We empirically investigate the welfare implications of intermediaries in oligopolistic markets, where intermediaries offer additional services to differentiate their products from the ones of the manufacturers. Our identification strategy exploits the unique circumstance that, in the outdoors advertising industry, there are two distribution channels: consumers can purchase the product either directly from manufacturers, or through intermediaries. We specify a differentiated products' equilibrium model, and estimate it using product-level data for the whole industry.

On the demand side, the model includes consumers who engage in costly search with preferences that are specific to the distribution channel. On the supply side, the model includes two competing distribution channels. One features two layers of activity, where manufacturers and intermediaries bargain over wholesale prices, and intermediaries compete on final prices to consumers. The other is vertically integrated. The estimated model is used to simulate counterfactual scenarios, where intermediaries do not offer additional services. We find that the presence of intermediaries increases welfare because the value of their services outweighs the additional margin charged.

Donna, Javier D. and Pereira, Pedro and Pires, Tiago and Trindade, Andre

The Ohio State University, Autoridade da Concorrencia, University of North Carolina, Chapel Hill, FGV EPGE Brazilian School of Economics and Finance

July 2018

Measuring the Welfare of Intermediation in Vertical Markets

Javier D. Donna Pedro Pereira Tiago Pires Andre Trindade*

First version: March, 2016.
This version: July 18, 2018.

The latest version of the paper is available here.

Abstract

We empirically investigate the welfare implications of intermediaries in oligopolistic markets, where intermediaries offer additional services to differentiate their products from the ones of the manufacturers. Our identification strategy exploits the unique circumstance that, in the outdoors advertising industry, there are two distribution channels: consumers can purchase the product either directly from manufacturers, or through intermediaries. We specify a differentiated products' equilibrium model, and estimate it using product-level data for the whole industry. On the demand side, the model includes consumers who engage in costly search with preferences that are specific to the distribution channel. On the supply side, the model includes two competing distribution channels. One features two layers of activity, where manufacturers and intermediaries bargain over wholesale prices, and intermediaries compete on final prices to consumers. The other is vertically integrated. The estimated model is used to simulate counterfactual scenarios, where intermediaries do not offer additional services. We find that the presence of intermediaries increases welfare because the value of their services outweighs the additional margin charged.

JEL Codes: L81; L42; D83; M37.

Keywords: Intermediaries, vertical markets, search frictions, bargaining, outdoor advertising

*Donna: The Ohio State University, donna.1@osu.edu. Pereira: Autoridade da Concorrência and CEFAGE, pedro.br.pereira@gmail.com. Pires: Department of Economics, University of North Carolina, tpires@email.unc.edu. Trindade: FGV EPGE Brazilian School of Economics and Finance, andre.trindade@fgv.br. Discussions with Jason Blevins, Chris Conlon, Daniel Ershov, Paul Grieco, José Luis Moraga-González, Aviv Nevo, Mathias Reynaert, Tiago Ribeiro, Bruce Weinberg, Kevin Williams, Renan Yoshida, as well as seminar participants at UCLA Anderson, Toulouse, FGV/EPGE, Nova SBE, U.Porto, Ohio State, CUNEF, the Barcelona GSE Summer Forum (Consumer Search and Switching Costs Workshop), the 15th and 16th IIOC, the 12th CRESSE conference, the 11th PEJ annual meeting, the 2018 SED Meeting, the 2017 NASAM of the Econometric Society, the 2018 GEE-Nova SBE conference, and the 2017 Triangle Microeconomics Conference in honor to Tiago Pires have greatly benefited this work. The opinions expressed in this article reflect only the authors' views and in no way bind the institutions to which they are affiliated. We gratefully acknowledge the allocation of computing time from the Ohio Supercomputer Center (grant PAS1350-2). P. Pereira is pleased to acknowledge financial support from Fundação para a Ciência e Tecnologia and FEDER/COMPETE (grant UID/ECO/04007/2013). The latest version of the paper is available at: https://www.dropbox.com/s/zy9jwxxl1u2g4uz/DPPT-Intermediation-Welfare.pdf?dl=1. All errors are our own.

1 Introduction

Intermediaries play an important role in contemporary economies. For example, in the U.S. they represent over a third of the value added to the economy.[1] They provide a wide variety of services to the consumers. Intermediaries often add value by transforming products (*e.g.* adding transportation, packaging, or assembling services), or by providing information and consulting services about the characteristics of the products sold by the manufacturers (Spulber 1996). Intermediaries also improve welfare to consumers by reducing search frictions, thus improving the coordination of the actions of consumers and manufacturers (Rubinstein and Wolinsky 1987). When negotiating with a manufacturer, intermediaries centralize transactions of multiple consumers, thus supplanting consumers' decentralized bargaining with the manufacturer (Spulber 1996). The resulting increase in intermediaries' bargaining power, translates in lower marginal costs for the intermediaries, which results in lower prices to the consumers. In the absence of market power, intermediaries improve consumers' welfare when they provide these additional services (see Spulber 1999 for a thorough analysis). When market power is present, however, intermediaries may also lead to *double marginalization*, whereby the product is marked up above the marginal cost of production twice, once by the manufacturer and once again by the intermediary.[2] In such a cases, intermediaries may reduce welfare. A natural question arises: What are the welfare implications of intermediaries in oligopolistic markets where intermediaries offer additional services to differentiate their products from the ones of the manufacturers?

There is a vast literature studying the role of intermediaries in different markets.[3] Two major explanations why intermediaries arise are to facilitate matching of buyers and sellers,

[1] U.S. Bureau of Economic Analysis (2017). The estimate corresponds to the year 2016 using the definition by Spulber (1996) in table 1, whereby intermediation includes retail trade (5.9 percent of GDP for the year 2016), wholesale trade (5.9 percent), finance and insurance (7.3 percent), business services (12.4 percent), and other services (2.3 percent). This estimate assumes that intermediation activities in all other sectors are zero, so it is a conservative estimate (see Spulber 1996).

[2] This is the well known result when there are successive monopolies at two layers of production, and goes back to Lerner (1934) (for further references see, *e.g.*, Spengler 1950 and Tirole 1988, pp. 174-6; see Luco and Marshall 2018 for a recent investigation on vertical integration with multiproduct firms). When the manufacturer and intermediary layers are both oligopolistic, and vertically integrated and unintegrated manufacturers coexist, the presence of intermediaries may increase or decrease the price of the final good (Salinger 1988). With consumer search the double marginalization problem is worsened, resulting in higher wholesale and retail prices due manufacturer's demand being more inelastic (Janssen and Shelegia 2015).

[3] See Spulber (1999) for a survey. See next subsection for the related literature.

and to guarantee quality. There is also a large empirical literature studying specific roles of intermediaries in many markets, such as online markets, two-sided platforms, financial markets, banking, asset pricing, labor markets, agrifood chains, and facilitating trade. However, there has been little empirical work to address the central question of what are the overall welfare implications of intermediaries in the industry when intermediaries offer simultaneously consulting, search, and purchase aggregation services to differentiate their products from the ones of the manufacturers. Yet, ignoring these additional services has significant consequences on the theoretical and empirical predictions for the determination of prices and consumer choices in these markets.

In this paper, we provide empirical estimates of the welfare of intermediation in vertical markets when intermediaries simultaneously provide consulting, search, and purchase aggregation services as defined in the model below. There are two major challenges to identifying the value of intermediaries in such cases. The first challenge arises due the non existence of a counterfactual scenario without intermediaries in industries where intermediaries are present.[4] This precludes evaluating the performance of the market without intermediaries. The second challenge arises due to the difficulty of observing all the transactions between manufacturers, intermediaries, and consumers in the industry. This may introduce a selection problem when evaluating the behavior of the unobserved participants, in addition to the previous complication. In both cases, recovering consumer demand preferences and supply marginal costs requires strong assumptions about market participants.

To address these issues we collected a novel dataset with all meaningful transactions among manufacturers, intermediaries, and consumers in the Portuguese outdoor advertising industry for the year 2013. This allows us to exploit two unique features of the industry that allow us to quantify the welfare effects of intermediaries in this industry. First, there are two distribution channels in the outdoor advertising industry: consumers can purchase the product either directly from manufacturers, or through intermediaries. This helps us overcome the first challenge mentioned above, by comparing instances where the same combination of display format and manufacturer is sold in both distribution channels. We then

[4] Alternatively, the counterfactual scenario with intermediaries is unobserved in industries where intermediaries are not present.

use the model described below to compute the counterfactual value that the consumer would have obtained had the purchased been made in a distribution channel different from the one actually observed. Second, we collected market level data directly from all the meaningful manufactures and intermediaries in the industry, which encompass more than 95 percent of the volume of transactions in the industry. This helps us overcome the second challenge mentioned above.

To quantify the value of intermediation we develop an econometric model of the industry. The model features two distribution channels where consumers can buy advertising: the direct sales channel (DSC), where consumers purchase directly from the manufacturers; and the vertical sales channel (VSC), where consumers purchase using the intermediaries. On the demand side, consumers have preferences that are specific to each distribution channel and engage in costly search. To model demand we use a random coefficient nested logit model with costly search, as described in subsection 3.1. On the supply side, the industry consists of two vertical layers modeled using a two stages game, as described in subsection 3.2. In the top layer, the manufacturers produce display formats for the display of outdoor advertising (*manufacture* products) that they sell to the intermediaries at wholesale prices. Manufacturers and intermediaries bargain over wholesale prices through Nash bargaining. This is the manufacture game. In the second layer, manufacturers and intermediaries sell the display formats (*final* products) to the consumers, competing on prices. This is the retail game.The distribution channels are represented by two types of retailers: DSC and VSC retailers. The DSC retailers are the manufacturers who sell final products to the consumers charging DSC prices. The VSC retailers are the intermediaries who charge VSC prices to the consumers. We use the terms "intermediaries" and "VSC retailers" interchangeably.

We estimate the model in two steps. First, we estimate the parameters that characterize demand without using the supply side model. For the estimation of the demand, we use an adapted version of the procedure proposed by Moraga-González, Sándor, and Wildenbeest (2015). This procedure adapts the nested fixed algorithm used by Berry, Levinsohn, and Pakes (1995) by allowing for an endogenous choice set for each consumer, which is the outcome of the search step. To identify the price coefficient, the heterogeneity parameters, and the search costs parameters, we rely on instruments with the exclusion restrictions discussed in subsection 4.1. Then, we estimate the parameters that characterize supply (retail and manufacture marginal costs, and bargaining weights) conditional on the demand estimates from the first step. To identify the supply side parameters we use the equilibrium conditions from the model, and additional restrictions using the vertical structure in our empirical setting. We use the first order conditions from manufacture and retail games, and that the manufacture marginal costs are the same for a display formats—the physical products in this industry—sold to VSC retailers and to consumers as discussed in subsection 4.2.

Our strategy to estimate the welfare implications of intermediaries consists of comparing circumstances where the same combination of display format and manufacturer is sold by DSC and VSC retailers, and using the model to estimate the value to consumers of each of the additional services provided by the VSC retailers. The VSC retailers provide three additional services to the consumers, and charge a margin for them. The additional services are: (i) *search services*, whereby VSC retailers provide information to consumers about display formats from multiple manufacturers, thus decreasing consumers' search costs; (ii) *purchase aggregation services*, whereby consumers benefit from quantity discounts that VSC retailers obtain because they aggregate purchases from multiple consumers; and (iii) *consulting services*, defined as the residual gross utility of buying from VSC retailers relative to a DSC retailer.[5] The market structures in the vertical layers determine the prices and margins charged by the manufacturers and retailers.

We use the estimated equilibrium model to simulate four counterfactual scenarios to quantify the value of intermediaries. First, we simulate the equilibrium of an industry where retailers do not offer consulting services. Second, we simulate the equilibrium of an industry where retailers do not offer search services. Third, we simulate the equilibrium of an industry where retailers do not offer purchase aggregation services. Finally, we remove simultaneously the three types of services to evaluate the total welfare impact of intermediaries. In the latter we find that the presence of intermediaries increases welfare because the value of their services outweighs the additional margin charged.

In summary, we make three main contributions. First, we combine a novel data set with a new econometric equilibrium model to estimate consumer demand preferences and

[5]See subsection 6.2 for the definition of these services in terms of the model.

marginal costs in the presence of intermediation, consumers' costly search, and bargaining between manufactures and intermediaries. The data includes all meaningful transactions in an industry where consumers can choose whether to use intermediaries. The model includes consumers who have preferences that are specific to each distribution channel and engage in costly search on the demand side; and two layers of activity with two distribution channels on the supply side. Second, we quantify the value of intermediaries in oligopolistic markets where intermediaries offer these additional services to differentiate their products from the ones of the manufacturers. Finally, from a normative perspective, our estimates show that the presence of intermediaries in the outdoor advertising industry is welfare improving because the benefits to consumers from the additional services provided by the intermediaries outweighs the additional margin charged by the intermediaries.

The rest of the paper is organized as follows. Section 2 describes the industry, the data, and presents stylized facts about the industry. Section 3 presents the equilibrium model. Section 4 discusses identification and estimation of the demand and supply. Section 5 presents the estimation results. The welfare analysis is performed in section 6. Section 7 concludes. Robustness analysis, extensions, and details about the data and the model are in the appendix.

1.1 Related Literature

This paper contributes to the literature that studies intermediaries. The role of firms as intermediaries has been studied extensively. Some papers include, *e.g.*, Yanelle (1989), Gehrig (1993), Rust and Hall (2003), Hagiu and Jullien (2011), Wright and Wong (2014), and Edelman and Wright (2015). Spulber (1999) provides a comprehensive study of intermediation, including how intermediaries alleviate problems associated with search costs and a thorough discussion of additional services provided by intermediaries. Some explanations why intermediaries arise are to facilitate matching of buyers and sellers as in Rubinstein and Wolinsky (1987), to guarantee quality as in Biglaiser (1993) and Spulber (1996), and recently as rent extraction (Farboodi, Jarosch, and Menzio 2017). Our case is closest to that in Rubinstein and Wolinsky (1987) and Spulber (1995, 1999) in that intermediaries create value by reducing search costs, and by providing additional services to the consumers.

The role of intermediaries has been studied a many markets. There is a large literature studying the role of intermediaries in online markets (*e.g.* Brynjolfsson and Smith 2000; Morton, Zettelmeyer, and Silva-Risso 2001; Brown and Goolsbee 2002; Brynjolfsson, Hu, and Smith 2003; Baye, Morgan, and Scholten 2003; Ellison and Ellison 2009; Quan and Williams 2016),[6] and in financial markets, banking, and asset pricing (*e.g.* James 1987, Diamond 1984, He and Krishnamurthy 2013; Brunnermeier and Sannikov 2014; Gavazza 2016).[7] Intermediation also plays an important role in labor markets (*e.g.* Stanton and Thomas 2016), agrifood chains (*e.g.* Lee, Gereffi, and Beauvais 2012), facilitating trade (*e.g.* Ahn, Khandelwal, and Wei 2011), and certifying information in markets with adverse selection (*e.g.* Biglaiser 1993; Lizzeri 1999; Biglaiser, Li, Murry, and Zhou 2017). Relative to these papers, our contribution is to estimate the welfare implications due to the presence of intermediaries in the industry, accounting for the change in the market structure created by the presence of the intermediaries, and the three additional services that intermediaries offer to consumers which differentiates their products from the ones of the manufacturers. The literature studying outdoor advertising is nonexistent. The only paper that we are aware is Pereira and Ribeiro (2018); they study capacity divestitures in this industry, not intermediation.

Our demand model in subsection 3.1 is related to the literature that uses models of discrete choice between differentiated products with costly search.[8] Our demand model is closest to De los Santos, Hortaçsu, and Wildenbeest (2012), Honka (2014), and Moraga-González, Sándor, and Wildenbeest (2015), who develop discrete choice models of demand in which consumers engage in costly search with fixed sample size.[9] None of these papers consider

[6]Some references in the growing literature on two-sided platforms as intermediaries include, *e.g.*, Bernard and Jullien (2003), Evans (2003), Rochet and Tirole (2003), Armstrong (2006), Galeotti and Moraga-González (2009), Boik (2016).

[7]See O'Hara (1995), Freixas and Rochet (2008), and Frankel, Galli, and Giovannini (2009) for surveys.

[8]For studies of the formation of consideration sets with fixed sample search see, *e.g.*, Roberts and Lattin (1991) and Mehta, Rajiv, and Srinivasan (2003) in the marketing literature.

[9]See also Goeree (2008), Salz (2017), Gavazza (2016), and Ershov (2018).

preferences specific to the distribution channel, which is the main focus this paper.[10] We incorporate these preferences using the distribution assumptions of the nested logit (*e.g.* Berry 1994; Cardell 1997), that we embed into a random coefficient discrete choice demand model with costly search.[11] For the estimation of the demand, we adapt the estimation procedure from Moraga-González, Sándor, and Wildenbeest (2015) to account for the additional term in the choice probability that the preferences for the distribution channel introduce, which modifies the computation of the market share function from the estimation algorithm (see subsection 4.1 for details).

On the supply side, our model in subsection 3.2 is related to the literature that models the vertical relations between manufacturers and intermediaries/retailers. Our model features two layers of activity (manufacturers and VSC retailers), and two distribution channels where consumers can purchase (VSC and DSC retailers). The two layers of activity are related vertically as in, *e.g.*, Brenkers and Verboven (2006), Mortimer (2008), Bonnet and Dubois (2010), Villas-Boas (2007), and Dubois and Sæthre (2016). The main difference between these papers and ours, is that in our model manufacturers and VSC retailers bargain over wholesale prices through Nash bargaining.[12] Our bargaining model is standard and similar to, *e.g.*, Crawford and Yurukoglu (2012), Draganska, Klapper, and Villas-Boas (2010), Grennan (2013), Gowrisankaran, Nevo, and Town (2015), Ho and Lee (2017), Crawford, Lee, Whinston, and Yurukoglu (2018). The main difference between the bargaining models in these papers and ours, is that in our model the retailers in both distribution channels can sell their products to the consumers. This occurs in our model after the bargaining process, where prices are set to consumers through standard Bertrand competition. Thus, the in-

[10] The information structure is also different in our model relative to these papers. In our model consumers face uncertainty over *both* the price and the realization of the random shock of each product (similar to Pires 2016), while in De los Santos, Hortaçsu, and Wildenbeest (2012) consumers only face uncertainty about the price of the product (not about the realization of the random shock), and in Moraga-González, Sándor, and Wildenbeest (2015) consumers only face uncertainty about the realization of the random shock (not about the price of the product). See subsubsection 3.1.3 for details.

[11] For other recent applications of the random coefficient nested logit model see, *e.g.*, Grennan (2013), Ciliberto and Williams (2014), Conlon and Rao (2015), and Miller and Weinberg (2017). None of these papers incorporate costly search.

[12] Bargaining models give rise to quantity discounts (or nonlinear pricing schemes). In our case, the quantity discounts between manufacturers and VSC retailers are then partially transferred to the consumers by the VSC retailers. There is a large literature studying quantity discounts and nonlinear pricing (see, *e.g.*, Miravete 2002; Busse and Rysman 2005; McManus 2007; Cohen 2008; Chu, Leslie, and Sorensen 2011; Miravete and Röller 2004a; Miravete and Röller 2004b; Nevo, Turner, and Williams 2016; Donna and Pires 2016).

struments and identifying assumptions to recover equilibrium margins as a function of the demand primitives, and bargaining power of VSC retailers and manufacturers are different (see subsection 4.2 for details). Similar to Grennan (2013) we do not estimate all bargaining and cost parameters because we do not estimate a cost function.

2 Portuguese Outdoor Advertising Industry

In this section, we describe: (i) the Portuguese outdoor advertising industry, (ii) the data set, and (iii) three stylized facts about the industry.

2.1 Industry Overview

In this subsection, we give an overview of the Portuguese outdoor advertising industry.

Agents. There are three main economic agents in the Portuguese outdoor advertising industry: (i) manufacturers, (ii) retailers, and (iii) consumers. A *manufacturer*, also called media owner, is a firm that installs and commercially exploits equipment for the display of outdoor advertising; examples include J.C. Decaux Group, Cemusa, Mop, *etc.* A *retailer*, also called media group, is an intermediary that buys advertising from the manufacturer on behalf of the consumer; examples include Omnicom Media Group, WPP Plc., Power Media Group Inc., *etc.* Retailers also offer consumers additional services such as consulting services, advertising planning campaigns, and information about the products of several manufacturers.[13] Finally, a *consumer*, also called advertiser, is a firm that demands advertising to promote its products. So consumers in this industry are firms that buy "exposure" in the manufacturer advertisement network.[14] For example, consumers buy 200 faces[15] distributed in the national network of J.C. Decaux Group, but they cannot choose specific 2 m^2 panels located at a particular place. Most of the purchases are in the national network, which is the

[13] A retailer is set of media agencies and a central purchasing agency with a common owner. Media agencies plan and buy advertising campaigns. The central purchasing agency aggregates purchases of media agencies and places orders to the manufacturers. A retailer might own several media agencies either because they specialize in different industries or to avoid confidentiality issues with advertisers.

[14] The advertisement network refers to the location of the different display formats available from each of the manufacturers

[15] *E.g.* 200 panels of 2 m^2, which is the most popular display format as explained below.

focus of this paper. The exposition is similar across manufacturers.[16]

Vertical Relations. Consumers make 85 percent of their purchases from the retailers and the remaining 15 percent directly from the manufacturers (table 1 discussed in subsection 2.2). Thus, in this industry there are two active distribution channels: (i) the *Vertical Structure Channel* (VSC), whereby consumers purchase manufacturers' products through the intermediation of retailers; and (ii) the *Direct Sales Channel* (DSC), whereby consumers purchase manufacturers' products directly from the manufacturers. We refer to the retailers in the VSC as "VSC retailers" and to the manufacturers that sell directly to the consumers in the DSC as "DSC retailers." Similarly, we refer to the price charged to the consumers by the VSC retailers (DSC retailers) as "VSC price" ("DSC"). [17] Figure 1 displays the vertical relations in the Portuguese outdoor advertising industry.

[*Figure* 1]

Retailers' Services. Retailers provide three main services to the consumers in this industry. First, they provide consulting services (Spulber 1996, 1999). Retailers assist consumers with their advertisement campaigns by helping them to make decisions such as the type and number of display formats to buy (*i.e.* type and total number of panels), their location, and the duration of the advertising campaign.

Second, retailers provide purchase aggregation services to the consumers. Retailers aggregate the purchases from multiple consumers when buying from the manufacturers. This gives retailers larger bargaining power to negotiate with the manufacturers, enabling them to obtain lower prices per unit relative to the consumers (*i.e.* "quantity discounts" as defined in subsubsection 2.3.1). These discounts are partially transferred to the consumers (subsubsection 2.3.1).

Finally, retailers provide search services to the consumers. When a consumer contacts a retailer, the retailer provides information about the prices available of all the products from the manufacturers in the industry. Consumers would need to contact each of the

[16] As explained below, we observe the manufacturers' capacity.

[17] The VSC price is typically referred to as retail price in other industries.

manufacturers separately to collect this information by their own. Retailers collect this information once per period (*e.g.* month) and use it for the orders of multiple consumers. This allows retailers to benefit from economies of scale relative to the consumers (we document this in subsubsection 2.3.3). In addition retailers have more experience than consumers collecting this information from the manufacturers, which allows them to do it more efficiently (*i.e.* lower search costs due to better search technology).

