antitrust authorities worldwide. To combat the hidden activities of cartels, many countries offer incentives to cartel members to come forward with evidence about their activities. These strategies are known as leniency programs, which not only aim to destabilize existing cartels but also to deter the formation of new ones. Its efficacy, nonetheless, remains unclear because the full population of cartels is not observed. In this paper, I explore how previous papers assessed the effect of the leniency programs on the detection and deterrence of cartels and effects of the leniency programs and propose a new methodology that considers only the sample of observed cartels.

I review a count data and selection correction model. The count data model captures the effects of the leniency program on cartel detection, but omits the effects of self selection of cartels into the program. On the other hand, the Heckman selection model corrects for this bias. However, this does not address weather the sample of observed cartels is representative from the population of cartels. When working with the discovered cartels, it is not possible to make inference about the unobserved cartels. This happens because it is not possible to ensure that detected cartels are representative of the not discovered ones. To address this issue I propose using a Hidden Markov Model.

The main contribution of this paper is to propose using a new methodology to the estimation of the effects of the leniency programs. Hidden Markov Models (Baum and Petrie, 1966; Baum et al., 1970) identify unobservable *hidden* or latent states by analyzing observable data that indirectly indicates underlying processes. The usefulness of Hidden Markov Models (HMMs) to deal with incomplete datasets where only partial observations exist has led them to be used in economics (Hamilton, 2010), finance (Mamon and Elliott, 2007) or marketing (Poulsen, 1990; Park and Gupta, 2011; Netzer et al., 2008). In the context of antitrust enforcement and the study of cartels, where the clandestine nature of collusion often leads to limited and incomplete data, a Hidden Markov Model offers a useful framework. By modeling the relationship between observed indicators, such as cartel activity, and latent states representing different states of market behavior (e.g., cartelized or competitive), HMM allows uncovering hidden patterns, even with sparse or imperfect data. This methodology enables the identification of cartel's formation and observation probabilities, themes that have not been handled the literature yet.

Leniency programs are crucial tools for competition authorities, not only for deterring cartels by disrupting internal trust and destabilizing agreements, but also for facilitating prosecution by bringing anticompetitive behavior to light. Despite its positive effects, there has been a significant decline in leniency applications worldwide between 2015 and 2021, with OECD jurisdictions experiencing a 58% drop.[2] The decrease in leniency applications raises concerns about cartel enforcement effectiveness, as competition authorities heavily rely on these applications. This trend prompts questions about the potential threats posed by the decline in cartel enforcement.

Finally, this paper also contributes to the policy debate about the efficacy of the enforcement measures against cartels. This is specially relevant for Latin America. As the World Bank states, *...empirical data suggests that, in advanced economies, there are between three and 10 times as many cartel agreements as competition authorities manage to detect. It is likely that the number of cartels that go undetected is even greater in Latin America.*[3] I propose using a HMM to analysis the introduction of the leniency programs in Brazil, Chile, Colombia, Mexico and Peru. I take advantage of the introduction of leniency programs around the year 2010 for the these countries to evaluate its efficacy.

By showing the positive effects on deterring cartels, this paper can bring evidence to policymakers of not releasing laws that conflict with the leniency programs. For example, unlike the US where civil and criminal proceedings are under the same roof of the Department of Justice, other jurisdictions that pursue cartels as criminal activities can put under risk the immunity benefits of the leniency programs. This was the case of Peru in 2021, were an enacted law allowed to criminally process individuals that participated of a cartel. This reduced the incentives from whistleblowing about cartel activities. It was until June 2023 that cartel whistleblowers were protected from criminal proceedings.[4]

In the next section starts with a review of the literature, which is followed by an examination of the regulatory framework surrounding leniency programs in Latin American countries. This provides an overview of the legal structures governing anticompetitive practices and outlines the data to be employed in future versions of this paper. Following this, section 4 presents three distinct models to capture the effects' leniency programs and propose its application to Latin American countries. Finally, section 5 concludes with overarching implications for policy, enforcement, and future research in competition law and regulation.

