Consumer Search and Prices in the Automobile Market*

José Luis Moraga-González[†]

Zsolt Sándor[‡]

Matthijs R. Wildenbeest[§]

December 2018

Abstract

In many markets consumers have imperfect information about the utility they derive from the products that are on offer and need to visit stores to find the product that is the most preferred. This paper develops a discrete choice model of demand with optimal sequential consumer search. Consumers first choose a product to search; then, once they learn the utility they get from the searched product, they choose whether to buy it or to keep searching. The set of products searched is endogenous and consumer specific. Therefore, substitution patterns are not only driven by variation in product characteristics but also by variation in search costs. We apply the model to the automobile industry. Our search cost estimate is highly significant and indicates that consumers conduct a limited amount of search. Estimates of own-price elasticities are lower and markups are higher than if we assume consumers have full information. Moreover, cross-price elasticity estimates for the search model indicate that consumers are more likely to switch to a brand that is located nearby than in the full information case.

Keywords: consumer search, differentiated products, demand estimation, automobiles
JEL Classification: C14, D83, L13

---

*We thank Mark Armstrong, Michael Choi, Jean-Pierre Dubé, Amit Gandhi, Elisabeth Honka, Jean-François Houde, Sergei Koulayev, Emanuel Marcu, Ariel Pakes, Steven Puller, Stephan Seiler, Marc Santugini, Philipp Schmidt-Dengler, Michelle Sovinsky, and Frank Verboven for their useful comments and suggestions. This paper has also benefited from presentations at the Düsseldorf Institute for Competition Economics, Illinois State University, Katholieke Universiteit Leuven, The Ohio State University, Penn State University, Texas A&M University, Tilburg University, Tinbergen Institute, University of Chicago, University of Illinois at Urbana-Champaign, University of Michigan, University of North Carolina, University of Zurich, 2009 Workshop on Search and Switching Costs at the University of Groningen, 2010 IIOC meeting in Vancouver, 2010 Marketing Science Conference in Cologne, 2011 European Meeting of the Econometric Society in Oslo, 2012 IOS ASSA Meeting in Chicago, 9th Invitational Choice Symposium in Noordwijk, 2014 Barcelona GSE Summer Forum, 2014 Workshop on Search and Switching Costs at Indiana University, 2014 MaCCI Workshop on Consumer Search in Bad Homburg, and the 2018 Tinbergen Institute Workshop on Empirical IO. Financial support from Marie Curie Excellence Grant MEXT-CT-2006-042471 and grant PN-II-ID-PCE-2012-4-0066 of the Romanian Ministry of National Education, CNCS-UEFISCDI is gratefully acknowledged.

[†]Vrije Universiteit Amsterdam, E-mail: `j.l.moragagonzalez@vu.nl`. Moraga-González is also affiliated with the University of Groningen, the Tinbergen Institute, the CEPR, and the Public-Private Sector Research Center at IESE (Barcelona).

[‡]Sapientia Hungarian University of Transylvania, E-mail: `sandorzsolt@cs.sapientia.ro`.

[§]Indiana University, Kelley School of Business, E-mail: `mwildenb@indiana.edu`.

1 Introduction

In many markets, such as those for automobiles, electronics, computers, and clothing, consumers typically have to visit stores to find out which product they like most. Though basic information about products sold in these markets is usually easy to obtain either from television, the Internet, newspapers, specialized magazines, or just from neighbors, family, and friends, consumers search because some relevant product characteristics are difficult to quantify, print, or advertise. In practice, since visiting stores involves significant search costs, most consumers engage in a limited amount of search.[1]

Earlier work on the estimation of demand models (Berry, Levinsohn, and Pakes, 1995, 2004; Nevo, 2001; Petrin, 2002) has proceeded by assuming that consumers have perfect information about all the products available in the market. In the market settings referred to above, the full information assumption is, arguably, unrealistic. In a study of the US computer industry, Sovinsky Goeree (2008) shows that departing from the perfect information assumption is important for obtaining realistic estimates of demand and supply parameters. In her model, firms distribute advertisements about the existence and characteristics of the computers they sell. Advertisements are perfectly informative and consumers differ in the likelihood with which they are exposed to them. As a result consumers end up having heterogeneous and limited information about the existing alternatives in the market. Yet, in the setting of Sovinsky Goeree (2008) consumers do not need to incur any search costs to evaluate the utility they derive from the alternatives they happen to be informed of via the advertisements.