Display Formats. There are three main display formats: (i) 2 m^2 panels,[18] (ii) Seniores, and (iii) Others. Panels of 2 m^2 include city information panels, bus shelters, kiosks, *etc.* A Senior is an advertising panel with an area between 8 and 24 m^2. The last category, "others," encompasses Transports and Special Formats. A Transport includes panels on moving vehicles (*e.g.* buses, trains, taxis, *etc.*) or transport hubs (*e.g.* airports, railway's stations, subways' stations, *etc.*). Finally, a Special Format is large panel typically made by special request to be displayed, *e.g.*, on buildings' gables. We define a product as a combination of display format, manufacturer, and retailer (see subsection 2.2).

Payment Schedules. Contracts and payment schedules between manufacturers and VSC retailers are negotiated because all participants in the industry are firms. Manufacturers charge a price schedule that consists of a linear price and quantity discounts in the form of rebates. Consumers' purchases in the DSC exhibit no quantity discounts (see table 4 described in subsubsection 2.3.1). However, when consumers purchase in the VSC, these rebates (that the VSC retailers obtain from the manufacturers) are partially transferred to the consumers. Payment schedules between retailers and consumers are posted prices from the consumers' perspective (Pereira and Ribeiro 2018).

Productive Capacity. In the short run the productive capacity of each manufacturer and, thus, of the industry is fixed. The capacity is measured by the installed display equipment available for outdoor advertising. To operate, manufacturers must first obtain the right to use the space where the display equipment is installed, either through a public tender or direct contracting. This right is obtained from the "site owners", which are the landlords

[18] Also referred as "mupis" in the industry of Romance countries.

of the physical space where the display equipment is installed. Site owners include transit authorities, airports, supermarkets, malls, and other private landlords. The rights between the manufacturers and the site owners are set by long term contracts that last up to 20 years. In this paper we focus on the year 2013, so the productive capacity is fixed. Moreover, the inspection of our data on manufacturers' installed capacity and monthly usage indicates that capacity limits are never attained in our sample for any of the manufacturers. So manufacturers always operate below capacity.

Market Concentration. The Portuguese outdoor advertising market is quite concentrated both at the manufacture and retail levels. At the manufacture level there are three large national firms that are responsible of 77.6 percent of the sales in the market. At the retail level the five largest VSC retailers are responsible of 48.2 percent of the sales. See appendix B.1 for details.

2.2 Data

The Data Set. The data were obtained from various sources. We obtained administrative data from all the meaningful manufactures and retailers in the industry for the year 2013 aggregated at the monthly market level. By meaningful we mean that our data encompass more than 95 percent of the volume of transactions in the industry. A product is a combination of display format, manufacturer, and retailer. We consider 3 display formats:[19] 2 m^2 panel, senior, and an additional category aggregating the remaining formats that have negligible weight individually.[20] We consider 4 manufacturers: the 3 main manufacturers in the industry (J.C. Decaux Group, Cemusa, and Mop) and an additional manufacturer that aggregates the smaller manufacturers. Finally we consider 9 retailers: the 5 main VSC retailers in the industry (Omnicom Media Group, WPP Plc., Power Media Group Inc., Havas Media Group, and Interpublic Group of Companies), 1 additional VSC retailer that aggregates the smaller VSC retailers, and 3 DSC retailers representing the direct sales of each of the 3 larger manufacturers (J.C. Decaux Group, Cemusa, and Mop).[21] Examples of products

[19] See previous subsection for a description of display formats.
[20] This category includes special and transport formats
[21] There are not direct sales through the other manufacturers.

are: J.C. Decaux Group's 2 m^2 panels sold by Havas Media Group, Cemusa's seniors sold by WPP Plc., and J.C. Decaux Group's 2 m^2 panels sold directly by J.C. Decaux Group. Henceforth, and for confidentiality reasons, we refer to the 3 main manufacturers as m_1, m_2, and m_3, not necessarily in the order above, to the additional manufacturer as m_4, to the retailers in the DSC as r_1^d, r_2^d, r_3^d, by the same order as the 3 main manufacturers, to the 5 main VSC retailers as $r_4^v, \ldots, r_8^v$, not necessarily in the order above, and to the additional VSC retailer as r_9^v. Figure 1 summarizes this information. See appendix A.1 for details about the procedures to clean the data.

Characteristics of the manufacturers and retailers were collected by inspecting the websites of the retailers and manufacturers. Google trends data—used to construct moment conditions that identify the search parameters on the demand side—were obtained from Google Trends Portugal. See appendix A.2 for details.

In each month and for each triplet of display format, manufacturer, and retailer we observe: the total sales, measured in Euros; the total quantity of advertising sold, measured in advertising faces and square meters; the wholesale prices charged from the manufacturers to the retailers, measured in Euros; the commissions, fees, and rebates paid to the manufacturers, measured in Euros; and the installed capacity, measured in advertising faces.[22] We also observe characteristics for each manufacturer and retailer, such as the number of offices.

$[Table\ 1]$

Products. We define a product as a combination of display format, manufacturer, and retailer.[23] Panel A in table 1 shows the percentage of sales to consumers by each combination of manufacturer-retailer in the sample. Panel B in table 1 shows the percentage of sales of each of the 57 inside products in the sample. Note that the total number of inside products in the sample, 57, is lower than the total possible products in the market, 81.[24] This is because:

[22] From the manufacturers we collected the data from the first week of each month.

[23] Including DSC and VSC retailers.

[24] The total possible number of products to the consumers in the market is (see table 1):

$$\underbrace{(\text{3 Display Formats}) \times (\text{4 Manufacturers}) \times (\text{6 VSC Retailers})}_{\text{VSC}} + \underbrace{(\text{3 Display Formats}) \times (\text{3 DSC Retailers})}_{\text{DSC}} = 81.$$

(i) some VSC retailers only sell a subset of display formats from certain manufacturers, the subset with which they contracted,[25] and (ii) some DSC retailers do not sell certain display formats directly to consumers.[26] Note that all of the VSC retailers contract with all of the three largest manufacturers. This rules out the possibility that some retailers do not negotiate with some of the manufacturers due to selection based on unobservables.

Wholesale and Retail Prices in the VSC. Table 2 reports summary statistics on wholesale and VSC prices for each display format (see figure 1 for definitions of the prices charged by each agent). VSC prices, *i.e.* retail prices, are higher than wholesale prices, as expected. Panel B shows large differences in prices across both manufacturers and retailers holding constant the display format. These price differences and the differences in the observed market shares suggest that differentiation is important. Table A1 in the appendix compares wholesale and retail prices in the VSC by manufacturer and by retailer for the display format 2 m^2 panel. There are substantial differences in VSC and wholesale prices across VSC retailers holding constant the display format and the manufacturer. For example, retailer r_9^v is the most expensive retailer, including DSC retailer, for 2 m^2 panels manufactured by m_2, but the cheapest retailer for seniors manufactured by m_2. Tables 2 and A1 also show that differences in wholesale and VSC prices are small. This suggest that most of the differences in VSC prices are explained by differences in wholesale prices and that profits margins of VSC retailers are small.

[*Table* 2]

Sales to Consumers in the VSC and DSC. Panel A in table 1 shows that 85.2 percent of the sales to consumers are made through VSC retailers and the remaining 14.8 percent are made through DSC retailers. There is substantial variation across months in the market shares of VSC and DSC sales (see figure A1 that is described in appendix B.2). Monthly sales in the DSC range between 13.9 and 42.6 percent (figure A1). DSC prices are higher than

[25] *E.g.* Panel B in table 1 shows that retailer r_8^v does not sell 2 m^2 panels manufactured by manufacturer m_3.

[26] *E.g.* Panel B in table 1 shows that retailer r_1^d, which corresponds to manufacturer m_1 selling directly to consumers, does not sell seniors in the DSC.

wholesale prices holding constant the manufacturer and the display format. This is because manufacturers offer quantity discounts to retailers; it may also suggest that manufacturers use direct sales as a price discrimination mechanism (in the DSC).

Table 3 shows that the median price paid by consumers is typically higher in the DSC than in the VSC. But occasionally prices in the DSC are lower than in the VSC (*e.g.* senior manufactured by m_1 in Panel B in table 3). This is the result of two effects. On the one hand VSC retailers aggregate the purchases of several consumers, allowing them to obtain lower prices per unit (due the quantity discounts) when negotiating with the manufacturers (see subsection 2.3). This lowers VSC prices and increases VSC price dispersion (higher standard deviation) relative to the DSC prices.[27] On the other hand VSC retailers offer additional services to the consumers (*e.g.* consulting services, advertising planning campaigns, information about the products of all manufacturers, etc.) that are not offered by DSC retailers. This increases VSC prices relative DSC because VSC retailers charge for these additional services.

Market Shares. We use the data described above to build a data set of products (defined as a combination of display format, manufacturer, and retailer) sold in the DSC and VSC for each month of the year 2013 and their characteristics. Market shares are defined by dividing volume sales by the total potential sales in a given month (*i.e.* market size). This potential sales (or market size) was assumed to be twenty percent greater than the maximum observed total monthly sales of the year 2013. The market share of the outside good was defined as the difference between one and the sum of the market shares of the inside goods in each month. The outside good can be conceptualized as including products outside the sample (*e.g.* special request panels), outdoor advertising sold by other manufacturers and retailers (*e.g.* small manufacturers and retailers that operate locally), and not buying outdoor advertising. An observation in this data set represents a market share of a product as defined above in a given month. We consider 12 markets, one for each month of the year, and a continuum of heterogeneous consumers in each market.

[27] Note, however, that conditional on quantity discounts, the distribution of prices in the VSC is less disperse than in the DSC, as discussed on page 16.

[*Table* 3]

2.3 Three Stylized Facts

In this subsection we present descriptive patterns from the data. The Portuguese outdoor advertising industry is characterized by: (i) quantity discounts in the VSC, (ii) seasonal effects and large variation in the market shares, and (iii) substantial price dispersion conditional on quantity discounts and seasonal effects. In the next section we use the patterns presented here to construct the structural model.

2.3.1 Quantity Discounts in the VSC

Consumers' purchases exhibit quantity discounts in the VSC, but not in the DSC. By *quantity discounts* we mean that the price paid per square meter decreases with the volume purchased. Table 4 presents evidence about quantity discounts. It displays OLS regressions of the price paid by consumers per square meter of advertising on the total volume of advertising in a logarithmic scale, denoted by "Log(m^2)." Column 1 shows that the price paid by consumers per square meter of advertising decreases nonlinearly with the volume purchased of advertising. In column 2 we include an interaction between "Log(m^2)" and "VSC." The variable "VSC" is a dummy variable that equals 1 if the consumer performed the purchase through a VSC retailer, and 0 if the consumer performed the purchase through a DSC retailer. The interaction term is negative and statistically different from zero, while the coefficient on "Log(m^2)" is no longer statistically different from zero. This means that the purchases made by consumers in the VSC exhibit quantity discounts, and the ones made in the DSC not. Columns 3 and 4 show that similar results when including fixed effects for manufacturers, retailers, display formats, and months. Columns 2 and 4 show that the effect of quantity discounts is only present for the purchases made in the VSC.[28]

The presence of quantity discounts only in the VSC arises because the retailers aggregate the purchases from multiple consumers when buying from the manufacturers (subsection

[28]We obtain similar results by regressing the price paid by consumers per square meter of advertising on a polynomial of the amount of square meters of advertising purchased. Results are available upon request.

2.1). This results in quantity discounts on the wholesale prices of the products bought by the VSC retailers from the manufacturers. The resulting quantity discounts are then partially transferred to the consumers by the VSC retailers. Although the consumer could negotiate directly with the manufacturer (*i.e.* DSC retailer), the individual quantity purchased by the consumer is substantially lower than the total quantity purchased by the VSC retailers (because retailers aggregate the volume purchased by many consumers). So the purchases made by consumers in the DSC exhibit no quantity discounts.

[*Table* 4]

2.3.2 Seasonalities and Monthly Variation

The Portuguese advertising industry is also characterized by seasonal variations. The total volume purchased in the Portuguese advertising industry increases during the summer. For the estimation, we use monthly indicator variables to account for these seasonal effects. See appendix B.2 for details.

2.3.3 Price Dispersion and Returns to Consumer Search

The Portuguese outdoor advertising industry is characterized by substantial price dispersion: (i) across retailers holding constant the display format (product heterogeneity), the month of the year (seasonal effects), and the volume percentile (quantity discounts); and (ii) across months holding constant the display format, the manufacturer, the retailer, and the volume percentile. This indicates that the returns to consumers' search (for product's prices) are high in this market.[29]

Price dispersion is lower in the VSC than in the DSC conditional on quantity dis-

[29] In principle, one can also measure the variation in prices across identical goods sold by the same retailer holding constant the manufacturer, the month of the year (seasonal effects), and the volume percentile (quantity discounts). However, we cannot hold constant both, the retailer-manufacturer and the month of the year, due to the nature of our data (*i.e.* we do not observe the individual transactions of the consumers, only the total transactions per month per product and per retailer-manufacturer). Figure 2 measures price dispersion across retailers holding constant the display format, the month of the year, and the volume percentile (quantity discounts). Figure A2 in the appendix measures price dispersion across months holding constant the display format, the manufacturer, the retailer, and the volume percentile. Similar results are obtained. In the rest of this subsection we focus on figure 2 for consistency with the structural model in next section. See also footnote 32.

counts (figures 2 and A2). The top panel in figure 2 displays the distribution of the coefficient of variation of prices (CV) holding constant the display format, the month, and the volume percentile (*i.e.* each CV is computed within the unit of analysis in the tuple (Display Format, Month, Volume Percentile)).[30] The mean CV (pooling together sales in the VSC and DSC) is 45 percent. The mean CV for sales made in the VSC is 43 percent and for sales made in the DSC is 54 percent.[31] This indicates that returns to consumers' search (for product's prices) vary substantially by distribution channel.[32]

The bottom panel in figure 2 shows that the empirical CDF for sales made in the DSC first order stochastically dominates the one for sales made in the VSC. This indicates that consumers who buy in the VSC face lower price dispersion consistently. Figure 2 is consistent with VSC retailers providing search services to the consumers (subsection 2.1). Buying in the VSC can provide substantial returns to consumers with large search costs in this market.

The observables and fixed effects included in the model explain 82.8 percent of the price dispersion. The latter number refers to the R^2 of regressing the CVs in figure 2 on observables and fixed effects (which include months, products, manufacturers, retailers, and volume percentile fixed effects).

[$Figure$ 2]

[30] We obtain similar results using other measures of price dispersion such as percentile differences (*e.g.* difference between the 95th and the 5th price percentiles, difference between the 90th and the 10th price percentiles, *etc.*), range, and price gap. Results are available upon request.

[31] Note that substantial price variation is explained by quantity discounts, as emphasized in subsubsection 2.3.1. Ignoring quantity discounts one would incorrectly conclude that the price distribution is more disperse for sales made in the VSC than for sales made in the DSC (bottom panel in figure A1).

[32] As emphasized in subsection 2.1, VSC retailers also offer additional services to consumers (*e.g.* consulting services, advertising planning campaigns, information about the products of all manufacturers, *etc.*) that are not offered by DSC retailers. Figure A2 in the appendix shows similar patterns to the ones in figure 2 when we measure price dispersion across months holding constant the display format, the manufacturer, the retailer, and the volume percentile (*i.e.* identical products sold by the same seller holding constant the volume percentile). Thus the additional services provided by the VSC retailers in this industry shift the distribution of prices charged by each VSC retailer, but do not affect price dispersion in the market.

3 Model

3.1 Consumers

3.1.1 Set Up

To model demand we use a random coefficient nested logit model with costly search. Consumers know the products available in each market, but do not know the price or the realization of the random shock associated with each product. To learn this information consumers engage in costly search for retailers.

Consumer's choice is a two step process. In the first step the consumer chooses the subset of retailers to search. After searching for a retailer, the consumer learns the information (prices and the realization of the random shocks) of the products sold by that retailer. This determines the choice set, or *consideration set*, for each consumer type. The consideration set is given by the subset of products sold by all the retailers searched and the outside product, as described below. In the second step, after observing the prices and random shocks of the products sold by the retailers searched, the consumer chooses the product to purchase (*i.e.* the consumer chooses among the subset of products from the retailers searched). This is a standard discrete choice problem (*e.g.* Berry, Levinsohn, and Pakes 1995), where the only difference is that the choice set for each consumer is endogenously determined in the first step. We refer to the first step as the "search step" and to the second step as the "purchase step." Below we describe each step starting with purchase step.

Assume that there are $t = 1, \ldots, T$ markets, each with $i = 1, \ldots, I_t$ types of rational, utility-maximizing consumers. Consumers have different preferences for different products. A market is defined as a month. In each market t, there are J_t horizontally differentiated and mutually exclusive inside products, indexed by $j = 1, \ldots, J_t$. We index with $j = 0$ the outside product that allows consumers not to purchase any of the inside products. In each market, each consumer purchases one inside product or the outside product.

3.1.2 Step 2: Purchase step

Consider consumer i who searched R_i retailers in the search step in market t.[33] The indirect utility of consumer i for inside product j in market t conditional on the set of retailers R_i searched, denoted by $U_{ijt|R_i}$, is:

$$U_{ijt|R_i} = -\alpha_i p_{jt} + x_{jt}\beta + \tau_d^D + \tau_m^D + \tau_r^D + \tau_t^D + \xi_{jt} + \hat{\varepsilon}_{ijt}, \tag{1}$$

$$i = 1, \ldots, I_t, \qquad j \in \hat{J}_{tR_i} = \{\hat{j} : \hat{j} \in J_t \text{ is sold by retailer } r \in R_i\} \cup \{0\}, \qquad t = 1, \ldots, T,$$

where R_i denotes the subset of retailers searched by consumer i in market t; $\hat{J}_{tR_i}$ is the consideration set of consumer i, given by the subset of products sold by all the retailers searched and the outside product; p_{jt} is the price of product j in market t; x_{jt} is a S-dimensional (row) vector of observable characteristics of product j in market t; τ_d^D, τ_m^D, τ_r^D, and τ_t^D capture the preferences for display format d, manufacturer m, retailer r, and monthly seasonal effects in market t, using fixed dummy variables for display format, manufacturer, retailer, and monthly seasonal effects, respectively; ξ_{jt} is the valuation of unobserved (by the econometrician) characteristics of product j in market t; $\hat{\varepsilon}_{ijt}$ is a stochastic term described below; α_i are individual-specific parameters that capture consumers' preferences for price as described below; and β is a S-dimensional vector of parameters. In each market t, we normalize the characteristics of the outside product, $j = 0$, such that $p_{0kt} = x_{0t} = \tau_0^D = \tau_0^D = \tau_0^D = \tau_t^D = \xi_{0t} = 0$ for all t. Denote by $\overline{U}_{i\hat{j}t} \equiv -\alpha_i p_{\hat{j}t} + x_{\hat{j}t}\beta + \tau_d^D + \tau_m^D + \tau_r^D + \tau_t^D + \xi_{\hat{j}t}$ the indirect utility of consumer i for product $\hat{j} \in \hat{J}_{tR_i}$ in market t, net of the stochastic term, $\hat{\varepsilon}_{ijt}$. We model the distribution of consumers' preferences for price as follows:

$$\alpha_i = \alpha + \Sigma\nu_i, \qquad \nu_i \sim P_\nu(\nu_i) = \mathcal{N}(0, 1),$$

where α and Σ are parameters, ν_i captures unobserved (by the econometrician) individual characteristics, and $P_\nu(\cdot)$ is a parametric distribution assumed to be a standardized Normal, $\mathcal{N}(0, 1)$, for the estimation. Denote by $\delta_{\hat{j}t} \equiv -\alpha p_{\hat{j}t} + x_{\hat{j}t}\beta + \tau_d^D + \tau_m^D + \tau_r^D + \tau_t^D + \xi_{\hat{j}t}$ the

[33]Note that the set of searched retailers is market specific. To simplify the notation we omit the market subscript and we refer to the set of searched retailers as R_i instead of R_{it}.

mean utility for product $\hat{j}$ in market t (*i.e.* the portion of the utility that is constant across types of consumers). Note that $\overline{U}_{i\hat{j}t} = \delta_{\hat{j}t} - \Sigma\nu_i p_{\hat{j}t}$ for all i, $\hat{j}$, and t.

Consumers have preferences that are specific to each distribution channel (or retailer type), and the outside product. We capture this by decomposing the stochastic term, $\hat{\varepsilon}_{ijt}$, using the distributional assumptions of the nested logit with a factor structure (*e.g.* Berry 1994, Cardell 1997):

$$\hat{\varepsilon}_{ijt} = \zeta_{igt} + (1-\lambda)\varepsilon_{ijt}, \quad g \in \{0,1,2\},$$

where $g \in \{0,1,2\}$ define three groups (or nests) of nonoverlapping products for the outside product (denoted $g = 0$ with only one product), the products sold by the DSC retailers (denoted $g = 1$), and the products sold by the VSC retailers (denoted $g = 2$); ζ_{igt} has a unique distribution such that $\hat{\varepsilon}_{ijt}$ is extreme value (see Cardell 1997), and λ is a nesting parameter such that $0 \leq \lambda < 1$. A larger value of λ corresponds to a greater correlation in preferences for products in the same distribution channel and the outside product. Thus, a larger value of λ is associated with less substitution between products in different distribution channels and the outside product. Similarly, when $\lambda = 0$ the model in the second step collapses to a standard random coefficient mixed logit model (*e.g.* Berry, Levinsohn, and Pakes 1995; Nevo 2001), with no preference heterogeneity for distribution channels or the outside product, but with endogenous choice set from the search step (*e.g.* Moraga-González, Sándor, and Wildenbeest 2015) as described below.

For the estimation in subsection 4.1 it is convenient to write the nested logit choice probability as the product of two standard logit probabilities. Denote by $\mathbb{P}_{i\hat{j}t|R_i}$ the nested logit probability that individual i chooses product $\hat{j}$ in period t conditional on the searched retailers, R_i. Then we have:

$$\begin{aligned}
\mathbb{P}_{i\hat{j}t|R_i} &= \quad \mathbb{P}_{i\hat{j}t|\hat{g}R_i} \quad \times \quad \mathbb{P}_{i\hat{g}t|R_i}, && \text{(2a)}\\
&= \frac{\exp(I_{i\hat{g}R_i})}{\exp(I_{\hat{g}R_i})} \times \frac{\exp\left(\frac{\overline{U}_{i\hat{j}t}}{1-\lambda}\right)}{\exp\left(\frac{I_{i\hat{g}R_i}}{1-\lambda}\right)}, && \text{(2b)}
\end{aligned}$$

$$i=1,\dots,I_t, \qquad \hat{j}\in(\hat{g}\cap\hat{J}_{tR_i}), \qquad \hat{g}\in\{0,1,2\}, \qquad t=1,\dots,T.$$

where the first equality follows from the law of total probability; $\mathbb{P}_{i\hat{j}t|\hat{g}R_i}$ is the conditional probability of choosing product $\hat{j}$ given that the product is in group $\hat{g}$ and in the consideration set, $\hat{J}_{tR_i}$; $\mathbb{P}_{i\hat{g}t|R_i}$ is the marginal conditional probability of choosing a product in group $\hat{g}$ given that the product is in the consideration set $\hat{J}_{tR_i}$; the last equality follows from the nested logit structure using the decomposition into two standard logit probabilities (see appendix C for details); and the *inclusive values*, $I_{i\hat{g}R_i}$ and $I_{\hat{g}R_i}$, are given by:

$$\begin{aligned} I_{i\hat{g}R_i} &\equiv (1-\lambda)\mathbb{E}\left[\max_{j\in(\hat{g}\cap\hat{J}_{tR_i})} U_{ijt|R_i}\right], \\ &= (1-\lambda)\log \sum_{j\in(\hat{g}\cap\hat{J}_{tR_i})} e^{\overline{U}_{ijt}/(1-\lambda)}, \end{aligned} \tag{3a}$$

$$I_{\hat{g}R_i} \equiv \log\left(1+\sum_{g=1}^{2} e^{I_{i\hat{g}R_i}}\right), \tag{3b}$$

where $\mathbb{E}(\cdot)$ is the expectation operator taken over the *i.i.d.* shocks within group $\hat{g}$; the equation in (3a) follows from the maximum expected value of the logit model (*i.e.* within group $\hat{g}$);[34] and the equation in (3b) follows because the inclusive value of the outside product is equal to zero.