3.2 Literature Review

First, this paper is related to the literature on the effects of leniency programs over cartel deterrence (Motta and Polo, 2003; Spagnolo, 2004; Aubert et al., 2006; Harrington Jr, 2008, 2013; Chen and Rey, 2013). The challenge in measuring the effects of the leniency program

[2] OECD (2023), p. 6.
[3] See https://rb.gy/ivb5gg. Last time seen it: April 14, 2024.
[4] Law No. 31775. See https://rb.gy/ezvzo7. Last time seen it: April 25, 2024.

arises from the potential sample bias arising from only observing cartels detected by antitrust authorities (Harrington Jr et al., 2006; Harrington Jr and Chang, 2009; Miller, 2009; Brenner, 2009; Ilzkovitz and Dierx, 2014). Harrington Jr and Chang (2009) show that, following the introduction of a new policy, such as a leniency program, a decrease in the duration of detected cartels can be informative of a decrease in the latent rate of cartel formation, because the observed cartels come from a sample of longer-lasting cartels.

Previous empirical papers have tested the theoretical results that are consistent with an effective leniency program. Miller (2009) examines a Markov transition process for cartels, finding an immediate increase in detected cartel cases following the DOJ's 1993 amendments to the leniency program. Subsequently, the number of detections readjusts below pre-leniency levels, indicating an increased detection rate and decreased formation rate. Brenner (2009) and De (2010) evaluate the impacts of the EC's leniency programs and test Harrington Jr and Chang (2009)'s theory on the EC's 1996 and 2002 leniency program changes, showing ambiguous results of the efficacy. Bigoni et al. (2012) find that leniency programs have both a deterrent and a stabilizing effect on cartels. I add to these papers that rather than testing predictions from theoretical models I account for the sample bias when studying the effects of the leniency programs on cartel detection and formation.

The theoretical literature has studied the process of birth and death of cartels (Harrington Jr and Chang, 2009; Harrington Jr and Wei, 2017). Empirically, Hyytinen et al. (2018) and Marvão et al. (2021) use a hidden Markov model (HMM) to model (legal) cartels births and death across the panel of Finish manufacturing firms. Jointly, both papers analyze the prevalence and longevity of 365 legal cartels from 1951 to 1990. They focus on characterizing the hidden cartel dynamics and analyze the prevalence of cartels as new antitrust regulation is introduced.

I follow these papers and use a HMM to model the transition from competition to collusion, and back, in different markets when the true state of the market is not observed. I study a setting where cartels are illegal such that observed cartels are the ones detected by the antitrust authority. The HMM allows handling the sample selection bias by accounting for the unobserved transitions in the states. The introduction of leniency programs in Latin America allows testing for the changes in time in the transition probabilities between competition and cartels. Even without the full population of cartels, the HMM enables to assess for the effects of leniency programs on cartel formation and detection.

3.3 Regulatory Framework and Data

Leniency programs provide incentives for firms engaged in anticompetitive behavior, such as cartel formation, to self-report their activities in exchange for immunity or reduced penalties.

Leniency programs are considered as powerful tools for detecting and deterring anticompetitive practices. This paper proposes applying a Hidden Markov Model to study the effects of the introduction of leniency laws in five key Latin American economies: Brazil, Chile, Colombia, Mexico, and Peru. Each country has enacted leniency legislation as part of broader efforts to combat cartel activity and promote competition within their respective markets.

Brazil: Although Brazil's leniency program started in 1994, it underwent significant changes in 2011.[5] The revised leniency program introduced several crucial provisions aimed at incentivizing cartel participants to come forward with information and evidence of anticompetitive behavior. The change in the structure of the program was oriented to align with global best practices in antitrust enforcement Furthermore, the amended legislation endowed Administrative Council for Economic Defense (CADE) with enhanced investigative powers and procedural mechanisms to expedite the resolution of cartel cases. These reforms sought to streamline the leniency application process and improve coordination between enforcement agencies.

Chile: Chile's leniency program can be traced back to February 16, 2004, when significant reforms were introduced to Chile's competition law landscape.[6] Under this legislation, the National Economic Prosecutor's Office (FNE) was entrusted with the authority to investigate and prosecute anticompetitive practices, including cartel behavior. On March 14, 2011, specific provisions related to the leniency in were further refined.[7] The provisions established detailed guidelines and procedures for leniency applicants, outlining the conditions under which immunity or reduced penalties would be granted in exchange for cooperation with the Antitrust Authority.