This paper adds to the literature on the structural estimation of demand models by presenting a discrete choice model of demand with optimal sequential consumer search. To the best of our knowledge our paper is the first to do this in a Berry, Levinsohn, and Pakes (1995) (BLP hereafter) framework. The distinctive feature of the BLP framework is that a product's utility depends on a structural error term, which is known as an unobserved product characteristic in this literature, and is crucial for modeling price endogeneity. The key difference between our demand model and that in BLP is that in our model consumers do not know all the relevant information about the products available and have to search in order to evaluate them. Search is costly and consumers search sequentially through the available options. The costs of searching vary across individuals and firms so consumers choose to visit distinct sellers even if they have similar preferences. In our model consumer product information is thus endogenous and consumer specific. As a result, substitution patterns across products are not only driven by product differentiation but also by the variation in consumer information sets generated by costly search. Similar to the effects of advertising in Sovinsky Goeree (2008), search frictions thus generate heterogeneous and limited consumer information.

We apply the model to the automobile market. The automobile market is precisely a market in which advertisements, reports in specialized magazines, television programs, and the Internet convey much but not

[1]Several recent empirical papers have found that consumers search relatively little. For instance, Honka (2014) reports that consumers obtain an average of 2.96 quotes when shopping for car insurance. De los Santos, Hortaçsu, and Wildenbeest (2012) find that over 75 percent of consumers visited only one online bookstore before buying a book online, whereas De los Santos, Hortaçsu, and Wildenbeest (2017) find that the mean number of online retailers searched is less than 3 for MP3 players. Some other examples of markets in which search frictions are found to be non-trivial are S&P 500 index funds (Hortaçsu and Syverson, 2004), automobiles (Moorthy, Ratchford, and Talukdar, 1997; Scott Morton, Silva-Risso, and Zettelmeyer, 2011), and the retail market for illicit drugs (Galenianos and Gavazza, 2017).

all the relevant information about the models available. As a result, a great deal of new car buyers visit dealerships to view, inspect, and test-drive cars. Given prior information about the products available at the various sellers (such as size, horsepower, fuel efficiency, price, and design of a car), as well as the costs of search (which we relate to dealership locations and certain consumer demographics), each consumer chooses the order in which she will visit the sellers and, after each visit, whether to stop searching and buy one of the alternatives inspected so far or else to continue searching at the next seller. Following Weitzman (1979), the solution of this search problem consists of ranking the sellers in terms of reservation utilities, visiting them in descending reservation utility order, and stopping search when the highest observed utility is above the reservation utility of the next option to be searched.

The complexity of the real-world setting to which we apply our model—specifically, many alternatives available in the market that differ in ex-ante observable characteristics and search costs—makes the computation of the consumers' buying probabilities a challenge. When there is a large number of alternatives available in the market, there are many ways through which a consumer may end up buying a particular product. For example, with just three alternatives, there are eleven distinct search paths a consumer may follow before deciding to buy a given product. As the number of alternatives increases, the number of search paths grows factorially, which makes Weitzman's solution difficult to implement directly in settings in which many alternatives are available. To address this challenge, we adapt recent findings from the theoretical search literature by Armstrong (2017) and Choi, Dai, and Kim (2018) that make it possible to compute the buying probability of a given alternative without having to go explicitly through the myriad of possible ways in which a consumer may end up considering the alternative in question.

We develop our search model in Section 2. Following Armstrong (2017) and Choi, Dai, and Kim (2018), we express a consumer's search and purchase decision as a discrete choice problem in which a consumer chooses the alternative that offers the highest minimum of the reservation value and realized utility among all available alternatives. We show how to use these probabilities to estimate the model, but, because these expressions are generally not closed form, estimation of the model is relatively slow. To speed up the process, we develop a method that significantly reduces the computational complexity of the model. We solve the model backwards: starting from a distribution for the minimum of the reservation value and realized utility that leads to closed-form expressions for the buying probabilities, we derive a parametric specification for the search cost distribution that rationalizes this distributional assumption. This alternative procedure is extremely practical and turns out to deliver estimates of demand, elasticities of demand, and markups that are very similar to those obtained for alternative search costs distributions in the more general procedure.

We discuss the estimation procedure as well as identification in Section 3. For the case in which search costs only depend on variables that are excluded from the utility function, search costs can be identified using aggregate data only, and the model can be estimated using a similar approach as in BLP. In the more general case in which similar covariates enter both the utility and search cost function, we show in the identification discussion in Section 3 that these common shifters can be separately identified by using search data in addition. Intuitively, the combined effect—as opposed to the separate effects—of utility and search cost variables on purchases can be identified based on purchase decisions only. According to economic theory, the reservation values that guide search decisions respond differently to changes in utility shifters than to changes in search cost shifters, and, as a result, variation in observed search decisions allows us to separate the combined effect into a utility part and a search part. For example, a high market share for the outside option could be driven by a high utility for the outside option relative to the inside goods, or relatively high search cost. Without observing search decisions we cannot distinguish between the two. However, if according to search data consumers are searching a lot but end up not buying, it must be that search costs are low and the utility for the outside option is high.