3.1.3 Step 1: Search step

Consumers know the products available in each market, but do not know the price, p_{jt}, or the realization of the random shocks, $\hat{\varepsilon}_{ijt}$, associated with each inside product.[35] Consumers can purchase an inside product only if they collect information about its price and random shock. They can engage in costly search to collect this information. A consumer who does not search, can only buy the outside product. The cost of consumers of collecting information about prices and random shocks from each VSC retailer is s^{VSC}, and from each DSC retailer is s^{DSC}.[36] We assume that if consumers search a retailer, they collect information about all

[34] See, *e.g.*, footnote 73 in appendix C for details.

[35] Before searching consumers only know the distributions of the prices, $\tilde{F}_{pt}(p)$, and random shocks, $\hat{\varepsilon}_{ijt}$. See page 23 and footnote 39 for details.

[36] The search cost includes the time spent to find and collect information about retailers, and processing costs (*e.g.* investigating in the retailer's webpage). Hence, our definition of search costs encompasses the cost of including a product at the purchase occasion and an evaluation cost (Hauser and Wernerfelt 1990).

the products sold by that retailer. Thus, our search costs are the cost of searching a retailer, not the cost of searching a product. The VSC retailers sell the products from multiple manufacturers (Panel B in table 1). Thus, searching for a VSC retailer allows consumers to collect the information about a larger set of products than searching for a DSC retailer. This allows us to rationalize the lower price dispersion observed in the VSC relative to the DSC (figures 2 and A2 discussed in subsubsection 2.3.3).[37]

We consider a fixed sample search process following De los Santos, Hortaçsu, and Wildenbeest (2012), Honka (2014), and Moraga-González, Sándor, and Wildenbeest (2015). First, consumers commit to search a fixed number of retailers. (The number could be zero, in which case the consumer buys the outside product in the purchase step.) This commitment is done before beginning the search. The search finishes after consumers searched all the retailers they committed to, even if they obtain a favorable search outcome early on.[38]

The expected net value for consumer i of searching a subset of retailers R_i in market t, denoted by V_{tR_i}, is the difference between the expected maximum utility of buying the most preferred product in that subset and the cost of searching for these retailers, denoted by SC_{itR_i}. That is:

$$V_{tR_i} = \int \max_{j \in \hat{J}_i} U_{ijt} \, \mathrm{d}F_{\hat{\varepsilon}}(\hat{\varepsilon}) \, \mathrm{d}\tilde{F}_{pt}(p) - SC_{itR_i} + \tilde{\varepsilon}_{itR_i} =$$
$$\int \log \left(1 + \sum_{\hat{g}=1}^{2} e^{I_{\hat{g}R_i}} \right) \mathrm{d}\tilde{F}_{pt}(p) + \hat{\gamma} - SC_{itR_i} + \tilde{\varepsilon}_{itR_i}, \quad (4)$$

where $\tilde{F}_{pt}(p)$ is the distribution of (inside) products' prices known by the consumers in market t, that we describe below; SC_{itR_i} is the cost of searching the subset of retailers R_i, that we describe below; $\tilde{\varepsilon}_{itR_i}$ is a random shock to the subset of searched retailers, that we described below; the equality in the second line follows from the expected maximum utility of the nested logit model conditional on the searched retailers (see appendix C for details); $\hat{\gamma} = 0.5772$ is

[37] See Wolinsky (1986) and Haan and Moraga-González (2011) for a theoretical analysis about the impact of search costs on the observability of prices.

[38] In our model consumers are firms demanding advertising. The decision of how many retailers to contact is typically made *ex ante*. This practice was explained to us by industry members. The fixed sample search assumption in our model is intended to capture such practice." See Morgan and Manning (1985) for a formal discussion of this behavior. For a discussion of the sequential and fixed sample search processes see, *e.g.*, Baye, Morgan, and Scholten (2006) and the references there.

the Euler's constant; and $I_{\hat{g}R_i}$ is the inclusive value of the set of products from the searched retailers that belong to subset $\hat{g}$ (excluding the outside product), and is given by:

$$I_{\hat{g}R_i} \equiv (1-\lambda) \log \sum_{j \in (\hat{g} \cap \hat{J}_{tR_i} \setminus \{0\})} e^{\overline{U}_{ijt}/(1-\lambda)}, \quad \hat{g} \in \{1, 2\}. \tag{5}$$

Before searching, consumers only know the distribution of prices of the (inside) products available in market t, $\tilde{F}_{pt}(p)$. We assume that consumers know the true distribution of prices in each market desegregated by distribution channel (or type of retailer). This is equivalent to say that consumers know *two* distributions of prices: the distribution of prices for the DSC retailers, and the distribution of prices for the VSC retailers. Consumers learn the prices of the set of products sold by each retailer when they search such retailer.[39]

We model the cost of searching a subset R_i of retailers, SC_{itR_i}, as:

$$SC_{itR_i} = \sum_{r \in R_i} sc_{irt},$$

$$sc_{irt} = \bar{S} + s \times \mathrm{g}_{rt}, \qquad r \in R_i,$$

where r index each of the searched retailers; sc_{irt} is the cost of searching for retailer r; $\bar{S}$ is a parameter measuring the fixed component of search; s is a parameter measuring the specific cost of searching for retailer $r \in R_i$; and g_{rt} is a variable denoting the visibility of retailer r in period t (for the estimation we use data on google searches).

We model the search problem of the consumer from an stochastic point of view. To do that, we add a random shock to the subset of searched retailers, $\tilde{\varepsilon}_{itR_i}$ in the equation in (4), as in De los Santos, Hortaçsu, and Wildenbeest (2012) and Moraga-González, Sándor, and Wildenbeest (2015). We do this for two reasons. First, the random shock has a natural interpretation in our setting. It captures consumer specific variation in the search costs for

[39] We have also performed the analysis using the following information structures for $\tilde{F}_{pt}(p)$: (i) Consumers only know the overall distribution of prices. That is, consumers only know *one* distribution with the prices for all the products in the market. (ii) Consumers know the distribution of each product in the market desegregated by distribution channel and by product. That is, consumers know 57 distribution of prices, where the number 57 corresponds to the number of inside products in the sample (see footnote 24).

Under the information structure in (i), consumers have less information than in the benchmark (*i.e.* where consumers know two distribution of prices, by distribution channel). Under the information structure in (ii), consumers have more information than in the benchmark. We obtained similar results in terms of the welfare analysis under the benchmark, and under (i) and (ii). Results are available upon request.

the retailers that are unobserved to the econometrician. The variance of this shock measures the degree of heterogeneity of consumers' search cost for retailers.[40] Second, solving for a consumer's optimal search strategy is a difficult problem. The consumer must simultaneously choose among a set of ranked stochastic options. Each choice is costly and only the best realized option is exercised. When there are many alternatives available in the market, finding the optimal choice set is extremely complex because there are many choice sets to be evaluated. By approaching this problem from an stochastic perspective, we smooth the choice set probabilities (of choosing a given subset of retailers), and we do not need to solve the search problem of every consumer. Instead we compute the probability that a given subset of retailers is searched by a consumer.[41] We assume that the term $\tilde{\varepsilon}_{itR_i}$ is drawn *i.i.d.* (across individuals, markets, and sets of retailers) from a type I extreme value distribution with location parameter $\mu_{\tilde{\varepsilon}} = 0$ and scale parameter $\sigma_{\tilde{\varepsilon}} > 0$.[42,43] Denote by $\Psi \equiv (\bar{S}, s^{DSC}, s^{VSC}, \sigma_{\tilde{\varepsilon}})$ the vector of search-related parameters.

Consumer i chooses the subset of searched retailers, R_i, that maximizes the expected net benefit of searching, V_{tR_i}. The probability that consumer i finds optimal to sample the subset of retailers R_i in market t, denoted by $\mathbb{P}_{R_i}$, is:

$$\mathbb{P}_{R_i} = \frac{e^{\overline{V}_{tR_i}/\sigma_{\tilde{\varepsilon}}}}{\sum_{\hat{R}_i \in \Lambda} e^{\overline{V}_{t\hat{R}_i}/\sigma_{\tilde{\varepsilon}}}}, \tag{6}$$

where $\overline{V}_{tR_i} \equiv V_{tR_i} - \tilde{\varepsilon}_{itR_i}$, is the expected value of searching a subset of retailers R_i net of the shock $\tilde{\varepsilon}_{itR_i}$, with V_{tR_i} given by the equation in (4); Λ is the powerset of all retailers; and the equality follows the well known logit choice probability.[44]

[40] Alternatively, it can be interpreted as measuring the degree of heterogeneity in the errors that consumers make when evaluating the net expected gain of a subset of retailers (De los Santos, Hortaçsu, and Wildenbeest 2012).

[41] An alternative procedure is provided by Chade and Smith (2006) with the Marginal Improvement Algorithm. This procedure is not computationally feasible in our empirical application, as explained above.

[42] *I.e.* the cumulative distribution function is: $F(\tilde{\varepsilon}_{ijt}) = e^{-e^{-\tilde{\varepsilon}_{ijt}/\sigma_{\tilde{\varepsilon}}}}$.

[43] The scale parameter, $\sigma_{\tilde{\varepsilon}}$, is identified because we have already normalized the scale of the ordinal utility of the consumer. This was done by normalizing the scale parameter of the stochastic term $\tilde{\varepsilon}_{ijt}$ in equation (1) to $\sigma_{\tilde{\varepsilon}} = 1$.

[44] See, *e.g.*, McFadden (1978).

3.1.4 Choice Probabilities and Market Shares

The unconditional choice probability of consumer type i for product j in market t is:

$$\mathbb{P}_{ijt} = \sum_{R_i' \in \Lambda} \mathbb{P}_{ijt|R_i'} \times \mathbb{P}_{R_i'}, \tag{7a}$$

$$= \sum_{R_i' \in \Lambda} \underbrace{\mathbb{P}_{ijt|gR_i'} \times \mathbb{P}_{igt|R_i'}}_{\mathbb{P}_{ijt|R_i'}} \times \mathbb{P}_{R_i'}, \tag{7b}$$

$$= \sum_{R_i' \in \Lambda} \underbrace{\underbrace{\frac{\exp(I_{igR_i'})}{\exp(I_{gR_i'})}}_{\mathbb{P}_{ijt|gR_i'}} \times \underbrace{\frac{\exp\left(\frac{\overline{U}_{ijt}}{1-\lambda}\right)}{\exp\left(\frac{I_{igR_i'}}{1-\lambda}\right)}}_{\mathbb{P}_{igt|R_i'}}}_{\mathbb{P}_{ijt|R_i'}} \times \underbrace{\frac{e^{\overline{V}_{tR_i'}/\sigma_{\tilde{\varepsilon}}}}{\sum_{\hat{R}_i \in \Lambda} e^{\overline{V}_{t\hat{R}_i}/\sigma_{\tilde{\varepsilon}}}}}_{\mathbb{P}_{R_i'}}, \tag{7c}$$

$$i = 1, \ldots, I_t, \qquad j \in J_t, \qquad g \in \{0, 1, 2\}, \qquad t = 1, \ldots, T.$$

where the equality in (7a) follows from the law of total probability; the equality in (7b) follows from the equation in (2a); and the equality in (7c) follows by replacing the expressions for $\mathbb{P}_{i\hat{j}t|\hat{g}R_i}$ and $\mathbb{P}_{i\hat{g}t|R_i}$ by the equation in (2b) (purchase step), and by replacing the expression for $\mathbb{P}_{R_i'}$ by the equation in (6) (search step).

Intuitively, equation (7) says that the unconditional probability that consumer type i chooses product j in market t is the weighted average of the conditional choice probability of each consideration set (or subset of retailers R_i'), where the weight of each consideration set is given by the unconditional probability that the consumer finds it optimal to sample that subset of retailers, $\mathbb{P}_{R_i'}$. In our model this weighted average of probabilities (equation 7a) can be written as the product of three standard logit formulas that are linked (equation 7c): (i) the conditional probability of choosing product j given that it belongs to group g and is sold by the subset of searched retailers R_i ($\mathbb{P}_{ijt|gR_i'}$); (ii) the conditional probability of choosing a product in group g given that is sold by the subset of searched retailers ($\mathbb{P}_{igt|R_i'}$); (iii) and the unconditional probability that the consumer finds optimal to sample the subset of retailers ($\mathbb{P}_{R_i'}$). The probabilities in (i) and (ii) are linked through the inclusive value I_{igR_i}. This inclusive value is the expected utility of consumer i from choosing a product among the ones in nest g, as can be seen from the equation in (3a).[45] The parameter λ reflects the degree of correlation in preferences for products in the same distribution channel. When $\lambda = 0$, the probabilities in (i) and (ii) collapse to a standard (random coefficients) logit model. Similarly, these probabilities are linked to the probability in (iii) through the subsets of searched retailers (or consideration sets). These subsets of retailers, R_i', enter in the value $\overline{V}_{tR_i'}$, and in the inclusive values $I_{igR_i'}$ and $I_{gR_i'}$. The parameters in the search costs, $SC_{itR_i'}$, determine the probability that the consumer finds it optimal to search R_i' retailers. When all search costs are zero, the consumer searches all retailers with probability one, and $\mathbb{P}_{R_i'} = 0$ for any other subset R_i' of retailers. Therefore, equation (7) collapses to $\mathbb{P}_{ijt} = \mathbb{P}_{ijt|g} \times \mathbb{P}_{igt}$, a standard random coefficients nested logit model without search.[46]

The market share function for product j in market t, denoted by s_{jt}, is obtained by integrating over the distribution of consumer types:[47]

$$s_{jt} = \int_{\nu_i} \mathbb{P}_{ijt} \, \mathrm{d}P_\nu(\nu_i), \tag{8}$$

where $\mathbb{P}_{ijt}$ is given by the equation in (7); and $P_\nu(\cdot)$ is a parametric distribution assumed to be a standardized Normal, $\mathcal{N}(0,1)$, for the estimation.[48]

3.2 Manufacturers and Retailers

In this subsection we present the supply side of the industry. The supply side model has two main characteristics. First, the industry consists of two layers of activity that are related vertically (*e.g.* Brenkers and Verboven 2006; Bonnet and Dubois 2010; Villas-Boas 2007), as displayed in figure 1. Second, there are two distribution channels (or retailers' types), where consumers can buy advertising: the direct sales channel (DSC) and the vertical sales channel (VSC). The game unfolds in two stages. In the first stage, manufacturers and VSC retailers

[45] This takes the standard logit formula because the choice of products within the nest g is a logit.

[46] If the nesting, price heterogeneity, and search costs parameters equal zero (*i.e.* if $\lambda = \Sigma = \bar{S}_t = s_{rt}^D = 0$ for all D, t), then the demand model collapses to a standard logit model, and $\mathbb{P}_{ijt} = {e^{\delta_{jt}}}/{\sum_{\hat{j}=0}^{J_t} e^{\delta_{\hat{j}t}}}$ in equation (7). Similarly, our random coefficient nested logit model with search collapses to the nested logit model if $\Sigma = \bar{S}_t = s_{rt}^D = 0$ for all D, t.

[47] Note that the market share is a function of the parameters and the characteristics of the products.

[48] For the estimation we approximate the integral in (8) by: $s_{jt} = \frac{1}{NS}\sum_{ns=1}^{NS} s_{jt}(v_{ns})$, where v_{ns} with $ns = 1, \ldots, NS$ are draws from $P_v(\cdot) = \mathcal{N}(0,1)$.

bargain over wholesale prices through Nash bargaining (*e.g.* Crawford and Yurukoglu 2012; Draganska, Klapper, and Villas-Boas 2010; Grennan 2013; Gowrisankaran, Nevo, and Town 2015). In the second stage VSC and DSC retailers set retail prices to consumers, through Bertrand competition.

3.2.1 Set Up

Consider an industry with a two layered vertical structure: (i) the manufacture layer, and (ii) the retail layer. In the manufacture layer, multi-product firms, called *manufactures*, produce basic production factors, called display formats, that they sell either to the VSC retailers or directly to the consumers (if they have a DSC channel). A production factor (display format) from a given manufacturer sold to different VSC retailers generates different *manufacture products*. In the retail layer, multi-product firms, called *VSC retailers*, combine manufacture products with their own retail production factors to produce *retail products*, also called display formats, that they sell to final consumers. Retailers have free access to all the information regarding the products that they carry in their portfolio.

We now present the notation to write the profit functions of the firms. There are $d = 1, \ldots, \mathcal{D}$ basic production factors, $m = m_1, \ldots, m_{\mathcal{M}}$ manufacturers and $r = r_1, \ldots, r_{\mathcal{R}}$ retailers. Each manufacturer may sell their product through one of the retailers or directly to the final consumer. Hence, there are, potentially, $\mathcal{J} = \mathcal{D} \times \mathcal{M} \times (\mathcal{R} + 1)$ final products.[49] Not all the manufacturers have a DSC channel. Therefore, it is convenient to divide them into *pure* manufacturers (those who sell only in the VSC) and *hybrid* manufacturers (those who sell in both the VSC and DSC). Denote by Ω_m^V the set of products that manufacturer m sells to VSC retailers, and by Ω_m^D the set of products that manufacturer m sells to consumers directly. Denote by Ω_r the set of products that retailer r sells to the final consumers. Denote by ω_j the wholesale price of manufacture product j, by ω the $\mathcal{J} \times 1$ vector of manufacture wholesale prices, by μ_j the marginal cost of manufacture product j, and by μ the $\mathcal{J} \times 1$ vector of manufacture marginal costs. Denote by p_j the price of retail product j, by $\mathbf{p}$ the $\mathcal{J} \times 1$ vector of retail prices, by ρ_j the marginal cost of retail product j, and by ρ the $\mathcal{J} \times 1$ vector

[49] Some of these products may not be offered because some VSC retailers may not carry the products of all manufacturers. As explained in subsection 2.2, the total number of inside products in our data is 57, and the total possible products in the market is 81. See footnote 24 and table 1 for details.

of retail marginal costs. The profit function of hybrid manufacturers contains two terms: the profits from selling the products in the VSC and in the DSC. For the profit of hybrid manufacturers in the DSC, we assume that the marginal cost is the manufacture cost plus a retail cost: $\mu_j + \rho_j$. Finally, denote by M the market size, by $s_j(\mathbf{p})$ the share of product j given by the equation in (8), and by $\mathbf{s}(\mathbf{p})$ the $\mathcal{J} \times 1$ share vector.[50]

The profit of a VSC retailer r is:

$$\Pi_r = \sum\nolimits_{j \in \Omega_r} (p_j - \omega_j - \rho_j)\, M s_j(\mathbf{p}),$$

The profit of a pure manufacturer m is:

$$\Pi_m^{pure} = \sum\nolimits_{j \in \Omega_m^V} (\omega_j - \mu_j)\, M s_j(\mathbf{p}),$$

The profit of an hybrid manufacturer m is:

$$\Pi_m^{hyb} = \sum\nolimits_{j \in \Omega_m^V} (\omega_j - \mu_j)\, M s_j(\mathbf{p}) + \sum\nolimits_{j \in \Omega_m^D} (p_j - \mu_j - \rho_j)\, M s_j(\mathbf{p}),$$

where we allow that the manufacture marginal costs of a given display format to differ depending on whether it is sold to the consumer or to a VSC retailer (*i.e.* we allow for $\rho_j > 0$ for $j \in \Omega_m^V$). In subsection 4.2 we set $\rho_j = 0$ for all $j \in \Omega_m^V$ to identify the supply system.

To account for the dual channel, but keep the equations compact, we define $\widetilde{\omega}_k$ as:

$$\widetilde{\omega}_k = \begin{cases} \omega_k, & k \in \Omega_m^V, \\ p_k - \rho_k, & k \in \Omega_m^D. \end{cases}$$

Then, the profit of manufacturers (pure or hybrid) can be written as:

$$\Pi_m = \sum\nolimits_{j \in \Omega_m} (\widetilde{\omega}_j - \mu_j)\, M s_j(\mathbf{p}),$$

where $\Omega_m = \Omega_m^V$ for pure manufacturers, and $\Omega_m = \Omega_m^V \cup \Omega_m^D$ for hybrid manufacturers.

[50] We omit the market subscript, t, for the variables in this subsection to simplify the notation.

3.3 Equilibrium

We construct the equilibrium by working backwards. The game unfolds in two stages. In the first stage manufacturers and VSC retailers bargain over manufacture prices (*i.e.* wholesale prices of display formats), in a Nash bargaining game. The equilibrium concept is Nash equilibrium in Nash bargains. We call this the "manufacture game." In the second stage VSC and DSC retailers set retail prices to the consumers, through a Nash Bertrand game. We call this the the "retail game." The equilibrium concept is subgame perfect equilibrium (equilibrium henceforth). We solve for the equilibrium by backward induction, starting with the retail game. Equilibrium prices are denoted with superscripts "*".

3.4 Stage 2: The Retail Game

The standard equilibrium first order necessary conditions for p_j are:

$$s_j(\mathbf{p}^*) + \sum\nolimits_{k\in\Omega_r} (p_k^* - \omega_k - \rho_k)\frac{\partial s_k(\mathbf{p}^*)}{\partial p_j} = 0, \tag{9a}$$

$$\sum\nolimits_{k\in\Omega_m^V} (\omega_k - \mu_k)\frac{\partial s_k(\mathbf{p}^*)}{\partial p_j} + s_j(\mathbf{p}^*) + \sum\nolimits_{k\in\Omega_m^D} (p_k^* - \mu_k - \rho_k)\frac{\partial s_k(\mathbf{p}^*)}{\partial p_j} = 0, \tag{9b}$$

The system of equations in (9) defines retail prices implicitly as a function of wholesale prices, $\mathbf{p}^* = \mathcal{P}(\omega)$ by applying the implicit function theorem to (9).

3.5 Stage 1: The Manufacture Game

Manufacturers and VSC retailers bargain bilaterally and simultaneously over wholesale prices, ω_j, as in Horn and Wolinsky (1988) and Collard-Wexler, Gowrisankaran, and Lee (2016).[51] The equilibrium concept is Nash equilibrium in Nash bargains: no manufacturer-retailer pair would like change their agreement, given all other agreements.[52] Motivated by our empirical setting, where VSC retailers negotiate with manufacturers each display format separately, we assume that all manufacturer-retailer pairs negotiate each wholesale price separately, as in Draganska, Klapper, and Villas-Boas (2010) and Bonnet, Bouamra-Mechemache, and Molina

[51] See also Collard-Wexler, Gowrisankaran, and Lee (2016) and Arie, Grieco, and Rachmilevitch (2016).

[52] Each pair of players maximizes the bilateral gains from trade, modeled by an asymmetric Nash bargaining solution, given the strategies of all others pairs.