Colombia: Initially introduced in 2009, the leniency program was reformed in 2015 to enhance cartel detection efficiency, including shortening the application timeframe for whistleblowers and offering new sentence reductions.[8] These reforms led to increased applications and agreements, resulting in the exposure of major cartels in various sectors. However, challenges arose, requiring further reforms. Cases like 'Chlorine-Soda' and cartel activities in the mining industry highlighted issues with whistleblower compliance. Despite initial success, the Antitrust Authority noted a decline in leniency program applications after 2015, attributing it to insufficient incentives and increased disincentives. Recent legislative reforms in 2022 expanded the leniency program benefits to cover broader anticompetitive practices and modified benefits for whistleblowers. These changes aim to promote non-cooperative responses from cartel participants and increase leniency program utilization.

[5] It was established under Law No. 8884/94 in 1994 and was reformed in 2011 by Law No. 12529/11.

[6] Law No. 19911

[7] Supreme Decree No. 21

[8] Introduced through Act No. 1340 and regulated by Decree No. 2896 in 2010. Reformed in 2015 by Decree No. 1523

Mexico: The leniency program was introduced in 2006, thourgh an amendment to Mexico's inaugural competition law.[9] This reform sought to align Mexico's competition enforcement practices with global standards by introducing a mechanism akin to leniency programs observed in other jurisdictions. In 2010 the first leniency program guide was published, aiming to improve clarity on application procedures.[10] Further reforms in 2011 introduced robust penalties for collusion, including criminal sanctions and income-based fines. This increased the number of applications to the program. In 2013 the Federal Economic Competition Commission (COFECE) was created autonomous entity. This constitutional reform expanded the powers to the Antitrust Authority, including the authority to conduct surprise verification visits, but also reinforced penalties for cartel conduct. The most recent phase of the program's reforms spanned from 2019 to 2021, culminating with a new leniency program guide.

Peru: The original leniency program in Peru was established within the Law for the Repression of Anti-competitive Behavior of 2008.[11] In September 2015, the firs significant reform was carried aiming at giving further clarity to the benefits of participating in the program.[12] The Leniency Program Guidelines were issued in August 2017, providing detailed instructions and requirements for leniency applicants. These guidelines aimed to ensure transparency, predictability, and reliability in the leniency application process.[13] Nonetheless, Act No. 31040 enacted in August 2020 introduced criminal sanctions for anticompetitive behavior without aligning incentives for cooperation in criminal proceedings with those in administrative proceedings, like the leniency program. This discrepancy led to an immediate decline in leniency applications, prompting the need for legislative reform to restore confidence and incentivize cooperation.

Table 1 shows the details for leniency programs in Peru, Chile, Colombia and Mexico.

3.4 The Model

Measuring the effects of the leniency programs is not free of challenges. Since most countries prosecute cartel formation, observed cartels are the ones detected by antitrust authorities (AAs). This sample of *observed* cartels is subject to two potential biases: detection and selection. The former one refers to the fact the only a subset of cartels observed are the ones detected by the AA. Detected cartels could be systematically different from those that manage

[9] Under the 'Ley Federal de Competencia Económica'(LFCE), Article 39

[10] Reform to the LFCE, Article 28

[11] Specifically ruled by Legislative Decree No. 1034

[12] Reformed by Legislative Decree No. 1205, focusing on Article 26.

[13] Additional amendments to the Law for the Repression of Anti-competitive Behavior where introduced in 2018 and 2021. This was done by Legislative Decree No. 1396 of 2018 and Act No. 31112 of 2021

Table 3.3.1: Characteristics of Leniency Programs in Latin America

Characteristics	Peru	Chile	Colombia	Mexico
Level of Sanction Reduction	1st applicant: up to 100% 2nd applicant: 30% to 50% 3rd applicant: 20% to 30% Subsequent: up to 20%	1st applicant: up to 100% Subsequent: up to 50%	1st applicant: up to 100% 2nd applicant: 30% to 50% Subsequent: up to 25%	1st applicant: up to 100% if submitted before investigation starts 1/3 to 2/3 if submitted after started
Quality of Information	Evidence proving cartel's existence	Evidence proving cartel's existence, or information allowing for requests such as surprise inspections, wiretaps, etc.	Complete information about conduct and evidence proving its existence	Complete information about conduct
Obligation to Submit Written Application	No	Yes	No	No
Benefits in Criminal Liability Cases	No criminal sanctions	Criminal sanctions exist. Leniency provides immunity in criminal proceedings for applicants meeting required terms	Only criminal sanctions for collusion in bids. Leniency can only reduce penalty by 1/3. However, if applicant negotiates with the Prosecutor, total immunity possible	Criminal sanctions exist. Leniency provides immunity in criminal proceedings for applicants meeting required terms
Benefits in Damages and Losses	No exemption or reduction for damages claimed in civil proceedings initiated by affected consumers	No exemption or reduction for damages claimed in civil proceedings initiated by affected consumers	No exemption or reduction for damages claimed in civil proceedings initiated by affected consumers	No exemption or reduction for damages claimed in civil proceedings initiated by affected consumers
Benefits Available for Ringleader	Ringleader prohibited from full immunity, but can obtain reductions in sanctions	Prohibits ringleader, cartel motivator, or market leader from benefiting from leniency	Prohibits ringleader, cartel motivator, or market leader from benefiting from leniency	Ringleader can obtain benefits similarly to other participants
Availability of Markers	Yes	Yes	Yes	Yes
Mandatory Termination of Cartel Participation	Applicant must cease participation unless instructed otherwise by the agency	Applicant must cease participation in cartel to apply for leniency	Only when ordered by the Competition Authority	Applicant must cease participation unless instructed otherwise by the agency
Confidentiality Rules	Rules protecting applicant's identity. Applicant's information or evidence considered confidential	Rules protecting applicant's identity, trade secrets, and strategic information, but not evidence presented by applicant	Rules protecting applicant's identity	Rules protecting applicant's identity. Applicant's information or evidence considered confidential

to stay undetected, leading to an over representation of them in the observed data. Regarding the selection bias, cartels that choose to participate in the leniency program may also differ systematically from those that do not. For example, long-lasting cartels or cartels facing higher enforcement risks may be more likely to apply for leniency, biasing estimates of cartel characteristics or dynamics.[14] In this section I discuss three different models to assess the effects of leniency programs on cartel formation while dealing with these issues.

First I discuss a count model, as in Miller (2009). These models are suitable for modeling events where the counts are non-negative integers. The dependent variable is assumed to have a Negative Binomial distribution, with the mean parameterized by a set of explanatory variables. Second, I discuss a Heckman correction method (Heckman, 1979) to address the sample selection bias. This bias arises when the sample used for estimation is not representative of the population of interest due to non-random selection. The Heckman correction involves a two-step estimation process, where the fist step estimates the probability of selection into the sample.

The third method I discuss is the Hidden Markov Model (Baum and Petrie, 1966; Baum et al., 1970), a useful method for analyzing scenarios with incomplete datasets, where direct observation of the underlying states is not possible. These *hidden*, or *latent*, states are inferred using observable data that typically only provides indirect insights. The model allows for transitioning between these states with certain probabilities. At each state, there is a probability

[14] It is possible to also consider a reporting bias. Even among detected cartels, there may be reporting bias if not all relevant information is disclosed or if the accuracy of reported data varies. This can lead to biased estimates of cartel characteristics, such as duration or size. In this paper I do not consider take this specific type.

distribution associated with the observable outputs, effectively making the observed data a probabilistic function of the hidden state.

The most recent literature on the topic of cartel dynamics Hyytinen et al. (2018); Marvão et al. (2021) has chosen Hidden Markov Models (HMMs) to study cartels' birth and death. In a setting where the 'cartelized' state of the economy cannot be observed, the HMM becomes useful. By estimating how the probability of transitioning from competition to cartel changes after the introduction of the leniency program, the enables the measurement of its effects without the issue

3.4.1 Count Data and Selection Models

Count Data: Poisson or negative binomial models are appropriate for count data, such as the number of cartel discoveries or indictments over time, which variations in time are common outputs of leniency programs (Miller, 2009; Marvão et al., 2021). Given the over dispersion of the data, I prefer the use of a negative binomial model. The number of cartel discoveries, $Y_t, \sim NegBin(\mathbb{E}\, Y_t, \alpha)$ and α is the over dispersion parameter.

$$\log(\mathbb{E}\, Y_t) = \beta_0 + \beta_1 L_t + \boldsymbol{X}_t + \tau_t + v_t, \tag{3.1}$$

where L_t is an indicator for leniency at time t, $\boldsymbol{X}_t$ is a matrix of economic conditions and τ_t is a time trend. The parameter β_1 captures the effect of the leniency programs on cartel discoveries. Nonetheless, this time series of detected cartels does not consider that some cartels apply to the leniency program, and they might be significantly different from those who do not. To account for the self-selection, I explore a Heckman Selection Model.