(2016).[53] If the negotiations over ω_j fail, manufacture and retail products j are not sold. If the negotiations over ω_j succeed, the profit of manufacturer m from manufacture product j is $\Pi_{mj}(\omega) = (\omega_j - \mu_j)\, Ms_j(\mathcal{P}(\omega))$, and the profit of retailer r from retail product j is $\Pi_{rj}(\omega) = \left(p_j^* - \omega_j - \rho_j\right) Ms_j(\mathcal{P}(\omega))$. Denote by $\Pi_{r,-j}$ and $\Pi_{m,-j}$ the disagreements payoffs of retailer r and manufacturer m, respectively, when they bargain over ω_j. Denote by ν_{rmj} the bargaining weight of retailer r when it bargains with manufacturer m over ω_j. The Nash product of manufacturer m and retailer r for ω_j is:

$$\begin{aligned}\mathcal{N}_{rmj} &\equiv \left[\sum_{k\in\Omega_r} \left(p_k^* - \omega_k - \rho_k\right) Ms_k(\mathcal{P}(\omega)) - \Pi_{r,-j}\right]^{\nu_{rmj}} \left[\sum_{k\in\Omega_w} \left(\widetilde{\omega}_k - \mu_k\right) Ms_k(\mathcal{P}(\omega)) - \Pi_{m,-j}\right]^{1-\nu_{rmj}}, \\ &\equiv \left[\Pi_r(\omega^*) - \Pi_{r,-j}\right]^{\nu_{rmj}} \left[\Pi_m(\omega^*) - \Pi_{m,-j}\right]^{1-\nu_{rmj}}. \end{aligned} \tag{10}$$

Denote by $\Omega_x \backslash \{j\}$, $x = r, m$, the set of products firm x sells minus product j. Denote by ω_{-j} the $(\mathcal{J}-1)\times 1$ vector of manufacture prices without element ω_j. Denote by $\Delta s_k^{-j}(\mathcal{P}(\omega_{-j}))$ the change in the market share of product k if product j is not offered. We assume that the disagreement profits for manufacture m and retailer r when they bargain over ω_j are the maximum profits each could earn if product j were not offered, where the parties assume that other contracts would not be renegotiated if they did not reach an agreement. Hence, the disagreements payoffs are:

$$\begin{aligned}\Pi_{r,-j} &\equiv \sum\nolimits_{k\in\Omega_r\backslash\{j\}} \left(p_k - \omega_k - \rho_k\right) Ms_k^{-j}(\mathcal{P}(\omega_{-j})), \\ \Pi_{m,-j} &\equiv \sum\nolimits_{k\in\Omega_m\backslash\{j\}} \left(\widetilde{\omega}_k - \mu_k\right) Ms_k^{-j}(\mathcal{P}(\omega_{-j})).\end{aligned}$$

Note that the bargaining game takes place only over products that are sold in the VSC. However, when *hybrid* manufacturers bargain with retailers, the Nash product expression includes profits from both the VSC and DSC. This is because a change in the wholesale price that the parties negotiate will affect sales in the DSC channel as well.

[53] An alternative assumption, followed by Bonnet, Bouamra-Mechemache, and Molina (2016), is that each pair of manufacturer-retailer negotiate all their products jointly.

The first order necessary equilibrium conditions for ω_j are:

$$\nu_{rmj}\left[\Pi_r(\omega^*)-\Pi_{r,-j}\right]^{\nu_{rmj}-1}\left[\Pi_m(\omega^*)-\Pi_{m,-j}\right]^{1-\nu_{rmj}}\frac{\partial\Pi_r(\omega^*)}{\partial\omega_j}+$$
$$\left(1-\nu_{rmj}\right)\left[\Pi_r(\omega^*)-\Pi_{r,-j}\right]^{\nu_{rmj}}\left[\Pi_m(\omega^*)-\Pi_{m,-j}\right]^{-\nu_{rmj}}\frac{\partial\Pi_m(\omega^*)}{\partial\omega_j}=0,\quad j=1,...,\mathcal{J}. \quad (11)$$

4 Estimation

4.1 Demand: Estimation and Instruments

We estimate the parameters that characterize demand without using the supply side model. For the estimation we use an adapted version of the procedure proposed by Moraga-González, Sándor, and Wildenbeest (2015, henceforth MSW), as described below. To identify the price coefficient, the heterogeneity parameters, and the search costs parameters, we rely on instruments with the exclusion restrictions that we discuss below. Finally, we identify the value to consumers of the services provided by the VSC retailers by using the model and comparing instances where the same combination of display format and manufacturer is sold by DSC and VSC retailers.

4.1.1 Estimation

We estimate the demand model using the data from subsection 2.2 adapting the procedure used by MSW. The procedure by MSW adapts the nested fixed algorithm used by Berry, Levinsohn, and Pakes (1995, henceforth BLP) by allowing for an endogenous choice set for each consumer type i, which is the outcome of the search step. There are four major differences in our procedure relative to the one by MSW. First, due to our research question, we account for consumers' preferences for the two distribution channels. This introduces the additional multiplicative term, $\mathbb{P}_{igt|R_i'}$, to the choice probability in equation (7c), which enters into the market share computation in equation (8). Second, we identify the parameters on the demand side without specifying a functional form for the supply side, while in MSW identification relies on the functional form of a supply equation, similar to BLP. Third, due to the nature of our data, we do not use micro moments. Finally, related to the previous points, the instruments and identifying assumptions are different.

The model is estimated by GMM (*e.g.* Hansen 1982) and relies on the moment condition $\mathbb{E}[Z' \cdot \omega(\theta^*)] = 0$, where Z is a matrix of instruments, $\omega(\cdot)$ is a structural error term defined below, and $\theta^* = (\alpha, \Sigma, \lambda, \Psi)$ is the true value of the parameters. The GMM estimate is:

$$\hat{\theta} = \arg\min_{\theta} \left[\omega(\theta)' Z A^{-1} Z' \omega(\theta)\right],$$

where A is a consistent estimate of $\mathbb{E}(Z'\omega\omega' Z)$.

We now describe the estimation procedure. For each candidate parameter vector, we use equation (8) with the choice probability in equation (7c) to compute the market shares as a function of the parameters.[54] We define the error term as the unobserved products' characteristics and compute it by solving for the mean utility level, δ_{jt} that equates:

$$s_{jt}(p_{jt}, x_{jt}, \delta_{jt}; \Sigma, \lambda, \Psi) = S_{jt}, \tag{12}$$

where $s_{jt}(\cdot)$ is the market share function given by the equation in (8); and S_{jt} are the observed market shares obtained from the data. We use a contraction mapping[55] to solve for the implicit system of equations in (12) and identify the vector of mean utility levels. After solving this system of equations, the structural error term is defined as $\omega \equiv \delta_{jt}(p, x, S; \Sigma, \lambda, \Psi) + \alpha p_{jt}$. For the estimation, we concentrate the linear price parameter, α, out of the optimization problem to reduce the dimensionality of the nonlinear search.

We compute the standard errors for the estimates using the standard procedures (*e.g.*, Hansen 1982, Newey and McFadden 1994), correcting them to account that the simulation draws are the same for all of the observations in a market (*e.g.* BLP).

[54] The main difference relative to BLP is that the choice probability and, hence, the market share function incorporates the search step (this is the last term in equation 7c, denoted by $\mathbb{P}_{R'_i}$). Once the market share function is computed, the estimation procedure resembles BLP, as developed by MSW.

[55] MSW show that the vector of unobserved characteristics can be computed as the unique fixed point of a contraction mapping, similar to the one in BLP. Our contraction mapping is similar to the one in MSW. The difference, relative to MSW, is that we have the additional multiplicative term, $\mathbb{P}_{igt|R'_i}$, to the choice probability in equation (7c), which enters into the market share computation in equation (8).

4.1.2 Instruments

We rely on instruments with exclusion restrictions to identify the price coefficient, the heterogeneity parameters, and the search costs parameters. Identification requires at least one exogenous instrument for price, each heterogeneity parameter (Berry and Haile 2014, henceforth BH), and search costs. The value added by the VSC retailers is identified by comparing instances where the same combination of display format and manufacturer is sold the DSC and VSC, and using the model to infer the value to consumers.

Price parameter, α. At least one instrument is needed to identify α due to price endogeneity concerns (*e.g.* Berry 1994; BLP; Nevo 2001; BH). The structural error may be correlated with prices because retailers make their pricing decisions with knowledge about the valuations of the consumers in each market, conditional on the search step. In our case the structural error term is the unobserved month specific deviation from the overall mean valuation of the product. The supply side model in subsection 3.2 assumes that retailers in the industry observe this deviation. It enters into the market-specific markup term in the pricing equation, thus introducing a bias in the estimate of the price sensitivity, α. To address price endogeneity we use prices of the same product in other markets as instruments for the price of the product in the current market (*e.g.* Hausman 1996; Nevo 2001). The identifying assumption is that month-specific valuations for a product are independent across time, after accounting for display format, manufacturer, retailer, and months fixed effects. The prices of the same product are correlated across months due to the common marginal cost, while they are uncorrelated with month-specific valuations due to the exclusion restriction. We use average retail and average wholesale prices (excluding in both cases the price of the product in the month being instrumented) in all months, and lagged wholesale prices.[56]

Heterogeneity parameters, Σ and λ. The parameter Σ governs the distribution of the random coefficients (or heterogeneity in consumer preferences) for price. An instrument is needed to identify this parameter due to the endogeneity problem arising from the unknown parameter Σ, interacting with the endogenous variables, (s_{jt}, p_{jt}) (see BH for details). We

[56] One could potentially use retail and wholesale prices in all other months as instruments. See Chamberlain (1982) for a discussion of optimal instruments.

use the differentiation instruments proposed by Gandhi and Houde (2016). We construct instruments defined by a proximity measure counting the number of competitors located within one standard deviation of product j. Specifically, we use the count of other products whose prices lie within five Euros of the own price, and the interaction of this variable with product and manufacturer dummy variables.[57]

The nested logit parameter, λ, governs the substitution within and between subgroups of products (or nests) sold by the retailers in the DSC, VSC, and the outside product. An instrument is needed due to the unknown parameter λ, interacting with the endogenous within-group share (see BH). We use the number of products in the market within each distribution channel as an instrument. This is a "BLP instrument"[58] that has been previously used to identify the nested logit parameter (*e.g.* Miller and Weinberg 2017). The identifying assumption is that the structural error term is uncorrelated with the number of products in the market within each distribution channel. The power of the instrument comes from the number of products in the market within each distribution channel being negatively correlated with the share of the products within that distribution channel.

Search costs parameters, Ψ. A consumer searches a firm either because the benefits of searching that firm are high, or because the search costs are low. Identification of search costs requires shifters that affect only the benefit of searching and not the search costs.[59] We use two set of instruments. First, we use changes in product availability over time as instrument. The exclusion restriction is that the market share from the previous period only affects the benefit of searching, and is uncorrelated with cost of searching.[60] Intuitively, if the search costs are high, consumers will only search for the more preferred retailers. Thus, the prices of the products sold by the less preferred retailers would likely not affect the market shares of the products sold by the more preferred retailers, because the consumers will not be aware of possible lower prices from the less preferred retailers. In contrast, price reductions of less preferred retailers can affect the market shares of most preferred retailers when search costs

[57] We have also experimented with a band of ten Euros, and obtained similar results.

[58] See, *e.g.*, BH for a discussion of these instruments.

[59] An alternative approach to the one followed in this paper, would be to identify search costs from functional forms (*e.g.* Honka 2014), rather than with exclusion restrictions.

[60] This identification argument for the search costs parameters is similar to the one by Seiler (2013) and Pires (2016), but without accounting for dynamics.

are low. In general, consumers have more incentives to search, the larger is the variance from the distribution of prices (known by the consumers *ex ante*) of the inside products available in the market, denoted by $\tilde{F}_{pt}(p)$ in the model. So the correlation between prices of less preferred retailers, and market shares of more preferred retailers will be larger when this variance is large.[61] Second, we use Google trends data about weekly searches for the names of the manufacturers and retailers in Google Portugal. The exclusion restriction is similar to the one from the previous set of instruments. Namely, that the Google trend searches capture the visibility of retailers and manufacturers, thus affecting the benefit of searching for the consumers but is uncorrelated with their cost of searching. In terms of power, we find that, for all retailers/manufacturers, the first result displayed by Google Portugal was the webpage of the corresponding retailer/manufacturer after performing a search in Google Portugal with the name of the retailer/manufacturer used to construct the variable. See appendix A.2 for details about the Google trends data and the definition of the variables used.

Value of VSC Retailers. We comment on the empirical variation that identifies the value of VSC retailers. The identification problem arises due to the non existence of a counterfactual scenario without the VSC retailers. We evaluate the performance of the market without the VSC retailers by constructing such counterfactual scenario using our model and the fact that, in the outdoor advertising industry, goods can be purchased by the final consumer from either VSC or DSC retailers. To do that we compare instances where the same combination of display format and manufacturer is sold by a DSC and a VSC retailer, and compute the counterfactual value that the consumer would have obtained had the purchased been made in a distribution channel different from the one actually used. Consider consumer i and suppose that a given combination of display format and manufacturer is sold by a DSC and a VSC retailer in a given market, and that both types of retailers are part of the

[61] Variables that affect both, utility and search costs, enter the purchse probability in different ways. This allows us to identify the effects of such variables on search costs and on utility. Moraga-González, Sándor, and Wildenbeest (2015) note that under some conditions a combination of aggregate and consumer-level search data can jointly identify the effect of variables on search costs as well as the effect of a non-parametric specification of these same variables on the utility function. The idea behind non-parametric identification of the effect of these variables on the utility is that certain ratios of choice-set probabilities do not depend on the effect of these variables on search costs.

consideration set of the consumer. The value of a VSC retailer to consumer i is determined by the extent to which the consumer purchases from the VSC retailer, rather than from DSC retailer. According to the model, there are three (non mutually exclusive) channels for such decision by the consumer: (1) The VSC retailer offered a lower price due its access to quantity discounts from the manufacturer (purchase aggregation services); (2) the VSC retailer reduced the search costs to the consumer due to its access to the products of multiple manufacturers (search services); and (3) the gross utility to the consumer of purchasing from a VSC retailer is larger than the gross utility of purchasing from a DSC retailer (consulting services). The model decomposes the value of VSC retailers to consumer i into (1), (2), and (3), by constructing counterfactual scenarios where we remove each of these channels at a time. We do the same for all consumer types, and then exploit the unique feature of our data that we observe all the transactions in the industry to infer the value of VSC retailers.

4.2 Supply: Identification and Estimation

4.2.1 Identification

We now discuss identification of the parameters from the supply side conditional on the data that we observe, and the estimated demand system. The parameters from the supply side are: (i) the vector of retail marginal costs, ρ; (ii) the vector of manufacture marginal costs, μ; and (iii) the vector of bargaining weights, δ. We observe the vector of retail prices, $\mathbf{p}^*$, the vector of wholesale prices, ω^*, the vector of market shares, $\mathbf{s}$, and size of the market, M. Using the procedure described in subsection 3.1, we have an estimate of the demand system, $(\mathbf{s}(\mathbf{p}^*), \nabla_p \mathbf{s}, \overline{s})$.

The first order conditions from the retail game in the equation in (9) provide a system of $\mathcal{J}$ equations that "just identifies" the $\mathcal{J}$ vector of retail marginal costs (ρ). The first order conditions from the manufacture game in the equation in (11) also provide a system of $\mathcal{J}$ equations. Thus, in general, equation (11) cannot be used to identify both, the vector of vector of manufacture marginal costs (μ) and bargaining weights (δ) which jointly has dimension $2 \times \mathcal{J}$, without additional restrictions.

In our case, we obtain these additional restrictions using the particularity of the vertical structure in our empirical setting. Namely, that manufacturers sell the same display format to both, consumers (charging DSC prices) and VSC retailers (charging wholesale prices). A natural set of restrictions justified by this structure is that the manufacture marginal costs of a given display format in a given month is the same whether it is sold to the consumer or to a VSC retailer. In terms of the model, this implies that the manufacture marginal costs, μ, could be recovered using the first order conditions from the retail game (in the system in equation (9), without using the first order conditions from the manufacture game (in the system in equation 11), because the manufacturers are DSC retailers. So, in the first step, we use the first order conditions from the retail game in the equation in (9) to identify the vector of retail *and* manufacture marginal costs, using the fact that manufacture marginal costs are the same for a display format sold to a VSC retailer and to the consumer. In other words, μ is contained in ρ in our setting. Then, in the second step, we use first order conditions from the manufacture game in the equation in (11) to identify the bargaining weights. The $\mathcal{J}$ vector of bargaining weights, δ, is "just identified" using the system in equation (11) conditional on the estimated manufacture marginal costs, ρ, from the retailer game. In the next subsubsection we estimate the bargaining weights without imposing additional restrictions to the manufacture game. Our identification arguments exploit the vertical structure in our empirical setting and, thus, are different from the ones used previously in the literature (*e.g.* Grennan 2013; Gowrisankaran, Nevo, and Town 2015; Draganska, Klapper, and Villas-Boas 2010).[62]

4.2.2 Estimation

Step 1: Estimation of marginal costs. We use the first order necessary conditions from the retail game in the equation in (9). To recover the vector of marginal costs, ρ, we need to

[62]For example, Draganska, Klapper, and Villas-Boas (2010) use observed prices and quantities to estimate demand parameters. In addition, they have cost data, which they use to estimate marginal costs. Finally, they use the first order conditions to estimate the bargaining parameters. Grennan (2013) also uses observed prices and quantities to estimate demand parameters. However, he does not have cost data. This means that he cannot estimate all the bargaining and marginal cost parameters. Hence, he imposes restrictions on marginal cost parameters, and then uses first order conditions to estimate bargaining parameters. We use a somewhat different approach. Prior knowledge of the industry structure allows us to reduce the dimension of the marginal costs parameters (so that manufacture marginal costs can be estimated with the retailer marginal costs). Then, we use the first order conditions from the bargaining game to estimate the bargaining parameters alone.

compute the element $\frac{\partial s_{kt}(\mathbf{p}*)}{\partial p_{jt}}$ of matrix $[\nabla_p \mathbf{s}]'$. Changing the price of a single product has a negligible impact upon the *ex ante* price expectations in each distribution channel, therefore not affecting the choice probabilities for each set of stores, *i.e.*, $\frac{\partial s_{R'_i}}{\partial p_{jt}} \approx 0$. Thus:

$$
\begin{aligned}
\frac{\partial s_{kt}(\mathbf{p}*)}{\partial p_{jt}} &= \int_i \frac{\partial s_{ijt}(\mathbf{p}*)}{\partial p_{jt}} \mathrm{d}v_i, \\
&= \int_i \left(\sum_{R'_i \in \Lambda} \frac{\partial s_{ijt|R'_i}}{\partial p_{jt}} \times s_{R'_i} \right) \mathrm{d}v_i.
\end{aligned}
$$

From Grigolon and Verboven (2014, p. 934) the last derivative is given by:

$$
\frac{\partial s_{ijt|R'_i}}{\partial p_{kt}} = \begin{cases}
\alpha_i \left(s_{ijt} + \frac{\sigma}{1-\sigma} s_{ijt|g} - \frac{1}{1-\sigma} \right) s_{ijt} & \text{if } j = k, \\
\alpha_i \left(s_{ikt} + \frac{\sigma}{1-\sigma} s_{ikt|g} \right) s_{ijt} & \text{if } j, k \text{ are in the same nest } g, \\
\alpha_i s_{ikt} s_{ijt} & \text{if } j, k \text{ are in different nests,}
\end{cases}
$$

where $\alpha_i > 0$.

Step 2: Estimation of bargaining weights. We use the first order necessary conditions from the manufacture game in the equation in (11). Applying the envelope theorem and simplifying (see appendix D):

$$
\delta = \left[\left(\Lambda^{\mathcal{M}} \odot \overline{s} \right) (\omega^* - \mu) \right] \left[\left(\Lambda^{\mathcal{R}} \odot \overline{s} \right) (\mathbf{p}^* - \omega^* - \rho) \right]^{-1}, \tag{14}
$$

where:

$$
\overline{s} \equiv \begin{bmatrix}
s_1 & -\Delta s_2^{-1} & \cdots & -\Delta s_{\mathcal{J}}^{-1} \\
-\Delta s_1^{-2} & s_2 & \cdots & -\Delta s_{\mathcal{J}}^{-2} \\
\vdots & \vdots & \ddots & \vdots \\
-\Delta s_1^{-\mathcal{J}} & -\Delta s_2^{-\mathcal{J}} & \cdots & s_J
\end{bmatrix},
$$

is the matrix of shares and changes in shares in Draganska, Klapper, and Villas-Boas (2010, p. 62, first matrix) with $\Delta s_k^{-j} \equiv s_k^{-j}(\mathcal{P}(\omega)) - s_k(\mathcal{P}(\omega))$ denoting the difference between the market share of product k if product j is offered and if it is not.

We use the expression in (14) for the estimation of the bargaining weights.

5 Estimation Results

5.1 Demand Estimates

Table 5 provides the estimates of the demand model from subsection 3.1 using the estimation procedure from subsubsection 4.1.1. The table displays estimates from three specifications. Model 3 corresponds to the full model from subsection 3.1, and is our preferred specification. This model features random coefficients for the price coefficient, and a consumer search costs. Search costs are assumed to be the exponent of a constant term and a linear function of the previous period market share. We also compare the estimates from our full model to those from more traditional models. Model 1 is a logit model without search costs. It restricts the price sensitivity to be the same for every consumer. Model 2 is a random coefficients mixed logit model without search costs. In models 1 and 2 consumer can search each retailer for free. All three models include the nested logit structure, that allows for different correlation in preferences for products in the same distribution channel and the outside product.

[*Table* 5]

The mean price coefficient in model 3 is -0.06, and is statistically different from zero. The dispersion of the price sensitivity across consumers is also statistically different from zero. Search costs are important, as the constant term of that function is statistically different from zero with an estimated value of 0.589. The market share in the previous period does not seem to affect the cost of the search. We use the estimates from model 3 for the welfare analysis in the next section.

6 Welfare

In this section we use our estimates from section 5 to quantify the welfare impact of retailers. We simulate four counterfactuals scenarios that we describe in subsection 6.2. These counterfactuals are "partial" in the sense that the price vector does not respond to the changes in the demand system.[63] We use superscript c to denote a counterfactual.

6.1 Welfare Measures

Now we describe the welfare measures used in the counterfactual analysis. The expected consumer surplus, in dollars, for consumer type i is given by (see appendix C for details):

$$\mathbb{E}(CS_i) = \frac{1}{\alpha_i} \log \left\{ \sum_{R_i' \in \Lambda} \exp \left[\log \left(1 + \sum_{\hat{g}=1}^{2} e^{I_{\hat{g}R_i}} \right) - SC_{itR_i} \right] \right\} + C, \tag{15}$$

where $\mathbb{E}(\cdot)$ denotes the expectation operator taken over both random shocks $\hat{\varepsilon}_{ijt}$ and $\tilde{\varepsilon}_{ijR_i}$; the inclusive value, $I_{\hat{g}R_i}$ is given by the equation in (5); and C is a constant.[64]

Consumer welfare for type i is defined as the change in the consumer surplus (or compensating variation, CV) that results from a change in the services offered by the retailers. We compute the difference between the consumer surplus before and after such change. We consider four changes in the services offered by the retailers that we describe in subsection 6.2. For the welfare results in subsection 6.3 we compute the total consumer surplus calculated as the weighted sum of $\mathbb{E}(CS_i)$ using the weights reflecting the number of consumers who face the same representative utilities as the sampled consumer. That is:

$$\mathbb{E}(CV) = \int_{\nu_i} \left[\mathbb{E}(CS_i^1) - \mathbb{E}(CS_i^0) \right] \, \mathrm{d}P_\nu(\nu_i), \tag{16}$$

where $\mathbb{E}(CV)$ denotes the weighted sum across types of consumers of the compensating variation; the superscripts 0 and 1 refer, respectively, to before and after the counterfactual change in the services offered by the retailers; and $\mathbb{E}(CS_i)$ is given by the equation in (15). We describe the computation of the counterfactuals in next subsection.