Heckman Selection Model: This model involves a two-step estimation process. The first step estimates the selection equation, and the second step estimates the outcome equation while correcting for sample selection. The Heckman correction involves including the inverse Mills ratio, obtained in the first step, to the outcome equation to address sample selection bias. The key assumption is the exclusion restriction, which states that the instrument used in the selection equation does not directly affect the outcome equation.

Define L_{it} is an indicator for leniency in market i at time t. On the other hand, L^*_{it} is a latent variable representing the probability of a cartel being subject to the leniency program. The selection equation can be expressed as,

$$L^*_{it} = \pi_0 + \pi_1 X_{it} + \pi_2 I_{it} + \xi_{it}, \tag{3.2}$$

where $L_{it} = \mathbb{1}\{L^*_{it} > 0\}$. The matrix of controls X_{it} represents observed characteristics that influence the probability of being subject to the leniency program such as concentration levels, number of participants of the cartel, type of infringement, export, import or trade restriction, and whether the limit of 10% firm's turnover applied to the case. Finally, I_{it} is an instrument that influences the probability of a cartel's decision to enter the leniency program but does not directly affect the number of cartel detections, except through this decision. Good candidates could be sudden changes in commodity prices or significant shifts in demand within markets, which could influence cartel behavior and the propensity to seek leniency. These shocks might motivate cartels to dissolve or seek leniency due to altered economic conditions.

With the results from the Heckman selection equation, it is possible to quantify the effects of the leniency programs on the number of observed cartels, Y_{it}.

$$\log(\mathbb{E}\, Y_{it}) = \beta_0 + \beta_1 L_{it} + \beta_2 \mathbb{1}\{t > t^*\} + \beta_3 L_{it} \times \mathbb{1}\{t > t^*\} + \varepsilon_{it} \tag{3.3}$$

$$\varepsilon_{it} = \gamma\rho(L^*_{it}) + \eta_{it}, \tag{3.4}$$

where ε_{it} is the error term in the outcome equation, and $\gamma\rho(L^*_{it})$ is the correlation coefficient between the error term in the selection equation (ξ_{it}) and the error term in the outcome equation, η_{it}. The inverse Mills ratio, $\rho(L^*_{it})$, is calculated based on the estimated parameters from the selection equation.

As a conclusion, the count data model captures the effects of the leniency program on cartel detection, but omits the effects of self selection into the program. On the other hand, the Heckman selection model corrects for this bias. However, this does not address weather the sample of observed cartels sample of observed cartels is representative from the population of cartels. To solve for this problem, I introduce a Hidden Markov Model.

3.4.2 Hidden Markov Model

Every period the true state of market i can be either 'collusion', c or 'no collusion' n. The binary latent variable Z_{it} represents the unobserved state of market i at time t, where $Z_{it} = 1$ indicates the presence of a cartel and $Z_{it} = 0$ indicates 'no collusion' is the true state. The hidden states $\{Z_{it}\}_{t=1}^{T}$ follow a Markov chain. Similarly, let us introduce a binary indicator variable Y_{it} to represent the observed state of the market at time t. Here, $Y_{it} = 1$ denotes the detection of a cartel, while $Y_{it} = 0$ indicates no detection. Note that no detection means that the state could either be non-collusive or competitive.

In a hidden Markov model, the evolution in time of the hidden state variable Z_{it} is governed by transition probabilities. The transition matrix A encapsulates the probabilities of transitioning from one state to another. Each element a_{it}^{jk} of the matrix represents the probability of transitioning from state j to state k. For example, a_{it}^{nc} is the probability of transitioning from the absence of a cartel ($Z_{i,t-1} = 0$) to the presence of a cartel ($Z_{it} = 1$).

$$A_{it} = \begin{bmatrix} a_{it}^{nn} & a_{it}^{nc} \\ a_{it}^{cn} & a_{it}^{cc} \end{bmatrix} = \begin{bmatrix} 1 - h_{1t} & h_{1t} \\ 1 - h_{2t} & h_{2t} \end{bmatrix} \tag{3.5}$$

Similarly, I introduce a binary indicator variable Y_{it} to represent the observed state of the market at time t, where $Y_{it} = 1$ denotes that a cartel is observed, and $Y_{it} = 0$ indicates no detection. For industry i at time t, the probability of observing state y, given the true hidden state is z, as $P(Y_{it} = y_{it}|Z_{it} = z_{it}) = \beta^y(z)$. In particular, the probability of observing a cartel when the true state is collusion is β_t^c and the probability of observing competition given that competition is the true state is β_t^n. Considering these probabilities, the emission matrix B provides the probabilities of observing each possible state from each possible hidden state in the model.