[63] We are working to relax this assumption.

[64] The constant indicates that the absolute level of utility cannot be measured.

Social welfare is defined as the sum of consumer welfare, retailers' profits, and manufacturers' profits.

6.2 Counterfactual Scenarios

Three channels through which VSC retailers affect consumers' welfare in the outdoor advertising industry are by providing: (i) *consulting* services, (ii) *search* services, and (iii) *purchase aggregation* services.[65] We consider four counterfactual scenarios where we turn off each of these channels, and an additional one where we turn off all three simultaneously. For each counterfactual scenario we compute the compensating variation using equation (16). Below we describe how we compute each welfare scenario.

6.2.1 No Consulting Services

In this scenario, consumers may use the VSC, but VSC retailers do not offer *consulting services*, defined as the gross utility of given display format of a given manufacturer. In subsection 3.1, the gross utility of the consumer of purchasing a display format from a given manufacturer differed according to whether it was purchased through the VSC or DSC. We interpret the difference in gross utilities between the VSC and DSC retailers as the consulting services provided by the VSC retailers.

This counterfactual is implemented by imposing that the gross utility of consuming a display format from a given manufacturer purchased through the VSC is the same as gross utility of consuming the same display format of the same manufacturer purchased through the DSC. According to the equation in (1), in the demand model, the gross utility for display format d, produced by manufacturer m, and sold by retailer r, in market t, is given by:

$$\tau_{dmrt} \equiv \tau_d^D + \delta_m^D + \delta_r^D + \delta_t^D.$$

The "no consulting services scenario", denoted with the superscript "c_1", is implemented by changing each component of that vector that corresponds to purchases made trough the

[65] See subsection 2.1.

VSC such that: $\tau^{c_1}_{dmrt} = \tau_{dmmt}$, for every m, r and t.[66]

6.2.2 No Search Services

In this scenario, consumers may use the VSC, but VSC retailers do not offer *search services*, defined as consumers' access of price quotes and random shocks from all the manufacturers' products that offer the VSC retailer. In subsection 3.1, we assumed that if consumers search a retailer, they collect information about all the products sold by that retailer, so that the cost of searching for retailer r, sc_{irt}, in the VSC was:

$$sc_{irt} = \bar{S}_t + s^{\text{VSC}}_{rt}.$$

Because the VSC retailers sell the products from multiple manufacturers, searching for a VSC retailer allows consumers to collect the information about a larger set of products than searching for a DSC retailer.[67]

The "no search services scenario", denoted with the superscript "c_2", is implemented by imposing that the total search cost of visiting a VSC retailer is proportional to the number of manufacturers whose products the retailer carries. Specifically, we set the search cost to:

$$sc^{c_2}_{irt} = \bar{S}_t + \left(s^{\text{VSC}}_{rt}\right) \times |W_r|,$$

where $|W_r|$ is the number of manufacturers whose products the VSC retailer r carries.[68]

This counterfactual leaves everything constant, including the number of "stores" and display formats, and only eliminates the search cost advantage of buying through a retailer

[66] Value τ_{dmrt} represents the gross utility of a display format d, produced by manufacturer m, and sold to consumers by retailer r. In this counterfactual scenario, that product has the gross utility of a display format d, produced by manufacturer m, and sold to consumers by manufacturer m (*i.e.* a DSC retailer), δ_{dmmt}.

[67] This allowed us to rationalize the lower price dispersion observed in the VSC relative to the DSC (figures 2 and A2 discussed in subsubsection 2.3.3).

[68] *E.g.* suppose that retailer r sells products from manufacturers m and m'. In the counterfactual scenario without search services, the consumer has to pay twice the search cost of retailer r to obtain information about the display formats from these two manufacturers from this retailer, m and m', *i.e.*, $|W_r| = 2$.

instead of a manufacturer.[69] Each consumer will have the same consideration sets to choose from.

6.2.3 No Purchase Aggregation Services

In this scenario, we compare the equilibrium prices from the model in the supply side (*i.e.* bargaining game, then Nash Bertrand) with the ones of two successive Nash Bertrand. To implement it we follow the following four steps. First, we compute the optimal retail price function (as a function of any wholesale price) given by the equation in (9). Second, we use the result from step 1 to solve numerically for the pass through matrix defined by $\partial p_k / \partial \omega_r$ for all k, r similar to the equation in (9) in Villas-Boas (2007, p. 634). Third, we solve optimal wholesale price in the two margins model, using the equation in (11) with $\nu_{rmj} = 0$, and step 2 to get $\nabla_w \mathbf{s} \equiv \partial \mathbf{s} / \partial \omega' = \partial \mathbf{s} / \partial \mathbf{p} \times \partial \mathbf{p} / \partial \omega$. Finally, we solve for the optimal wholesale prices, ω, using the equation in (11) as obtained from step 3, which is similar to the expression in (9) in Draganska, Klapper, and Villas-Boas (2010, p. 62). This gives an expression that is a function of ω and $\mathbf{p}$ that can be solved for ω using the implicit function theorem applied to (11) because $\mathbf{p}^* = \mathcal{P}(\omega)$.

6.2.4 Consumers Only Purchase in the DSC

In this scenario, both VSC and DSC retailers operate, but VSC retailers do not offer either: consulting, nor search, nor purchase aggregation services, as defined above. This is implemented by implementing the three previous counterfactuals simultaneously. To evaluate the welfare under the different scenarios we simulate the choice outcomes predicted by the demand model.

[69] One possible interpretation for this counterfactual is the following. Imagine that initially a single product manager is in charge of all the display formats of all the manufacturers carried by a retailer. Under the counterfactual scenario, different managers are responsible for the display formats of different manufacturers. If a consumer wants to become informed about the prices of the display formats of the manufacturers carried by the retailer, he will now have to contact the various associated managers. Hence, paying a higher search cost.

6.3 Counterfactual Results

Table 6 reports the results from the counterfactual scenarios detailed above. We compare those results with the baseline predictions from the model. We report the following outcome variables, which shed light on the relative importance of each of the potential benefits of VSC retailers: fraction of consumers who purchase, number of visits to the retailer, number of searches, and the relative share of the DSC among the consumers who buy a product in the market.

[*Table* 6]

In the first counterfactual we remove the consulting services differential between the VSC and the DSC. We observe an increase in the total purchases, with a larger fraction coming from the direct channel. The reason behind these results is that the demand estimates show a larger gross utility for the average consumer from buying from the DSC. This is consistent, for example, with higher brand recognition of those firms.

In the second counterfactual, we remove the search services advantage of the VSC retailers. This causes a substantial decrease in the amount of purchases. In this counterfactual, approximately half of these purchases come from the DSC (while in the baseline case only 20 percent came from the DSC).

In the third counterfactual we remove the purchase aggregation services of the VSC retailers. Interestingly, the amount of total purchases does not decreases substantially. However, there is a large shift towards purchases in the DSC.

Finally, in the fourth counterfactual that is presented in the last column, we simultaneously remove all three services provided by the VSC retailers. We interpret the welfare results from this counterfactual as the impact that VSC retailers have on consumer welfare.

7 Concluding Remarks

In this paper we propose and estimate an empirical framework to quantify the welfare effects due to intermediation in vertical markets. The model includes a demand side, where consumers face costly search to choose among differentiated products. It also includes a supply side, where wholesalers and retailers bargain over the intermediate prices.

We apply our empirical framework to the Portuguese outdoor advertising industry. We recover the primitives of the industry, and simulate counterfactual scenarios to quantify the welfare of the different services that retailers offer. We find that the presence of intermediaries increases substantially the welfare in this industry, because the value of their services outweighs the additional margin charged.

Given that the outdoor advertising industry looks similar to other vertical markets in several dimensions (U.K. Office of Fair Trading 2011), the results of our policy studies can be used to learn about the effect of mergers in vertical markets, where intermediaries provide additional services to consumers. Recent examples in the U.S. include, *e.g.*, disputes between Tesla and Automobile Dealer Association, or the proposed merger of between Aetna and CVS.[70] In both cases, however, the value of intermediaries is related to the value of their services to the consumers.

Our empirical framework can be used to evaluate the efficiency and redistributive implications of vertical mergers, when intermediaries offer additional services to differentiate their products from the ones of the manufacturers. Our model combines features that are present in other industries, like the ones discussed in the previous paragraph. These features include consumers who have unobserved preferences that are specific to each distribution channel and engage in costly search on the demand side, and two layers of activity—where manufacturers and intermediaries bargain over wholesale prices—with two distribution channels—where the two distribution channels compete a la Bertrand—on the supply side.

References

Ahn, J., A. K. Khandelwal, AND S.-J. Wei (2011): "The role of intermediaries in facilitating trade," *Journal of International Economics*, 84(1), 73–85. (Cited on page 6.)

Arie, G., P. L. E. Grieco, AND S. Rachmilevitch (2016): "Generalized Insurer Bargaining," *Working Paper. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2677394 (Accessed on April 3, 2017.)*. (Cited on page 29.)

Armstrong, M. (2006): "Competition in two-sided markets," *The RAND Journal of Economics*, 37(3), 668–691. (Cited on page 6.)

[70] Regarding the U.S. car industry, in 48 states franchise laws prohibit or limit auto manufacturers from making sales directly to consumers, requiring sales to be done through the intermediation of car dealers. This has resulted in disputes between Tesla Inc. and state auto dealer associations (Sibilla 2017). Regarding the proposed merger between Aetna and CVS, one of the arguments in favor is that the merged CVS/Aetna would not need CVS/Caremark to function as intermediary, thus benefiting consumers by eliminating intermediaries' markup (Frakt 2017).

Baye, M. R., J. Morgan, and P. Scholten (2003): "The value of information in an online consumer electronics market," *Journal of Public Policy & Marketing*, 22(1), 17–25. (Cited on page 6.)

——— (2006): "Information, search, and price dispersion," *Handbook on economics and information systems*, 1, 323–77. (Cited on page 22.)

Bernard, C., and B. Jullien (2003): "Chicken and Egg: Competition among Intermediation Service Providers," *RAND Journal of Economics*, 34(2), 521. (Cited on page 6.)

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile prices in market equilibrium," *Econometrica*, pp. 841–890. (Cited on pages 3, 18, 20, and 31.)

Berry, S. T. (1994): "Estimating discrete-choice models of product differentiation," *The RAND Journal of Economics*, pp. 242–262. (Cited on pages 7, 20, and 33.)

Berry, S. T., and P. A. Haile (2014): "Identification in differentiated products markets using market level data," *Econometrica*, 82(5), 1749–1797. (Cited on page 33.)

Biglaiser, G. (1993): "Middlemen as experts," *The RAND journal of Economics*, pp. 212–223. (Cited on pages 5 and 6.)

Biglaiser, G., F. Li, C. Murry, and Y. Zhou (2017): "Middlemen as Information Intermediaries: Evidence from Used Car Markets," *Working Paper*. (Cited on page 6.)

Boik, A. (2016): "Intermediaries in Two-Sided Markets: An Empirical Analysis of the US Cable Television Industry," *American Economic Journal: Microeconomics*, 8(1), 256–282. (Cited on page 6.)

Bonnet, C., Z. Bouamra-Mechemache, and H. Molina (2016): "Vertical Relationships in Bilateral Oligopolies: A Structural Bargaining Model with Limited Data," *Working Paper. Available at: https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=EARIE43&paper_id=218 (Accessed on June 20, 2017.)*. (Cited on pages 29 and 30.)

Bonnet, C., and P. Dubois (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. (Cited on pages 7 and 26.)

Brenkers, R., and F. Verboven (2006): "Market definition with differentiated products: lessons from the car market," *in Choi, J.P. (Ed.), Recent Developments in Antitrust: Theory and Evidence*. (Cited on pages 7 and 26.)

Brown, J. R., and A. Goolsbee (2002): "Does the Internet make markets more competitive? Evidence from the life insurance industry," *Journal of political economy*, 110(3), 481–507. (Cited on page 6.)

Brunnermeier, M. K., and Y. Sannikov (2014): "A macroeconomic model with a financial sector," *The American Economic Review*, 104(2), 379–421. (Cited on page 6.)

Brynjolfsson, E., Y. Hu, and M. D. Smith (2003): "Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers," *Management Science*, 49(11), 1580–1596. (Cited on page 6.)

Brynjolfsson, E., and M. D. Smith (2000): "Frictionless commerce? A comparison of Internet and conventional retailers," *Management science*, 46(4), 563–585. (Cited on page 6.)

Busse, M., and M. Rysman (2005): "Competition and Price Discrimination in Yellow Pages Advertising," *RAND Journal of Economics*, pp. 378–390. (Cited on page 7.)

Cardell, N. S. (1997): "Variance components structures for the extreme-value and logistic distributions with application to models of heterogeneity," *Econometric Theory*, 13(02), 185–213. (Cited on pages 7 and 20.)

Chade, H., and L. Smith (2006): "Simultaneous search," *Econometrica*, 74(5), 1293–1307. (Cited on page 24.)

Chamberlain, G. (1982): "Multivariate regression models for panel data," *Journal of econometrics*, 18(1), 5–46. (Cited on page 33.)

Chu, C. S., P. Leslie, and A. Sorensen (2011): "Bundle-Size Pricing as an Approximation to Mixed Bundling," *American Economic Review*, 101, 263–303. (Cited on page 7.)

Ciliberto, F., and J. W. Williams (2014): "Does multimarket contact facilitate tacit collusion? Inference on conduct parameters in the airline industry," *The RAND Journal of Economics*, 45(4), 764–791. (Cited on page 7.)

Cohen, A. (2008): "Package size and price discrimination in the paper towel market," *International Journal of Industrial Organization*, 26(2), 502–516. (Cited on page 7.)

Collard-Wexler, A., G. Gowrisankaran, and R. S. Lee (2016): "'Nash-in-Nash' Bargaining: A Microfoundation for Applied Work," *NBER Working Paper No. 20641*. (Cited on page 29.)

Conlon, C. T., and N. Rao (2015): "The Price of Liquor is Too Damn High: Alcohol Taxation and Market Structure," *Working Paper. Available at: http://wagner.nyu.edu/files/faculty/publications/Liquor_Paper_Submission.pdf (Accessed on June 20, 2017.)*. (Cited on page 7.)

Crawford, G. S., R. S. Lee, M. D. Whinston, and A. Yurukoglu (2018): "The welfare effects of vertical integration in multichannel television markets," *Econometrica,* forthcoming. (Cited on page 7.)

Crawford, G. S., and A. Yurukoglu (2012): "The welfare effects of bundling in multichannel television markets," *The American Economic Review*, 102(2), 643–685. (Cited on pages 7 and 27.)

De los Santos, B., A. Hortaçsu, and M. R. Wildenbeest (2012): "Testing models of consumer search using data on web browsing and purchasing behavior," *The American Economic Review*, 102(6), 2955–2980. (Cited on pages 6, 7, 22, 23, and 24.)

Diamond, D. W. (1984): "Financial intermediation and delegated monitoring," *The Review of Economic Studies*, 51(3), 393–414. (Cited on page 6.)

Donna, J., and T. Pires (2016): "Quantity Discounts, Price Discrimination, and Incentive Compatibility Constraints," *Working Paper, The Ohio State University Discussion Papers. Available at: https://dl.dropboxusercontent.com/u/4794458/_WebPage/Papers/DP-Quantity-Discounts.pdf (Accessed on June 12, 2017.)*. (Cited on page 7.)

Draganska, M., D. Klapper, and S. B. Villas-Boas (2010): "A larger slice or a larger pie? An empirical investigation of bargaining power in the distribution channel," *Marketing Science*, 29(1), 57–74. (Cited on pages 7, 27, 29, 37, 38, 43, and A-12.)

Dubois, P., and M. Sæthre (2016): "On the Role of Parallel Trade on Manufacturers and Retailers Profits in the Pharmaceutical Sector," *Working Paper. Available at: https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=EEAESEM2016&paper_id=1349 (Accessed on June 12, 2017.)*. (Cited on page 7.)

Edelman, B., and J. Wright (2015): "Price coherence and excessive intermediation," *The Quarterly Journal of Economics*. (Cited on page 5.)

Ellison, G., and S. F. Ellison (2009): "Search, obfuscation, and price elasticities on the internet," *Econometrica*, 77(2), 427–452. (Cited on page 6.)

Ershov, D. (2018): "The effect of consumer search costs on entry and quality in the mobile app market," Discussion paper, Mimeo. Available at: https://drive.google.com/open?id=12LrT2boly9aBFeW6Vt5VEksPQnTSzyB5 (Accessed on April 6, 2018.). (Cited on page 6.)

Evans, D. S. (2003): "The antitrust economics of multi-sided platform markets," *Yale J. on Reg.*, 20, 325. (Cited on page 6.)

Farboodi, M., G. Jarosch, and G. Menzio (2017): "Intermediation as Rent Extraction," *NBER Working Paper No. 24171. Available at: https://www.nber.org/papers/w24171 (Accessed on January 2, 2017.).* (Cited on page 5.)

Frakt, A. (2017): "Why the CVS-Aetna Merger Could Benefit Consumers.," The New York Times, Dec. 3, 2017. Available at: https://www.nytimes.com/2017/12/03/upshot/why-the-giant-cvs-aetna-merger-could-benefit-consumers.html (accessed on Dec. 3, 2017). (Cited on page 45.)

Frankel, J. A., G. Galli, and A. Giovannini (2009): *The microstructure of foreign exchange markets*. University of Chicago Press. (Cited on page 6.)

Freixas, X., and J.-C. Rochet (2008): *Microeconomics of banking*. MIT press. (Cited on page 6.)

Galeotti, A., and J. L. Moraga-González (2009): "Platform intermediation in a market for differentiated products," *European Economic Review*, 53(4), 417–428. (Cited on page 6.)

Gandhi, A., and J.-F. Houde (2016): "Measuring substitution patterns in differentiated products industries," *Working Paper. Available at: https://pdfs.semanticscholar.org/63a9/b2cd63c4dc08524a6892a800bba302047fad.pdf (Accessed on June 20, 2017.).* (Cited on page 34.)

Gavazza, A. (2016): "An empirical equilibrium model of a decentralized asset market," *Econometrica*, 84(5), 1755–1798. (Cited on page 6.)

Gehrig, T. (1993): "Intermediation in search markets," *Journal of Economics & Management Strategy*, 2(1), 97–120. (Cited on page 5.)

Goeree, M. S. (2008): "Limited information and advertising in the US personal computer industry," *Econometrica*, 76(5), 1017–1074. (Cited on page 6.)

Gowrisankaran, G., A. Nevo, and R. Town (2015): "Mergers when prices are negotiated: Evidence from the hospital industry," *American Economic Review*, 105(1), 172–203. (Cited on pages 7, 27, and 37.)

Grennan, M. (2013): "Price discrimination and bargaining: Empirical evidence from medical devices," *The American Economic Review*, 103(1), 145–177. (Cited on pages 7, 8, 27, and 37.)

Grigolon, L., and F. Verboven (2014): "Nested logit or random coefficients logit? A comparison of alternative discrete choice models of product differentiation," *Review of Economics and Statistics*, 96(5), 916–935. (Cited on page 38.)

Haan, M., and J. L. Moraga-González (2011): "Consumer Search with Observable and Hidden Characteristics," *Working Paper. Available at: http://www.webmeets.com/files/papers/earie/2011/490/eariewauthors.pdf (Accessed on June 20, 2017.).* (Cited on page 22.)

Hagiu, A., and B. Jullien (2011): "Why do intermediaries divert search?," *The RAND Journal of Economics*, 42(2), 337–362. (Cited on page 5.)

Hansen, L. P. (1982): "Large sample properties of generalized method of moments estimators," *Econometrica*, pp. 1029–1054. (Cited on page 32.)

Hauser, J. R., and B. Wernerfelt (1990): "An evaluation cost model of consideration sets," *Journal of consumer research*, 16(4), 393–408. (Cited on page 21.)

Hausman, J. A. (1996): "Valuation of new goods under perfect and imperfect competition," in *The economics of new goods*, pp. 207–248. University of Chicago Press. (Cited on page 33.)

He, Z., and A. Krishnamurthy (2013): "Intermediary asset pricing," *The American Economic Review*, 103(2), 732–770. (Cited on page 6.)

Ho, K., and R. S. Lee (2017): "Insurer Competition in Health Care Markets," *Econometrica*, 85(2), 379–417. (Cited on page 7.)

Honka, E. (2014): "Quantifying search and switching costs in the US auto insurance industry," *The RAND Journal of Economics*, 45(4), 847–884. (Cited on pages 6, 22, and 34.)

Horn, H., and A. Wolinsky (1988): "Bilateral monopolies and incentives for merger," *The RAND Journal of Economics*, pp. 408–419. (Cited on page 29.)

James, C. (1987): "Some evidence on the uniqueness of bank loans," *Journal of financial economics*, 19(2), 217–235. (Cited on page 6.)

Janssen, M., and S. Shelegia (2015): "Consumer search and double marginalization," *American Economic Review*, 105(6), 1683–1710. (Cited on page 1.)

Lee, J., G. Gereffi, and J. Beauvais (2012): "Global value chains and agrifood standards: challenges and possibilities for smallholders in developing countries," *Proceedings of the National Academy of Sciences*, 109(31), 12326–12331. (Cited on page 6.)

Lerner, A. P. (1934): "The Concept of Monopoly and the Measurement of Monopoly Power," *The Review of Economic Studies*, 1(3), 157–175. (Cited on page 1.)

Lizzeri, A. (1999): "Information revelation and certification intermediaries," *The RAND Journal of Economics*, pp. 214–231. (Cited on page 6.)

Luco, F., and G. Marshall (2018): "Vertical Integration with Multiproduct Firms: When Eliminating Double Marginalization May Hurt Consumers," *Working Paper. Available at: https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID3110038_code2579292.pdf?abstractid=3110038&mirid=1 (Accessed on April 27, 2018.).* (Cited on page 1.)

McFadden, D. (1978): "Modeling the choice of residential location," *Transportation Research Record*, (673). (Cited on pages 24, A-7, A-9, and A-10.)

McManus, B. (2007): "Nonlinear pricing in an oligopoly market: the case of specialty coffee," *RAND Journal of Economics*, 38(2), 512–532. (Cited on page 7.)

Mehta, N., S. Rajiv, and K. Srinivasan (2003): "Price uncertainty and consumer search: A structural model of consideration set formation," *Marketing science*, 22(1), 58–84. (Cited on page 6.)

Miller, N. H., and M. C. Weinberg (2017): "Understanding the Price Effects of the MillerCoors Joint Venture," *Econometrica,* forthcoming. (Cited on pages 7 and 34.)

Miravete, E. J. (2002): "Estimating demand for local telephone service with asymmetric information and optional calling plans," *The Review of Economic Studies*, 69(4), 943–971. (Cited on page 7.)