$$B_{it} = \begin{bmatrix} b_{it}^n(n) & b_{it}^n(c) \\ b_{it}^c(n) & b_{it}^c(c) \end{bmatrix} = \begin{bmatrix} P(Y_{it} = 0|Z_{it} = 0) & P(Y_{it} = 0|Z_{it} = 1) \\ P(Y_{it} = 1|Z_{it} = 0) & P(Y_{it} = 1|Z_{it} = 1) \end{bmatrix} = \begin{bmatrix} \beta_t^n & 1 - \beta_t^n \\ 1 - \beta_t^c & \beta_t^c \end{bmatrix} \tag{3.6}$$

Considering both the transition and emission matrixes, the likelihood function can be written as,

$$L(\boldsymbol{h_{1t}}, \boldsymbol{h_{2t}}, \boldsymbol{\beta^n}, \boldsymbol{\beta^c}; \boldsymbol{Y}) = \Pi_{i=1}^{M} \{D1 (\Pi_{t=1}^{T} D2_{it}) \mathbf{1}\}, \tag{3.7}$$

where $D2$ is a 2×2 matrix, which elements are $d_{it}^{jk}(Y_{it}) = a_{it}^{jk} b_{it}^k(Y_{it})$, $\mathbf{1}$ is a 2×1 vector of ones and $D1$ is a 2×1 matrix of initial conditions with terms $d_{i1}^k(Y_{it}) = \tau_i^k b_{i1}^k(Y_{it})$. The initial distribution of Z_{it} is assumed, where the probability of market i is in the unobserved state k in the initial period is:

$$\tau_i^k = P(Z_{i1} = k) \tag{3.8}$$

Note that the value of $d_{it}^{jk}(Y_{it})$ depends on the value of the observed state. For example if the observed state is competition, then the upper right value of matrix $D2$ is $d_{it}^{nc}(Y_{it} = 0) = a_{it}^{nc} b_{it}^c(Y_{it} = 0) = h_{1t} b_{it}^c(n) = h_{1t}(1 - \beta_t^c)$.

As discussed before, one of the main challenges in the literature related to the leniency programs is the measurement of its efficacy under partial observation of the population of cartels. Using an HMM allows to circumvent this problem, and estimate the transition probability from competition to collusion, h_{1t}; i.e., the probability of cartel formation in an economy. Nonetheless, notice that even if the HMM manages to avoid the detection and selection biases, the conclusions from the model are in average for the whole economy and not for specific markets. This is useful for assessing the effects of the leniency programs. Nonetheless, this avoids making conclusions about individual markets. For example, it is not possible to make further predictions about specific industries, even if these have history about being cartelized.

On the other hand, the estimation of β_t^c gives the probability of observing the cartels. After the introduction of a leniency program it is possible to expect an increase in the probability of observing a cartel. This includes cartels observed due to leniency programs and those detected by the antitrust authority. Then, observing a cartel is different from detecting it. Following Miller (2009), it is possible to disentangle the probability of transitioning from cartel to competition, $1 - h_{2t}$ into detection probability and probability of naturally dropping the cartel. Nonetheless, data on the full sample of cartels would be required to estimate these probabilities. Making this distinction becomes specially relevant considering that when cartels apply to the leniency program they free resources from the Antitrust Authority that can be used to detect unobserved cartels. Similarly, institutional changes can affect positively the detection abilities of the authorities.

3.4.3 Application to Latin America

As it was reviewed in section 3.3, countries in Latin America adopted leniency programs adopted o re-introduced leniency programs at 2009 (Colombia), 2011 (Brazil, Chile, Mexico) or 2015 (Peru). Currently, there are no papers studying at great scale the effects of leniency programs in Latin American countries.

Due to the lack of a unified dataset about the cartel cases in Latin America, there is no easy access to information for researches interested in cartel related topics. I am currently requiring access to the Antitrust Authorities of the countries analyzed in section 3.3. I plan to apply the HMM setting in this dataset to study the effects of the leniency programs on cartel formation and detection.