Miravete, E. J., and L.-H. Röller (2004a): "Competitive Nonlinear Pricing in Duopoly Equilibrium: The Early U.S. Cellular Telephone Industry," *Working Paper. Available at: http://eugeniomiravete.com/papers/EJM-LHR.pdf (Accessed on April 11, 2017.).* (Cited on page 7.)

——— (2004b): "Estimating Markups under Nonlinear Pricing Competition," *Journal of the European Economic Association*, 2, 526–535. (Cited on page 7.)

Moraga-González, J. L., Z. Sándor, and M. R. Wildenbeest (2015): "Consumer search and prices in the automobile market," *Working Paper. Available at: http://www.kelley.iu.edu/mwildenb/searchautomobiles_old.pdf (Accessed on June 4, 2018.).* (Cited on pages 3, 6, 7, 20, 22, 23, 31, and 35.)

Morgan, P., and R. Manning (1985): "Optimal search," *Econometrica*, pp. 923–944. (Cited on page 22.)

Mortimer, J. H. (2008): "Vertical contracts in the video rental industry," *The Review of Economic Studies*, 75(1), 165–199. (Cited on page 7.)

Morton, F. S., F. Zettelmeyer, and J. Silva-Risso (2001): "Internet car retailing," *The Journal of Industrial Economics*, 49(4), 501–519. (Cited on page 6.)

Nevo, A. (2001): "Measuring market power in the ready-to-eat cereal industry," *Econometrica*, 69(2), 307–342. (Cited on pages 20 and 33.)

Nevo, A., J. L. Turner, and J. W. Williams (2016): "Usage-Based Pricing and Demand for Residential Broadband," *Econometrica*, 84(2), 411–443. (Cited on page 7.)

Newey, W. K., and D. McFadden (1994): "Large sample estimation and hypothesis testing," *Handbook of econometrics*, 4, 2111–2245. (Cited on page 32.)

O'Hara, M. (1995): *Market microstructure theory*, vol. 108. Blackwell Cambridge, MA. (Cited on page 6.)

Ohio Supercomputer Center (1987): "Ohio Supercomputer Center," http://osc.edu/ark:/19495/f5s1ph73. (Cited on page .)

Pereira, P., and T. Ribeiro (2018): "Evaluating Partial Divestitures When Vertical Relations are Important," *Review of Industrial Organization*. (Cited on pages 6 and 10.)

Pires, T. (2016): "Costly search and consideration sets in storable goods markets," *Quantitative Marketing and Economics*, 14(3), 157–193. (Cited on pages 7 and 34.)

Quan, T. W., and K. R. Williams (2016): "Product Variety, Across Market Demand Heterogeneity, and the Value of Online Retail," *Working Paper. Cowles Foundation Discussion Paper No. 2054.* (Cited on page 6.)

Roberts, J. H., and J. M. Lattin (1991): "Development and testing of a model of consideration set composition," *Journal of Marketing Research*, pp. 429–440. (Cited on page 6.)

Rochet, J.-C., and J. Tirole (2003): "Platform competition in two-sided markets," *Journal of the european economic association*, 1(4), 990–1029. (Cited on page 6.)

Rubinstein, A., and A. Wolinsky (1987): "Middlemen," *The Quarterly Journal of Economics*, 102(3), 581–593. (Cited on pages 1 and 5.)

Rust, J., and G. Hall (2003): "Middlemen versus market makers: A theory of competitive exchange," *Journal of Political Economy*, 111(2), 353–403. (Cited on page 5.)

Salinger, M. A. (1988): "Vertical mergers and market foreclosure," *The Quarterly Journal of Economics*, 103(2), 345–356. (Cited on page 1.)

Salz, T. (2017): "Intermediation and Competition in Search Markets: An Empirical Case Study," *Working Paper. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2961795 (Accessed on June 20, 2017.).* (Cited on page 6.)

Seiler, S. (2013): "The impact of search costs on consumer behavior: A dynamic approach," *Quantitative Marketing and Economics*, 11(2), 155–203. (Cited on page 34.)

Sibilla, N. (2017): "Connecticut Should Be Tesla Country," The New York Times, Jul. 7, 2017. Available at: https://www.nytimes.com/2017/07/07/opinion/connecticut-tesla-dealership-laws.html (accessed on Jul. 7, 2017). (Cited on page 45.)

Spengler, J. J. (1950): "Vertical integration and antitrust policy," *Journal of political economy*, 58(4), 347–352. (Cited on page 1.)

Spulber, D. F. (1995): "Market Making by Price Setting Firms," *Working paper, Northwestern University Discussion Papers.* (Cited on page 5.)

——— (1996): "Market microstructure and intermediation," *The Journal of Economic Perspectives*, 10(3), 135–152. (Cited on pages 1, 5, and 9.)

——— (1999): *Market microstructure: intermediaries and the theory of the firm*. Cambridge University Press. (Cited on pages 1, 5, and 9.)

Stanton, C. T., and C. Thomas (2016): "Landing the first job: The value of intermediaries in online hiring," *The Review of Economic Studies*, 83(2), 810–854. (Cited on page 6.)

Tirole, J. (1988): *The theory of industrial organization*. MIT press. (Cited on page 1.)

U.K. Office of Fair Trading (2011): "Outdoor Advertising: An OFT market study (OFT 1304, Feb. 2011)," https://www.lgcplus.com/Journals/3/Files/2011/2/3/oft1304.pdf (accessed on June 14, 2017). (Cited on page 45.)

U.S. Bureau of Economic Analysis (2017): "Value Added by Industry as a Percentage of Gross Domestic Product," https://www.bea.gov/iTable/iTableHtml.cfm?reqid=51&step=51&isuri=1&5114=a&5102=5 (accessed on June 14, 2017). (Cited on page 1.)

Villas-Boas, S. B. (2007): "Vertical relationships between manufacturers and retailers: Inference with limited data," *The Review of Economic Studies*, 74(2), 625–652. (Cited on pages 7, 26, and 43.)

Wolinsky, A. (1986): "True monopolistic competition as a result of imperfect information," *The Quarterly Journal of Economics*, 101(3), 493–511. (Cited on page 22.)

Wright, R., and Y.-Y. Wong (2014): "Buyers, Sellers, And Middlemen: Variations On Search-Theoretic Themes," *International Economic Review*, 55(2), 375–397. (Cited on page 5.)

Yanelle, M.-O. (1989): "The strategic analysis of intermediation," *European Economic Review*, 33(2-3), 294–301. (Cited on page 5.)

Table 1: Sales Percentage to Consumers by Manufacturer, Retailer, and Product.

Panel A: Sales to Consumers by Manufacturer and Retailer (as percentage of total sales in the sample)

	Seller	Manufacturer m_1	m_2	m_3	m_4	Total
VSC Retailers	r_4^v	1.06	2.34	0.63	2.34	6.36
	r_5^v	0.54	1.24	0.66	0.96	3.41
	r_6^v	3.31	5.35	1.26	5.32	15.25
	r_7^v	1.43	4.61	1.68	13.53	21.25
	r_8^v	0.18	1.36	0.37	–	1.91
	r_9^v	6.97	23.86	5.86	0.30	37.00
DSC Retailers	r_1^d	1.51	–	–	–	1.51
	r_2^d	–	8.79	–	–	8.79
	r_3^d	–	–	4.52	–	4.52
Total		15.01	47.56	14.98	22.45	100.00

Panel B: Sales to Consumers by Manufacturer, Retailer, and Display Type (as percentage of total sales in the sample)

	Seller	2 m^2 panel m_1	m_2	m_3	m_4	Senior m_1	m_2	m_3	m_4	Other m_1	m_2	m_3	m_4	Total 3	Total 4
VSC Retailers	r_4^v	1.06	0.94	0.24	–	–	1.40	0.10	–	–	–	0.29	2.34	6.36	85.18
	r_5^v	0.54	1.10	0.28	–	–	0.14	0.22	0.17	–	–	0.16	0.79	3.41	
	r_6^v	3.31	3.20	0.97	1.39	–	2.15	0.19	–	–	–	0.10	3.93	15.25	
	r_7^v	1.43	3.42	0.68	0.28	–	1.06	0.45	–	–	0.13	0.55	13.25	21.25	
	r_8^v	0.18	0.22	–	–	–	1.10	0.16	–	–	0.04	0.21	–	1.91	
	r_9^v	6.93	17.55	3.79	0.10	0.03	5.82	0.53	0.15	0.01	0.49	1.55	0.06	37.01	
DSC Retailers	r_1^d	1.51	–	–	–	–	–	–	–	–	–	–	–	1.51	14.83
	r_2^d	–	5.10	–	–	–	3.49	–	–	–	0.21	–	–	8.80	
	r_3^d	–	–	1.61	–	–	–	0.21	–	–	–	2.70	–	4.52	
Total 1		14.97	31.54	7.56	1.77	0.03	15.17	1.86	0.32	0.01	0.87	5.56	20.36	100	
Total 2		55.84				17.37				26.80					100

Notes: Each cell in Panels A and B corresponds to the percentage of sales to consumers (relative to the total sales' volume to consumers sold in year 2013 in the whole sample) by the corresponding combination of: (1) Manufacturer and Seller in Panel A and; (2) Manufacturer, Seller, and Display Format in Panel B. Thus, in each panel, all the numbers sum to 100 (excluding the rows and columns labeled as "Total"). A cell displays the symbol "–" when no sales are observed for such combination. In Panel B there are a total of 57 cells with positive sales (*i.e* without the symbol "–'), that corresponds to the 57 inside products (see footnote 24 for details). In panel B, "Total 1" refers to the total sum by manufacturer m_i, $i = 1, \ldots, 4$; "Total 2" refers to the total by display format (2 m^2 panel, Senior, and Other); "Total 3" refers to the total sum by retailer r_j^v, $j = 4, \ldots, 9$ and r_j^d, $j = 1, 2, 3$; "Total 4" refers to the total by VSC Retailers (*i.e.* sum over r_j^v, $j = 1, \ldots, 6$) and by DSC Retailers (*i.e.* sum over r_j^d, $j = 1, 2, 3$). "DSC" stands for Direct Sales Channel. "VSC" stands for Vertical Sales Channel. See figure 1 for definitions of prices and vertical relations in the industry.

Table 2: Wholesale and Retail Prices in the VSC.

Panel A: All Manufacturers and All VSC Retailers

Display Format	Wholesale Price					VSC Price (Retail Price)				
	Median	Mean	SD	Min	Max	Median	Mean	SD	Min	Max
2 m^2 panel	8.40	12.18	12.93	0.81	93.38	9.11	13.28	14.13	0.86	99.39
Senior	12.84	17.11	18.35	0.81	158.06	13.50	18.14	19.59	0.83	165.80
Other	24.62	28.61	26.02	0.34	148.60	26.91	31.70	29.46	0.37	171.99

Panel B: By Manufacturer and All VSC Retailers

Display Format	Manu-facturer	Wholesale Price					VSC Price (Retail Price)				
		Median	Mean	SD	Min	Max	Median	Mean	SD	Min	Max
2 m^2 panel	m_1	8.29	9.41	7.25	1.42	51.64	8.92	10.62	8.15	1.50	56.18
	m_2	10.79	12.64	10.73	2.33	64.80	11.46	13.78	12.54	2.41	83.22
	m_3	6.17	8.57	11.66	1.04	79.45	6.78	9.32	12.65	1.09	81.52
	m_4	12.38	19.80	18.88	0.81	93.38	13.29	21.10	20.22	0.86	99.39
Senior	m_1	16.09	15.17	4.45	9.07	24.13	16.78	15.90	4.54	9.46	25.17
	m_2	6.32	10.82	14.12	0.81	98.39	6.51	11.17	14.49	0.83	100.57
	m_3	8.80	21.73	28.76	2.09	158.06	9.46	23.86	30.98	2.17	165.80
	m_4	18.71	21.07	10.25	6.11	46.45	19.40	22.03	10.85	6.49	53.85
Other	m_1	48.69	42.41	24.94	3.61	83.26	52.58	48.05	29.76	5.09	106.60
	m_2	34.31	36.73	18.19	0.34	117.09	36.97	38.83	18.99	0.37	120.03
	m_3	13.71	28.60	33.88	0.45	148.60	15.35	32.66	39.28	0.64	171.99
	m_4	13.48	15.44	14.73	0.38	57.28	14.44	17.59	16.88	0.41	66.93

Notes: Panel A reports summary statistics of wholesale and VSC prices (*i.e.* retail prices) for each display format (2 m^2 panel, Senior, and Other) across manufacturers (m_1, m_2, m_3, and m_4) and VSC retailers ($r_4^v, r_5^v, \ldots, r_9^v$), respectively. Panel B reports summary statistics of wholesale and VSC prices across all VSC retailers for each combination of display format and manufacturer. See table A1 for a comparison of summary statistics of wholesale and VSC prices by manufacturer and by VSC retailer for the display format 2 m^2 panel. Similar tables for the other display formats (seniors and others) are available upon request. "DSC" stands for Direct Sales Channel. "VSC" stands for Vertical Sales Channel. See figure 1 for definitions of prices and vertical relations in the industry.

Table 3: Price Paid by Consumers to in the DSC and VSC.

Panel A: By Display Format, All Manufacturers

Display Format	Sale's Channel	Median	Mean	SD	Min	Max
2 m^2 panel	DSC	10.05	11.97	10.82	1.67	66.90
	VSC	9.11	13.28	14.13	0.86	99.39
Senior	DSC	13.69	14.55	6.67	6.30	40.44
	VSC	13.50	18.14	19.59	0.83	165.80
Other	DSC	5.85	14.74	16.65	1.36	63.62
	VSC	26.91	31.70	29.46	0.37	171.99

Panel B: By Display Format and by Manufacturer

Display Format	Manufacturer	Sale's Channel	Median	Mean	SD	Min	Max
2 m^2 panel	m_1	DSC	11.84	12.73	2.91	9.38	19.04
		VSC	8.92	10.62	8.15	1.50	56.18
	m_2	DSC	13.65	16.85	17.37	1.67	66.90
		VSC	11.46	13.78	12.54	2.41	83.22
	m_3	DSC	6.88	6.83	2.20	2.67	10.16
		VSC	6.78	9.32	12.65	1.09	81.52
Senior	m_1	DSC	13.97	14.19	2.67	11.65	19.04
		VSC	16.78	15.90	4.54	9.46	25.17
	m_2	DSC	14.85	14.60	4.10	7.87	21.31
		VSC	6.51	11.17	14.49	0.83	100.57
	m_3	DSC	9.97	14.67	9.25	6.30	40.44
		VSC	9.46	23.86	30.98	2.17	165.80
Other	m_2	DSC	18.50	24.04	21.04	1.36	63.62
		VSC	36.97	38.83	18.99	0.37	120.03
	m_3	DSC	5.46	6.78	3.22	3.81	14.33
		VSC	15.35	32.66	39.28	0.64	171.99

Notes: The table reports summary statistics of the price paid by consumers on sales made in the DSC and in the VSC (column labeled "Sale's Channel"). Panel A reports the summary statistics by display format (2 m^2 panel, Senior, and Other). DSC prices refer to the prices over all sales from manufacturers to consumers in that display format. VSC prices refer to the prices over all sales from retailers that are not manufacturers for that display format. Panel B reports the summary statistics by display format and by manufacturer. DSC prices in Panel B refer to the manufacturer in each cell (*i.e.* there is only one manufacturer in each of these cells). VSC prices refer to the prices over all sales from retailers that are not manufacturers for the display format and manufacturer in the relevant cell. Manufacturer m_4, that corresponds to the additional manufacturer that aggregates smaller manufacturers, is not included because it does not perform any sale to the consumers directly (*i.e.* manufacturer m_4 does not participate in the DSC). Similarly manufacturer m_1 is not included for the display format "Other" because it does not perform any sale to the consumers directly. "DSC" stands for Direct Sales Channel. "VSC" stands for Vertical Sales Channel. See figure 1 for definitions of prices and vertical relations in the industry.

Table 4: Quantity Discounts in the VSC, but not in the DSC.

Price paid by consumers per m^2	(1)	(2)	(3)	(4)
Log(m^2)	-7.0708***	-1.8348	-6.9948***	-1.5502
	(0.4472)	(1.2105)	(0.4511)	(1.1810)
Log(m^2) $\times$ VSC		-6.0297***		-6.2510***
		(1.2990)		(1.2576)
Manufacturers Fixed Effects	Yes	Yes	Yes	Yes
Retailers Fixed Effects	Yes	Yes	Yes	Yes
Display Formats Fixed Effects	No	No	Yes	Yes
Months Fixed Effects	No	No	Yes	Yes
Adjusted R^2	0.4081	0.4291	0.4493	0.4723
Number of Observations	570	570	570	570

Notes: All regressions are OLS specifications. The sample is the same sample used for the structural estimation, and it corresponds to all purchases of all display formats made by the consumers in the industry. The dependent variable is the price paid by consumers per square meter of advertising in a given month, labeled as "Price paid by consumers per m^2." The variable "Log(m^2)" corresponds to the total square meters of advertising purchased by consumers in that month on a logarithmic scale. The variable "VSC" is a dummy variable that equals 1 if the consumer performed the purchase through a VSC retailer, and 0 if the consumer performed the purchase through the DSC retailer. "DSC" stands for Direct Sales Channel. "VSC" stands for Vertical Sales Channel. Standard errors are in parentheses. *p<0.10; **p<0.05; ***p<0.01.

Table 5: Demand Model Estimates

	Model 1		Model 2		Model 3	
	Coeff	SE	Coeff	SE	Coeff	SE
Price:						
Mean (α)	-0.005	0.002	-0.031	0.006	-0.063	0.017
Std dev (σ)			0.027	0.000	0.027	0.012
Firm Dummies:						
Wholesaler 1	-0.037	0.032	-0.129	0.085	-0.316	0.231
Wholesaler 2	0.069	0.040	0.084	0.105	0.502	0.287
Wholesaler 3	-0.031	0.026	-0.165	0.068	-0.234	0.186
Retailer 1	-1.538	0.096	-1.264	0.251	1.576	0.683
Retailer 2	-0.055	0.035	-0.127	0.093	1.584	0.252
Retailer 3	-0.057	0.037	-0.163	0.098	1.276	0.267
Retailer 4	-1.591	0.078	-1.273	0.205	1.487	0.559
Retailer 5	-0.035	0.029	-0.119	0.077	1.900	0.209
Retailer 6	-1.582	0.083	-1.291	0.218	1.416	0.593
Retailer 7	-0.036	0.029	-0.061	0.076	2.191	0.207
Retailer 8	-0.129	0.069	-0.338	0.182	0.704	0.496
Product Dummies:						
2 m^2 panel	0.032	0.027	0.149	0.070	0.264	0.192
Senior	-0.059	0.034	-0.086	0.088	-0.434	0.239
Nest Parameter	0.920	0.039	0.788	0.102	0.348	0.279
Search Parameters:						
Constant					0.589	0.131
Previous Market Share					2.294	86.153
Objective function						
N	570		570		570	

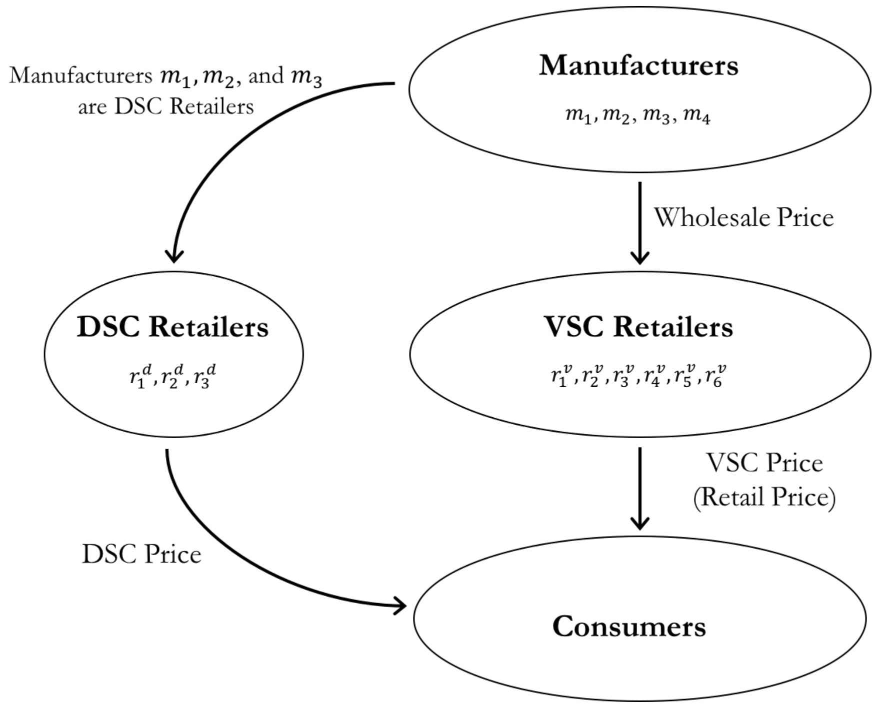
Notes: This table presents the structural estimates of our demand models. Model 1 is the simplest and restricts the price coefficient to be homogeneous across consumers. In Model 2 we introduce a random coefficient on prices. In model 3 we add both random coefficient on prices and a search cost function. All our specifications include wholesaler, retailer, product type and month fixed effects.

Table 6: Counterfactual Results

Variable	**Baseline**	**CF1 (no consult)**	**CF2 (no search)**	**CF3 (no QD)**	**CF4 (cf1+cf2+cf3)**
Inside Share	61.6%	58.1%	61.6%		
DSC fraction (of inside)	18.2%	22.7%	18.1%		
Mean price	16.85	16.85	16.85		
Mean price (weight)	8.76	9.34	8.76		
Number of Visits	5.12	5.11	5.12		
Count of Search Costs	5.12	5.11	15.48		
Δ Cons Surplus (euros/m2)		-3.79	-0.92		

Notes: This table presents the results of the simulation of our model in the baseline scenario, as well as under each of the counterfactual scenarios described in section 6. In the first row, we report the fraction of the total potential size of the market that resulted in actual sales. The second row reports the fraction of those actual sales that were made through the Direct Channel (DSC). Finally, in the third and forth rows, we report the average number of retail visits and the average number of searches, respectively.

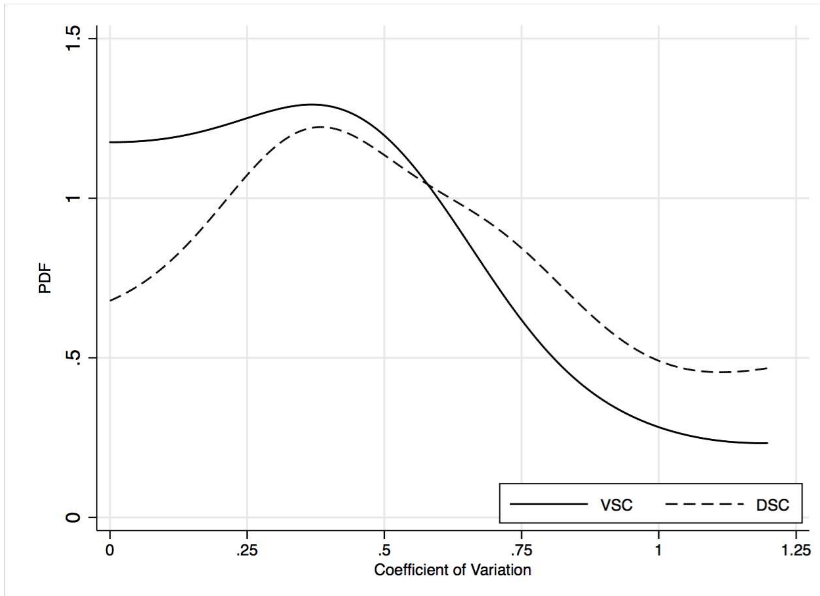
Figure 1: The Portuguese Outdoor Advertising Industry.