3.5 Conclusions and Next Steps

In conclusion, this paper highlights the importance of leniency programs in antitrust enforcement, particularly in a setting with declining applications and challenges in detecting cartel activities. By proposing the use of a Hidden Markov Model (HMM), this paper introduces a novel methodology to assess the impact of these programs more accurately. This approach deals with two critical problems when dealing with data that comes only from the pool of detected cartels: detection and selection bias. The HMM models the probability of a market transitioning from competition to cartel, avoiding the issue of assuming that observed cartels are representative of unobserved cartels. Additionally, the HMM allows the estimation of the probability of observing a cartel.

Finally, I propose using the HMM to the case of the introduction of leniency programs in Latin America, where these programs were re-introduced around the year 2010. The insights derived from the application of HMMs to the leniency programs in various Latin American countries could provide guidance for policymakers. As the international landscape of antitrust law continues to evolve, this research could influence legislative and regulatory decisions, ensuring that leniency programs remain a robust tool in the fight against cartels.

Bibliography

Aubert, C., Rey, P., and Kovacic, W. E. (2006). The impact of leniency and whistle-blowing programs on cartels. *International Journal of industrial organization*, 24(6):1241–1266.

Baum, L. E. and Petrie, T. (1966). Statistical inference for probabilistic functions of finite state markov chains. *The annals of mathematical statistics*, 37(6):1554–1563.

Baum, L. E., Petrie, T., Soules, G., and Weiss, N. (1970). A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains. *The annals of mathematical statistics*, 41(1):164–171.

Bigoni, M., Fridolfsson, S.-O., Le Coq, C., and Spagnolo, G. (2012). Fines, leniency, and rewards in antitrust. *The RAND Journal of Economics*, 43(2):368–390.

Brenner, S. (2009). An empirical study of the european corporate leniency program. *International Journal of Industrial Organization*, 27(6):639–645.

Chen, Z. and Rey, P. (2013). On the design of leniency programs. *The Journal of Law and Economics*, 56(4):917–957.

Hamilton, J. D. (2010). Regime switching models. In *Macroeconometrics and time series analysis*, pages 202–209. Springer.

Harrington Jr, J. E. (2008). Optimal corporate leniency programs. *The Journal of Industrial Economics*, 56(2):215–246.

Harrington Jr, J. E. (2013). Corporate leniency programs when firms have private information: the push of prosecution and the pull of pre-emption. *The Journal of Industrial Economics*, 61(1):1–27.

Harrington Jr, J. E. and Chang, M.-H. (2009). Modeling the birth and death of cartels with an application to evaluating competition policy. *Journal of the European Economic Association*, 7(6):1400–1435.

Harrington Jr, J. E. et al. (2006). How do cartels operate? *Foundations and Trends® in Microeconomics*, 2(1):1–105.

Harrington Jr, J. E. and Wei, Y. (2017). What can the duration of discovered cartels tell us about the duration of all cartels? *The Economic Journal*, 127(604):1977–2005.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, pages 153–161.

Hyytinen, A., Steen, F., and Toivanen, O. (2018). Cartels uncovered. *American Economic Journal: Microeconomics*, 10(4):190–222.

Ilzkovitz, F. and Dierx, A. (2014). *Ex-post economic evaluation of competition policy enforcement: A review of the literature*. Publications Office of the European Union.

Mamon, R. S. and Elliott, R. J. (2007). *Hidden Markov models in finance*, volume 4. Springer.

Marvão, C. M., Le Coq, C., and Forsbacka, T. (2021). Cartel birth and death dynamics: empirical evidence. *Available at SSRN 3873351*.

Miller, N. H. (2009). Strategic leniency and cartel enforcement. *American Economic Review*, 99(3):750–768.

Motta, M. and Polo, M. (2003). Leniency programs and cartel prosecution. *International journal of industrial organization*, 21(3):347–379.

Netzer, O., Lattin, J. M., and Srinivasan, V. (2008). A hidden markov model of customer relationship dynamics. *Marketing science*, 27(2):185–204.

OECD (2023). The future of effective leniency programmes: Advancing detection and deterrence of cartels. OECD Competition Policy Roundtable Background Note.

Park, S. and Gupta, S. (2011). A regime-switching model of cyclical category buying. *Marketing Science*, 30(3):469–480.

Poulsen, C. S. (1990). Mixed markov and latent markov modelling applied to brand choice behaviour. *International Journal of Research in Marketing*, 7(1):5–19.

Spagnolo, G. (2004). Divide et impera: Optimal leniency programs. *Available at SSRN 716143*.