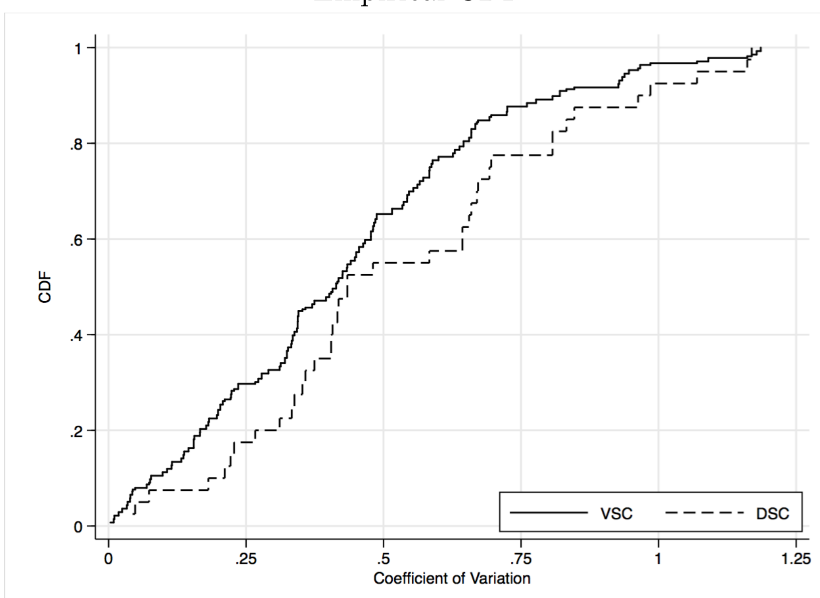
Notes: The figure displays the vertical relations in the Portuguese outdoor advertising industry. "DSC" stands for Direct Sales Channel. "VSC" stands for Vertical Sales Channel. The manufacturers $(m_1, \ldots, m_4)$ sell their products to the VSC retailers $(r_4^v, \ldots, r_9^v)$ charging wholesale prices. The VSC retailers sell to consumers, charging VSC prices (or retail prices). The three main manufacturers (m_1, m_2, m_3) also sell directly to the consumers through the DSC; this is captured in the diagram by the DSC retailers (r_1^d, r_2^d, r_3^d), which correspond to the large manufacturers charging a DSC prices to the consumers.

Figure 2: Distribution of Coefficient of Variation.

Density Estimate



Empirical CDF



Notes: The figure displays the kernel density estimate (top panel) and empirical cumulative distribution (bottom panel) of coefficient of variation of prices (CV) for sales to consumers in the VSC and DSC, conditional on quantity discounts. To perform the estimation we proceed in three steps. First, we define the unit of analysis as a tuple (Display Format, Month, Volume Percentile), where "Display Format" are the display formats as defined in subsection 2.1, "Month" are the months of the year, and "Volume Percentile" are the percentiles in the volume variable (to account for quantity discounts). Second, for each unit of analysis (*i.e.* tuple as defined above) we compute the CV (*i.e.* the variation of prices is within tuple). Third, we estimate the kernel density and empirical cumulative distribution as follows. Let cv denote realized CV in each tuple $j \in \{1, \ldots, J\}$. We estimate the probability density function for sales made to consumers through retailers and manufacturers, $f(cv)$, as: $\hat{f}_K(cv; h) = \frac{1}{J\,h} \sum_{j=1}^{J} K\left(\frac{cv - cv(j)}{h}\right)$, where $K(z)$ is a standard univariate gaussian kernel function, h is the bandwidth that we choose by cross validation, and $cv(j)$, $j = 1, \ldots, J$ are the CV in each tuple. Given that the price distribution has its domain bounded we use a renormalization method to deal with the boundaries when estimating the probability density function of CV. We estimate the empirical cumulative distribution of CV, $F(cv)$, as: $\hat{F}_J(cv) = \frac{1}{J} \sum_{j=1}^{J} \mathbf{1}\{cv(j) \leq cv\}$, where $\mathbf{1}\{A\}$ is the indicator function of the event A. "DSC" stands for Direct Sales Channel. "VSC" stands for Vertical Sales Channel. See figure 1 for definitions of prices and vertical relations in the industry.

Appendix

A Data Appendix

A.1 Procedures to Clean the Data

Below we describe the procedure we use to clean the raw data about the manufacturers and retailers.

We collected data from 5 retailers. The information collected from the retailers includes for each product and for each month of the year 2013: (i) the sales values, (ii) the quantity sold in number of advertising faces and m^2, (iii) the cost of the products transacted, (iv) the commissions, fees, and rebates received and paid. A product is a combination of (i) display format, (ii) manufacturers, and (iii) retailer. The data collected from retailers consists of 4807 observations (*i.e.* pairs product-month). For our analysis we combine manufacturers other than JCD, Cemusa, and Mop in a composite manufacturer. We further combine display formats other than MUPIs, Seniors, and FEs. We then aggregate the data so that an observation is a tuple manufacturer-retailer-display format-month. The number of observations after all these procedures is 726.

We complement the retailer data with information for direct sales of manufacturers (*i.e.*, sales of manufacturers to consumers) and sales of manufacturers to retailers other than the 5 retailers for which we collected data. This information was obtained from a survey to 3 manufacturers (JCD, Cemusa, and Mop) and includes the same information previously described for retailers. This additional data includes 176 observations and thus the combined data consists of 902 observations.

We dropped observations with (i) a ratio of median absolute deviation (MAD) of m^2 sold (in logs) to the standard deviation larger than 3 (11 observations dropped),[71] (ii) a ratio of MAD of wholesale price (in logs) to the standard deviation larger than 3 (12 observations dropped), (iii) a ratio of MAD of retail price (in logs) to standard deviation larger than 3 (6 observations dropped). We further aggregate all monthly sales through groups that not the

[71] The median and standard deviation are always specific to each display format.

ones that we surveyed in a single product (54 observations collapsed).

A.2 Google Trends Data

We use Google trends data to generate instruments that we use to construct the moment conditions that identify the search parameters on the demand side. The Google trends data were obtained directly from Google Trends Portugal from the link below, and accessed on August 30, 2017:

https://trends.google.com/trends/explore?date=2013-01-01%202013-12-31&geo=PT

The raw Google trends data corresponds to weekly searches, for the period under study, in Google Portugal (https://www.google.pt/) of the names of the retailers spelled as follows: cemusa, havas, ipg, jcdecaux, mop, Omnicom, pmg, and wpp. These retailers correspond, respectively, to the following retailers in the data: Cemusa, Havas, Megameios, JCDecaux, Mop, Opusopera, Powermedia, and Group M. For robustness, we performed the search in Google Portugal for each of these retailers with the names spelled as above, and in all cases the first result displayed by Google Portugal was the webpage of the retailer. This gives us confidence about the "power" of the instrument in that consumers who performed such searches in Google Portugal obtained the relevant information about the retailers.

We use these raw data to generate the instrumental variables defined next.

Google Trends Variables Definitions

- Mean weekly searches current month. Mean per month of weakly searches using the current month.
- Mean weekly searches previous month. Mean per month of weakly searches using the previous month.
- Mean weekly searches previous 3 months. Mean per month of weakly searches using the mean over previous 3 months.
- Mean weekly searches previous 6 months. Mean per month of weakly searches using the mean over previous 6 months.
- Mean weekly searches moving average. Mean per month of weakly searches using a three period moving average as adaptive expectation.
- Standard deviation weekly searches current month. Standard deviation per month of weakly searches using the current month.
- Standard deviation weekly searches previous 3 month. Standard deviation per month of weakly searches using the previous 3 months.
- Total searches current month. Total searches per month using the current month.
- Total searches previous month. Total searches per month using the previous month.
- Total searches previous 3 months. Total searches per month using the mean of the previous 3 months.
- Total searches previous 6 months. Total searches per month using the mean of the previous 6 months.
- Total searches moving average. Total searches per month using a three period moving average as adaptive expectation.

The idea behind the variables using moving averages is to capture an adaptive expectation of what the searches will be based on information on these lags. That is, they are a proxy for what could consumers forecast for the current visibility of a retailer based on the past values, using the fixed weighting scheme determined by the moving average.

B Additional Description of the Industry and the Data

B.1 Market Concentration.

The Portuguese outdoor advertising market is quite concentrated both at the manufacture and retail levels. At the manufacture level there are three large national firms, m_1, m_2, and m_3, that are responsible of 77.6 percent of the sales in the market (Panel A in table 1). The other small local manufacturers are responsible of the reminder of the sales. Manufacturer m_2 is the largest manufacturer with 47.6 percent of the sales. At the retail level the five largest VSC retailers, *i.e.* retailers $r_4^v, \ldots, r_8^v$, are responsible of 48.2 percent of the sales. Retailer r_7^v is the largest retailer with 21.2 percent of the sales, and also larger than the DSC retailers. The most popular display format are 2 m^2 panels, that encompass 55.8 percent of the sales (Panel B in table 1). The largest manufacturer, m_2, is responsible of 56.5 percent ($^{31.5}/_{55.8}$) of the sales of 2 m^2 panels in the market. The largest VSC retailer, r_7^v, is responsible of 10.4 percent ($^{(1.4+3.4+0.7+0.3)}/_{55.8}$) of the sales of 2 m^2 panels in the market. There is no cross-ownership between manufacturers, nor between retailers, nor between manufacturers and retailers. Geographically, all manufacturers and retailers operate in the same market. This follows from Portugal being a small country, where the population is concentrated along the coast.

B.2 Seasonalities and Monthly Variation

Market Shares and Quantities. The top panel in figure A1 shows monthly seasonal variations in total volume and substantial variation the market shares within month. The top panel shows the total volume sold each month (right vertical axis) distinguishing the sales to consumers in the VSC and DSC. The total volume purchased in the Portuguese advertising industry increases during the summer. Total volume varies substantially by month, reflecting the monthly variation in the purchases made in the VSC. The figure also shows the distribution of products' market shares (left vertical axis) for each month distinguishing sales to consumers in the VSC and DSC. There is large variation in market shares both in the VSC and DSC.

Seasonal variations in the Portuguese advertising industry have two main components. First, a deterministic component, whereby the demand for advertising increases during certain months of the year (*e.g.* summer vacation). Second, a non deterministic component, whereby individual firms make specific advertising decisions on certain months of the year based on their needs (*e.g.* firms launching a new product in September, or advertising in December before Christmas, *etc.*). The deterministic component explains the increase in advertising during the summer. The non deterministic component explains the monthly volatility.

Prices. The bottom panel in figure A1 shows that the within monthly variation in prices is larger than the variation across months. The bottom panel displays the distribution of prices paid by consumers each month distinguishing sales to consumers in the VSC and DSC. The distribution of prices of sales made in the VSC is more disperse than the one in the DSC. This is due to the presence of quantity discounts on the sales made in the VSC that are not present in the DSC (table 4). Conditional on quantity discounts, the distribution of prices of sales made in the VSC is less disperse as discussed in the paper.

[*Figure A1*]

C Random Coefficient Nested Logit Model with Search

In this appendix we compute the choice probabilities, maximum expected value, and welfare for the nested logit random coefficient model with search. To facilitate the reading we repeat the notation of the model in subsection C.1.

C.1 Notation review

The indirect utility of consumer i for inside product j in market t, conditional on the set of retailers searched, R_i, is:

$$U_{ijt|R_i} = -\alpha_i p_{jt} + x_{jt}\beta + \tau_d^D + \tau_m^D + \tau_r^D + \tau_t^D + \xi_{jt} + \hat{\varepsilon}_{ijt},$$

$$i = 1, \ldots, I_t, \qquad j \in \hat{J}_{tR_i} = \{\hat{j} : \hat{j} \in J_t \text{ is sold by retailer } r \in R_i\} \cup \{0\}, \qquad t = 1, \ldots, T,$$

where R_i denotes the subset of retailers searched; $\hat{J}_{tR_i}$ is the consideration set of consumer i, given by the subset of products sold by all the retailers searched and the outside product; p_{jt} is the price of product j in market t; x_{jt} is a S-dimensional (row) vector of observable characteristics of product j in market t; τ_d^D, τ_m^D, τ_r^D, and τ_t^D capture the preferences for display format d, manufacturer m, retailer r, and monthly seasonal effects in market t, using fixed dummy variables for display format, manufacturer, retailer, and monthly seasonal effects, respectively; ξ_{jt} is the valuation of unobserved (by the econometrician) characteristics of product j in market t; $\hat{\varepsilon}_{ijt} = \zeta_{igt} + (1-\lambda)\varepsilon_{ijt}$, is a stochastic term; $g \in \{0, 1, 2\}$ define three groups (or nests) of nonoverlapping products for the outside product (denoted $g = 0$ with only one product), the products sold by the DSC retailers (denoted $g = 1$), and the products sold by the VSC retailers (denoted $g = 2$); ζ_{igt} has a unique distribution such that $\hat{\varepsilon}_{ijt}$ is extreme value; λ is a nesting parameter such that $0 \leq \lambda < 1$; $\alpha_i = \alpha + \Sigma\nu_i,\ \nu_i \sim P_\nu(\nu_i) = \mathcal{N}(0,1)$, are individual-specific parameters that capture consumers' preferences for price; and β is a S-dimensional vector of parameters. In each market t, we normalize the characteristics of the outside product, $j = 0$, such that $p_{0kt} = x_{0t} = \tau_0^D = \tau_0^D = \tau_0^D = \tau_t^D = \xi_{0t} = 0$ for all t. Denote by $\overline{U}_{ijt} \equiv -\alpha_i p_{jt} + x_{\hat{q}t}\beta + \tau_d^D + \tau_m^D + \tau_r^D + \tau_t^D + \xi_{jt}$, and by $\delta_{jt} \equiv -\alpha p_{jt} + x_{jt}\beta + \tau_d^D + \tau_m^D + \tau_r^D + \tau_t^D + \xi_{jt}$. Note that $\overline{U}_{ijt} = \delta_{jt} - \Sigma\nu_i p_{jt}$.

C.2 Maximum expected value

Following McFadden (1978),[72] the expected maximum utility conditional on the set of retailers R_i searched is:[73]

$$\mathbb{E}\left[\max_{j\in\hat{J}_t}\left(\overline{U}_{ijt}+\hat{\varepsilon}_{ijt}\right)\right]=\log\ \sum_{\tilde{g}=0}^{2}\left[\sum_{\tilde{j}\in\tilde{g}}\left(e^{\overline{U}_{i\tilde{j}t}}\right)^{1/1-\lambda}\right]^{1-\lambda}+\hat{\gamma},$$

where $\mathbb{E}(\cdot)$ is the expectation operator taken over $\hat{\varepsilon}_{ijt}$; $\log(\cdot)$ denotes the natural logarithm function; and $\hat{\gamma}=0.5772$ is the Euler's constant.

Denote by $I_{\hat{g}R_i}$ the *inclusive value* of the set of products from the searched retailers that belong to subset $\hat{g}$ excluding the outside product:

$$I_{\hat{g}R_i}\equiv(1-\lambda)\log\sum_{j\in(\hat{g}\cap\hat{J}_{tR_i}\setminus\{0\})}e^{\overline{U}_{ijt}/(1-\lambda)},\quad \hat{g}\in\{1,2\}. \tag{C.1}$$

Then:

$$\begin{aligned}\mathbb{E}\left[\max_{j\in\hat{J}_t}\left(\overline{U}_{ijt}+\hat{\varepsilon}_{ijt}\right)\right]&=\log\ \sum_{\tilde{g}=0}^{2}\left[\sum_{j\in\tilde{g}}\left(e^{\overline{U}_{ijt}}\right)^{1/1-\lambda}\right]^{1-\lambda}+\hat{\gamma},\\ &=\log\ \sum_{\tilde{g}=0}^{2}\left[e^{log\left(\sum_{j\in\tilde{g}}e^{\overline{U}_{ijt/1-\lambda}}\right)}\right]^{1-\lambda}+\hat{\gamma},\\ &=\log\ \sum_{\tilde{g}=0}^{2}e^{I_{\tilde{g}R_i}}+\hat{\gamma},\\ &=\log\ \left(1+\sum_{\hat{g}=1}^{2}e^{I_{\hat{g}R_i}}\right)+\hat{\gamma},\end{aligned} \tag{C.2}$$

where the last equality follows because the inclusive value of the outside product is equal to zero.

Then the expected net value for consumer i of searching a subset of retailers R_i in market

[72] Corollary to theorem 1 on pages 82-3, equations (14) and (17).

[73] For the case of the logit model, where ε_{ijt} is a standardized type I extreme value, this expression specializes to $\mathbb{E}\left[\max_{j\in\hat{J}_t}\left(\overline{U}_{ijt}+\varepsilon_{ijt}\right)\right]=\log\sum_{j\in\hat{J}_i}e^{\overline{U}_{ijt}}$, a well known result.

t is:

$$V_{tR_i} = \int \max_{j \in \hat{J}_i} U_{ijt} \, \mathrm{d}F_{\hat{\varepsilon}}(\hat{\varepsilon}) \, \mathrm{d}F_p(p) - SC_{itR_i} + \tilde{\varepsilon}_{itR_i} =$$

$$\int \log \left(1 + \sum_{\hat{g}=1}^{2} e^{I_{\hat{g}R_i}} \right) \, \mathrm{d}F_p(p) + \hat{\gamma} - SC_{itR_i} + \tilde{\varepsilon}_{itR_i},$$

where SC_{itR_i} is the cost of searching the subset of retailers R_i; $\tilde{\varepsilon}_{itR_i}$ is a random shock to the subset of searched retailers that is drawn $i.i.d.$ from a type I extreme value distribution with location parameter $\mu_{\tilde{\varepsilon}} = 0$ and scale parameter $\sigma_{\tilde{\varepsilon}} > 0$; and the last equality follows from the expression in equation (C.2) with the inclusive value, $I_{\hat{g}R_i}$, defined by the equation in (C.1).

C.3 Choice probabilities

The probability that consumer i chooses product $\hat{j}$ in group $\hat{g}$ conditional on the searched retailers, R_i, denoted by $P_{i\hat{j}t|R_i}$, is given by:

$$P_{i\hat{j}t|R_i} = \frac{\exp\left(\frac{\overline{U}_{i\hat{j}t}}{(1-\lambda)}\right)\left[\sum_{\tilde{j}\in g}\exp\left(\frac{\overline{U}_{i\tilde{j}t}}{1-\lambda}\right)\right]^{-\lambda}}{\sum_{\tilde{g}=0}^{2}\left[\sum_{\tilde{j}\in\tilde{g}}\exp\left(\frac{\overline{U}_{i\tilde{j}t}}{1-\lambda}\right)\right]^{1-\lambda}}, \tag{C.3a}$$

$$= \frac{\exp\left(\frac{\overline{U}_{i\hat{j}t}}{1-\lambda}\right)}{\exp\left\{\log\left[\sum_{\hat{j}\in\hat{g}}\exp\left(\frac{\overline{U}_{i\hat{j}t}}{(1-\lambda)}\right)\right]\right\}} \times \frac{\exp\left\{\log\left[\sum_{\hat{j}\in\hat{g}}\exp\left(\frac{\overline{U}_{i\hat{j}t}}{(1-\lambda)}\right)\right]^{1-\lambda}\right\}}{\exp\left\{log\left[\sum_{\hat{g}=0}^{2}\exp\left(\log\left\{\sum_{\tilde{j}\in\tilde{g}}\exp\left[\frac{\overline{U}_{i\tilde{j}t}}{(1-\lambda)}\right]\right\}^{1-\lambda}\right)\right]\right\}},$$

$$= \frac{\exp\left(\frac{\overline{U}_{i\hat{j}t}}{1-\lambda}\right)}{\exp\left\{\log\left[\sum_{\hat{j}\in\hat{g}}\exp\left(\frac{\overline{U}_{i\hat{j}t}}{(1-\lambda)}\right)\right]\right\}} \times \frac{\exp\left\{(1-\lambda)\log\left[\sum_{\hat{j}\in\hat{g}}\exp\left(\frac{\overline{U}_{i\hat{j}t}}{(1-\lambda)}\right)\right]\right\}}{\exp\left\{log\left[\sum_{\hat{g}=0}^{2}\exp\left((1-\lambda)\log\left\{\sum_{\tilde{j}\in\tilde{g}}\exp\left[\frac{\overline{U}_{i\tilde{j}t}}{(1-\lambda)}\right]\right\}\right)\right]\right\}},$$

$$= \frac{\exp\left(\frac{\overline{U}_{i\hat{j}t}}{1-\lambda}\right)}{\exp\left(\frac{I_{i\hat{g}R_i}}{1-\lambda}\right)} \times \frac{\exp(I_{i\hat{g}R_i})}{\exp(I_{\hat{g}R_i})}, \tag{C.3b}$$

$$= \quad \mathbb{P}_{i\hat{g}t|R_i} \quad \times \quad \mathbb{P}_{i\hat{j}t|\hat{g}R_i}, \tag{C.3c}$$

$$i = 1, \ldots, I_t, \qquad \hat{j} \in (g \cap \hat{J}_{tR_i}), \qquad g \in \{0,1,2\}, \qquad t = 1, \ldots, T,$$

where:

$$I_{i\hat{g}R_i} \equiv (1-\lambda)\mathbb{E}\left[\max_{j\in(\hat{g}\cap\hat{J}_{tR_i})} U_{ijt|R_i}\right],$$

$$= (1-\lambda)\log \sum_{j\in(\hat{g}\cap\hat{J}_{tR_i})} e^{\overline{U}_{ijt}/(1-\lambda)}, \tag{C.4a}$$

$$I_{\hat{g}R_i} \equiv \log\left(1+\sum_{g=1}^{2} e^{I_{i\hat{g}R_i}}\right), \tag{C.4b}$$

and where the first equality in (C.3a) follows from the nested logit choice probability (*e.g.* McFadden 1978, equation (18)); the equality in (C.3b) follows from replacing the definitions of the *inclusive values*, $I_{i\hat{g}R_i}$ and $I_{\hat{g}R_i}$, given by equations in (C.4a) and (C.4b), respectively;[74] and where $\mathbb{P}_{i\hat{g}t|R_i}$ is the marginal conditional probability of choosing a product in group $\hat{g}$ given that the product is in the consideration set $\hat{J}_{tR_i}$, and $\mathbb{P}_{i\hat{j}t|\hat{g}R_i}$ is the conditional probability of choosing product $\hat{j}$ given that the product is in group $\hat{g}$ and in the consideration set $\hat{J}_{tR_i}$.

C.4 Welfare Measures

The expected consumer surplus, in dollars, for consumer i is given by:

$$\begin{aligned}
\mathbb{E}(CS_i) &= \frac{1}{\alpha_i}\ \mathbb{E}\left\{\max_{R_i'\in\Lambda}\left[\max_{j\in\hat{J}_i}\left(\overline{U}_{ijt}+\hat{\varepsilon}_{ijt}\right) - SC_{itR_i} + \tilde{\varepsilon}_{itR_i}\right]\right\}, \\
&= \frac{1}{\alpha_i}\int_{\tilde{\varepsilon}_{ijt}}\left\{\max_{R_i'\in\Lambda}\left[\int_{\hat{\varepsilon}_{ijt}}\left(\max_{j\in\hat{J}_i}\left\{\overline{U}_{ijt}+\hat{\varepsilon}_{ijt}\right\}\ \mathrm{d}F_{\hat{\varepsilon}_{itR_i}}(\hat{\varepsilon}_{itR_i})\right) - SC_{itR_i} + \tilde{\varepsilon}_{itR_i}\right]\right\}\mathrm{d}F_{\tilde{\varepsilon}_{itR_i}}(\tilde{\varepsilon}_{itR_i}), \\
&= \frac{1}{\alpha_i}\int_{\tilde{\varepsilon}_{ijt}}\left\{\max_{R_i'\in\Lambda}\left[\log\left(1+\sum_{\hat{g}=1}^{2} e^{I_{\hat{g}R_i}}\right) - SC_{itR_i} + \tilde{\varepsilon}_{itR_i}\right]\right\}\ \mathrm{d}F_{\tilde{\varepsilon}_{itR_i}}(\tilde{\varepsilon}_{itR_i}) + C_1, \\
&= \frac{1}{\alpha_i}\log\left\{\sum_{R_i'\in\Lambda}\exp\left[\log\left(1+\sum_{\hat{g}=1}^{2} e^{I_{\hat{g}R_i}}\right) - SC_{itR_i}\right]\right\} + C_2,
\end{aligned} \tag{C.5}$$

where the expectation in the first line is taken over both random shocks, $\hat{\varepsilon}_{ijt}$ and $\tilde{\varepsilon}_{ijR_i}$; the third equality follows by computing the maximum expected utility over the shocks $\hat{\varepsilon}_{ijt}$ using the expression in equation (C.2) with the inclusive value, $I_{\hat{g}R_i}$, defined by the equation in (C.1); the fourth equality follows by computing the maximum expected utility over the shocks

[74] As before, $\mathbb{E}(\cdot)$ is the expectation operator taken over the *i.i.d.* shocks within group $\hat{g}$; the equation in (C.4a) follows from the maximum expected value of the logit model (*i.e.* within group $\hat{g}$) (see, *e.g.*, footnote 73); and the equation in (C.4b) follows because the inclusive value of the outside product is equal to zero.

$\tilde{\varepsilon}_{ijt}$ (*e.g.* McFadden 1978); and C_1 and C_2 are two constants.

D Details about the Supply

Step 2: Estimation of the bargaining weights.

In this section we compute the equations that we use for the estimation in step 2 for the supply side. We derive the first order necessary conditions from the bargaining problem and transform them into an expression that we use to recover the bargaining weights.

Next equation repeats the objective problem to facilitate the reading:

$$\begin{aligned}\mathcal{N}_{rmj} &\equiv \left[\sum_{k\in\Omega_r}\left(p_k^* - \omega_k - \rho_k\right) M s_k(\mathcal{P}(\omega)) - \Pi_{r,-j}^{\text{Dist}}\right]^{\nu_{rmj}} \left[\sum_{k\in\Omega_w}\left(\omega_k - \mu_k\right) M s_k(\mathcal{P}(\omega)) - \Pi_{m,-j}^{\text{Dist}}\right]^{1-\nu_{rmj}}, \\ &\equiv \left[\Pi_r^{\text{Dist}}(\omega^*) - \Pi_{r,-j}^{\text{Dist}}\right]^{\nu_{rmj}} \left[\Pi_m(\omega^*) - \Pi_{m,-j}\right]^{1-\nu_{rmj}}. \end{aligned} \tag{D.1}$$

The first order necessary conditions for ω_j from the equation in (D.1) are:

$$\begin{aligned}\nu_{rmj}\left[\Pi_r^{\text{Dist}}(\omega^*) - \Pi_{r,-j}^{\text{Dist}}\right]^{\nu_{rmj}-1} \left[\Pi_m(\omega^*) - \Pi_{m,-j}\right]^{1-\nu_{rmj}} \frac{\partial \Pi_{rj}^{\text{Dist}}(\omega^*)}{\partial \omega_j} + \\ (1-\nu_{rmj})\left[\Pi_r^{\text{Dist}}(\omega^*) - \Pi_{r,-j}^{\text{Dist}}\right]^{\nu_{rmj}} \left[\Pi_m(\omega^*) - \Pi_{m,-j}\right]^{-\nu_{rmj}} \frac{\partial \Pi_{mj}(\omega^*)}{\partial \omega_j} = 0. \end{aligned} \tag{D.2}$$

From the envelope theorem $\frac{\partial \Pi_m(\omega^*)}{\partial \omega_j} = -\frac{\partial \Pi_m^{\text{Dist}}(\omega^*)}{\partial \omega_j} = M s_j(\mathcal{P}(\omega))$, which means that the equation in (D.2) simplifies to:

$$\begin{aligned}\nu_{rmj}\left[\Pi_r^{\text{Dist}}(\omega^*) - \Pi_{r,-j}^{\text{Dist}}\right]^{\nu_{rmj}-1} \left[\Pi_m(\omega^*) - \Pi_{m,-j}\right]^{1-\nu_{rmj}} - \\ (1-\nu_{rmj})\left[\Pi_r^{\text{Dist}}(\omega^*) - \Pi_{r,-j}^{\text{Dist}}\right]^{\nu_{rmj}} \left[\Pi_m(\omega^*) - \Pi_{m,-j}\right]^{-\nu_{rmj}} = 0. \end{aligned} \tag{D.3}$$

Simplifying the equation in (D.3):

$$
\begin{aligned}
\nu_{rmj}\left[\Pi_m(\omega^*)-\Pi_{m,-j}\right]^{1-\nu_{rmj}}-\left(1-\nu_{rmj}\right)\left[\Pi_r^{\text{Dist}}(\omega^*)-\Pi_{r,-j}^{\text{Dist}}\right]\left[\Pi_m(\omega^*)-\Pi_{m,-j}\right]^{-\nu_{rmj}}&=0,\\
\nu_{rmj}\left[\Pi_m(\omega^*)-\Pi_{m,-j}\right]-\left(1-\nu_{rmj}\right)\left[\Pi_r^{\text{Dist}}(\omega^*)-\Pi_{r,-j}^{\text{Dist}}\right]&=0,\\
\left[\Pi_m(\omega^*)-\Pi_{m,-j}\right]-\frac{\left(1-\nu_{rmj}\right)}{\nu_{rmj}}\left[\Pi_r^{\text{Dist}}(\omega^*)-\Pi_{r,-j}^{\text{Dist}}\right]&=0,\\
\left[\Pi_m(\omega^*)-\Pi_{m,-j}\right]-\delta_{rmj}\left[\Pi_r^{\text{Dist}}(\omega^*)-\Pi_{r,-j}^{\text{Dist}}\right]&=0.
\end{aligned}
\tag{D.4}
$$

Now look at the components in the equation in (D.4), namely, the difference between the profits with agreement and disagreement:

$$
\begin{aligned}
\Pi_m(\omega^*)-\Pi_{m,-j}&=\sum_{k\in\Omega_m}\left(p_k^*-\omega_k-\rho_k\right)Ms_k(\mathcal{P}(\omega))-\sum_{k\in\Omega_m\setminus j}\left(p_k^*-\omega_k-\rho_k\right)Ms_k^{-j}(\mathcal{P}(\omega)),\\
&=\left(p_j^*-\omega_j-\rho_j\right)Ms_j(\mathcal{P}(\omega))-\sum_{k\in\Omega_m\setminus j}\left(p_k^*-\omega_k-\rho_k\right)M\underbrace{\left[s_k^{-j}(\mathcal{P}(\omega))-s_k(\mathcal{P}(\omega))\right]}_{\Delta s_k^{-j}},\\
&=\left(p_j^*-\omega_j-\rho_j\right)Ms_j(\mathcal{P}(\omega))-\sum_{k\in\Omega_m\setminus j}\left(p_k^*-\omega_k-\rho_k\right)M\Delta s_k^{-j}.
\end{aligned}
\tag{D.5}
$$

where $\Delta s_k^{-j} \equiv s_k^{-j}(\mathcal{P}(\omega)) - s_k(\mathcal{P}(\omega))$ denotes the difference between the market share of product k if product j is offered and if it is not. Note that this expression corresponds to equation (9) in Draganska, Klapper, and Villas-Boas (2010, p. 62) that in our case does not has closed-form solution. It corresponds to the proportion of the market share of product j that is allocated to the other products carried by the retailer or manufacturer. Similarly, for the difference in retail profits, an expression analogue to the one in (D.5) is obtained.

Next replace the expression in (D.5) and its analogue for the the difference in retail profits into (D.4), and divide by M to obtain:

$$
\begin{aligned}
&\left[\left(p_j^*-\omega_j-\rho_j\right)s_j-\sum_{k\in\Omega_m\setminus j}\left(p_k^*-\omega_k-\rho_k\right)\Delta s_k^{-j}\right]\\
&\qquad-\delta_{rmj}\left[\left(p_j^*-\omega_j-\rho_j\right)s_j-\sum_{k\in\Omega_r\setminus j}\left(p_k^*-\omega_k-\rho_k\right)\Delta s_k^{-j}\right]=0.
\end{aligned}
\tag{D.6}
$$

Denote that matrix of shares and changes in shares by:

$$\overline{s} \equiv \begin{bmatrix} s_1 & -\Delta s_2^{-1} & \cdots & -\Delta s_{\mathcal{J}}^{-1} \\ -\Delta s_1^{-2} & s_2 & \cdots & -\Delta s_{\mathcal{J}}^{-2} \\ \vdots & \vdots & \ddots & \vdots \\ -\Delta s_1^{-\mathcal{J}} & -\Delta s_2^{-\mathcal{J}} & \cdots & s_J \end{bmatrix}. \tag{D.7}$$

Finally rewrite the equation in (D.6) in matrix form using (D.7):

$$\begin{aligned} \left(\Lambda^{\mathcal{M}} \odot \overline{s}\right)(\omega^* - \mu) - \delta \left(\Lambda^{\mathcal{R}} \odot \overline{s}\right)(\mathbf{p}^* - \omega^* - \rho) = 0, \\ \left[\left(\Lambda^{\mathcal{M}} \odot \overline{s}\right)(\omega^* - \mu)\right] \left[\left(\Lambda^{\mathcal{R}} \odot \overline{s}\right)(\mathbf{p}^* - \omega^* - \rho)\right]^{-1} = \delta. \end{aligned} \tag{D.8}$$

We use the expression in (D.8) for the estimation in step 2 of the supply side.

E Additional Tables and Figures

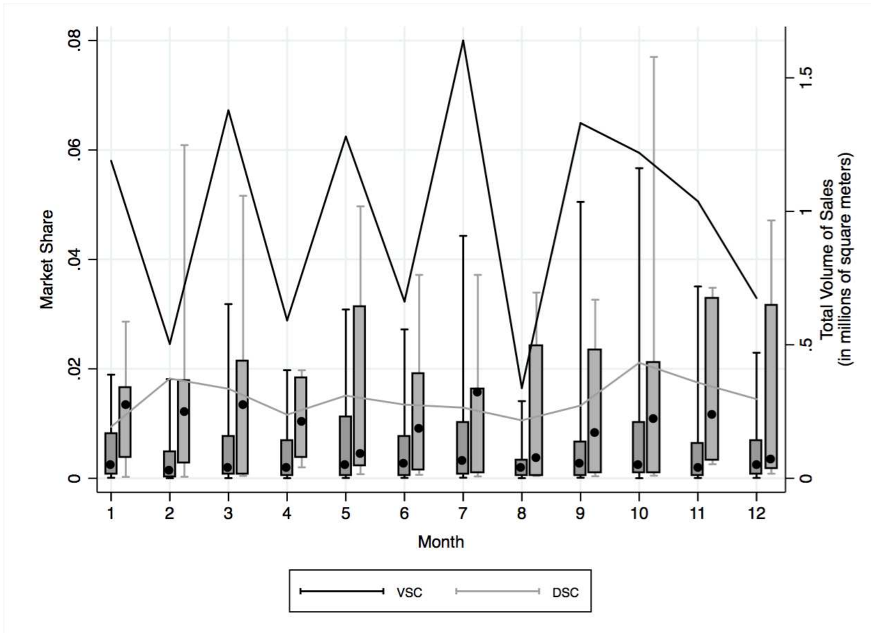
Table A1: Wholesale and Retail Prices in the VSC: By Manufacturer and by VSC Retailer, Display Format: 2 m^2 panel.

Manu-facturer	VSC Retailer	Wholesale Price Median	Mean	SD	Min	Max	VSC Price (Retail Price) Median	Mean	SD	Min	Max
m_1	r_4^v	8.01	7.75	1.02	5.68	9.41	8.40	8.30	1.11	6.11	10.27
	r_5^v	9.84	9.29	2.77	3.36	13.97	11.13	11.06	2.35	7.42	15.40
	r_6^v	2.94	3.72	1.89	1.42	7.25	3.20	4.01	2.05	1.50	8.19
	r_7^v	5.92	6.00	1.31	3.51	8.29	6.20	6.27	1.35	3.64	8.61
	r_8^v	10.50	11.58	2.92	8.10	16.27	11.54	12.94	3.46	8.94	19.42
	r_9^v	12.68	18.29	13.08	5.55	51.64	19.11	21.36	13.89	7.38	56.18
m_2	r_4^v	11.95	11.60	2.71	6.92	16.23	12.28	11.97	2.77	7.31	16.70
	r_5^v	11.62	12.70	5.58	5.76	29.02	11.94	13.42	5.51	7.02	29.92
	r_6^v	5.10	5.59	2.76	2.33	11.34	5.28	5.77	2.82	2.41	11.62
	r_7^v	8.64	9.64	4.17	4.79	20.64	9.01	10.02	4.27	4.94	21.19
	r_8^v	9.54	17.05	17.52	3.79	55.53	10.59	18.34	18.77	3.92	60.76
	r_9^v	14.21	19.99	16.30	8.19	64.80	16.96	23.92	20.32	10.97	83.22
m_3	r_4^v	7.07	7.50	1.14	6.47	9.93	7.28	7.71	1.15	6.79	10.28
	r_5^v	6.07	7.99	5.44	4.20	22.56	7.32	8.92	5.65	4.88	24.28
	r_6^v	3.42	3.78	2.30	1.04	8.79	3.58	4.02	2.41	1.09	9.34
	r_7^v	6.29	14.64	22.20	2.96	79.45	6.49	15.14	22.74	3.17	81.52
	r_8^v	4.51	5.28	2.24	3.57	11.16	4.94	5.72	2.29	4.05	11.63
	r_9^v	6.97	12.86	15.85	5.24	51.92	9.44	15.63	19.78	5.49	64.36
m_4	r_4^v	11.48	21.56	20.72	5.29	61.70	11.60	22.35	21.71	5.34	61.65
	r_5^v	40.44	34.72	14.47	18.48	56.84	43.47	37.48	16.88	18.92	64.88
	r_6^v	3.42	3.17	1.68	0.81	6.04	3.52	3.27	1.72	0.86	6.20
	r_7^v	11.42	12.41	3.32	7.43	18.40	11.98	13.29	3.75	7.94	19.94
	r_9^v	30.91	31.17	23.52	9.94	93.38	33.10	33.42	24.67	11.91	99.39

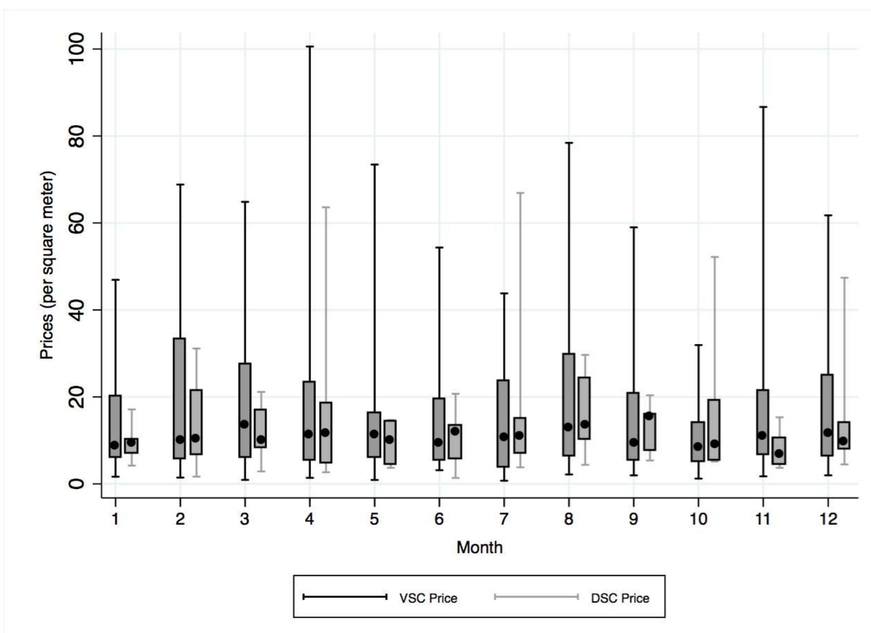
Notes: The table reports summary statistics of wholesale and VSC prices (*i.e.* retail prices) in the VSC for the display format 2 m^2 panel for each combination of manufacturer (m_1, m_2, m_3, and m_4) and VSC retailer ($r_4^v, r_5^v, \ldots, r_9^v$) across months of the year. (Note that this table corresponds to table 2, Panel B, sub-panel 2 m^2 panel, desegregated by retailer that is not manufacturer.) Similar tables for the other display formats (seniors and others) are available upon request. "DSC" stands for Direct Sales Channel. "VSC" stands for Vertical Sales Channel. See figure 1 for definitions of prices and vertical relations in the industry.

Figure A1: Market Shares, Total Volume, and Prices by Month.

Market Shares and Total Volume



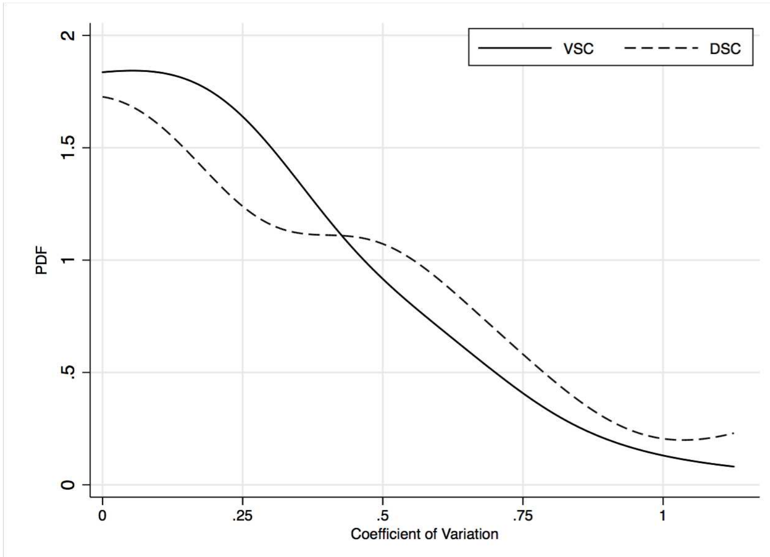
Prices



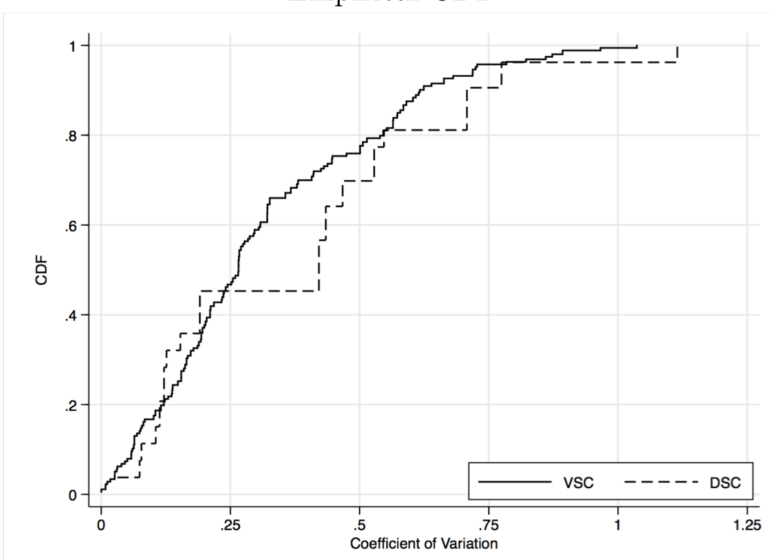
Notes: The top panel displays the distribution of market shares and total volume by month. The left vertical axis shows the distribution of market shares each month, distinguishing the sales to consumers in the VSC and DSC. The right vertical axis shows the total sales' volume from each month (horizontal series) distinguishing the sales to consumers in the VSC and DSC. The bottom panel displays the distribution of prices (per square meter) each month, distinguishing the sales to consumers in the VSC and DSC. Each vertical box displays the 95th percentile (upper whisker), 75th percentile (upper hinge), median (black circle marker), 25th percentile (lower hinge), and 5th percentile (lower whisker). The maximum market share by month are as follows (the first number refers to the sales on the VSC and the second number refers to the sales in the DSC): January (0.188, 0.029); February: (0.028, 0.061); March:(0.218, 0.052); April: (0.024, 0.020); May: (0.142, 0.050); June: (0.032, 0.037); July: (0.164, 0.037); August: (0.016, 0.034); September: (0.139, 0.033); October: (0.066, 0.077); November: (0.139, 0.035); December: (0.038, 0.047). "DSC" stands for Direct Sales Channel. "VSC" stands for Vertical Sales Channel. See figure 1 for definitions of prices and vertical relations in the industry.

Figure A2: Distribution of Coefficient of Variation.

Density Estimate



Empirical CDF



Notes: The figure displays the kernel density estimate (top panel) and empirical cumulative distribution (bottom panel) of coefficient of variation of prices (CV) for sales made to consumers through retailers and manufacturers, conditional on quantity discounts. To perform the estimation we proceed in three steps. First, we define the unit of analysis as a tuple (Display Format, Manufacturer, Retailer, Volume Percentile), where "Display format" are the display formats as defined in subsection 2.1, "Manufacturer" are the manufacturers of the product ($M_1, \ldots, M_4$), "Retailer" are the VSC retailers ($r_4^v, \ldots, r_9^v$) and DSC retailers (r_1^d, r_2^d, r_3^d), and "Volume Percentile" are the percentiles in the volume variable (to account for quantity discounts). Second, for each unit of analysis (*i.e.* tuple as defined above) we compute the CV (*i.e.* the variation of prices is within tuple). Third, we estimate the kernel density and empirical cumulative distribution as follows. Let cv denote realized CV in each tuple $j \in \{1, \ldots, J\}$. We estimate the probability density function for sales made to consumers through retailers and manufacturers, $f(cv)$, as: $\hat{f}_K(cv; h) = \frac{1}{J\,h} \sum_{j=1}^{J} K\left(\frac{cv - cv(j)}{h}\right)$, where $K(z)$ is a standard univariate gaussian kernel function, h is the bandwidth that we choose by cross validation, and $cv(j)$, $j = 1, \ldots, J$ are the CV in each tuple. Given that the price distribution has its domain bounded we use a renormalization method to deal with the boundaries when estimating the probability density function of CV. We estimate the empirical cumulative distribution of CV, $F(cv)$, as: $\hat{F}_J(cv) = \frac{1}{J} \sum_{j=1}^{J} \mathbf{1}\{cv(j) \leq cv\}$, where $\mathbf{1}\{A\}$ is the indicator function of the event A. "DSC" stands for Direct Sales Channel. "VSC" stands for Vertical Sales Channel. See figure 1 for definitions of prices and vertical relations in the industry.