We estimate our model's main specification using aggregate data on characteristics and market shares, as well as individual-level survey data that contains information on purchases and related search behavior. Following Berry, Levinsohn, and Pakes's (2004) two-step procedure, we use the aggregate data to obtain mean utilities, which capture the linear part of utility. These mean utilities and the individual-level survey data are used to estimate the nonlinear parameters of the model, which include random coefficients as well as the parameters of the search cost specification. The moments that we use to estimate the nonlinear parameters are defined as the difference between a variety of predicted and observed (from the survey) purchase and search probabilities. As we discuss in more detail in the identification section, the exact moments we use for estimation are picked such that there is a clear relation between them and the specific nonlinear variable. In a second step we use the estimated mean utilities to estimate the linear utility parameters, using an instrumental variables approach to deal with price endogeneity. An advantage of using the two-step procedure in order to obtain estimates of search costs and other nonlinear parameters is that we do not need to impose restrictions on the joint distribution of unobserved heterogeneity and observed utility and search cost shifters, making estimation of our model less susceptible to misspecification bias.

In Section 4 we apply our model to the Dutch market for new cars and present our estimation results. Survey data reveal two important facts. First, consumers visit a limited number of car dealers before buying a car—on average two for new car purchases—and the number of visits varies substantially across consumers. Second, a great deal of the dealer visits involve test-driving cars.[2] We interpret these two facts as being consistent with our search model. We also provide some reduced-form evidence that search behavior is related to demographics such as income, family size, age, and distances to dealerships.

Section 4 also discusses the estimation results for our main model. Our search cost estimates are highly significant and suggest consumers' search costs are affected by distances to dealership and other demographics such as income, age, and household composition. Moreover, taking into account search costs leads to less elastic demand estimates and higher estimates of price-cost margins compared to the standard BLP setting. According to our estimates, substitution patterns are not only driven by car characteristics but also by search costs, and our cross-price elasticity estimates indicate that consumers are relatively more likely to switch to a brand that is located nearby than is the case in the full information model. We conclude that accounting for costly search and its effects on generating heterogeneity in consumer choice sets is important

[2]According to survey data discussed in Section 4, respondents that were looking to buy a new car made a test drive in 45 percent of dealer visits, and 75 percent made at least one test drive at one of the visited dealerships.

for explaining variability in purchase patterns.

In Section 4 we also use the estimates to perform several counterfactual analyses. Specifically, taking our estimates as a starting point, we perform simulations that allow us to study the effects of lowering the costs of visiting dealerships on prices and profits. By offering home-delivered test drives, these traveling costs can be brought down to near zero.[3] Our simulation results indicate that average prices decrease by one percent when distance-related search costs go down to zero, with price decreases exceeding three percent for some cars. We find that the further away a dealer is (on average) from consumers, the less its price will be affected by changes in distance-related search costs. This could be explained by the fact that buyers of brands that have very few dealers tend to have relatively strong preferences for those brands.

We also simulate the competitive effects of changes in the way manufacturers use their dealer networks. When a manufacturer incorporates a new brand to the group of brands it sells, it has the choice whether to reorganize business and start retailing the new brand together with the existing ones. This is typically a long run decision because such reorganizations involve the re-design and refurbishing of showrooms, which may involve large sunk costs. Our estimates of the gains from such business reorganizations provide lower bounds for these sunk costs. We find that mergers of existing dealership networks could have nontrivial effects on sales, prices, and profits.

Related literature

Our paper builds on the theoretical and empirical literature on consumer search. At least since the seminal article of Stigler (1961) on the economics of information a great deal of theoretical and empirical work has revolved around the idea that the existence of search costs has nontrivial effects on market equilibria. Part of the effort has gone into the study of the effects of costly search in homogeneous product markets (see for instance Burdett and Judd, 1983; Reinganum, 1979; Stahl, 1989). In this literature a fundamental issue has been the existence of price dispersion in market equilibrium. Another tradition has been the study of costly search in markets with product differentiation. In a seminal contribution, Wolinsky (1986) notes that search costs generate market power even in settings with free entry of firms. More recent contributions investigate how product diversity (Anderson and Renault, 1999), product quality (Wolinsky, 2005), and product design (Bar-Isaac, Caruana, and Cuñat, 2012) are affected by costly search. As in our model, in this literature consumers search for a good product fit, and not for lower prices. Our search model is most closely related to the framework of Wolinsky (1986) but we allow for asymmetric multi-product firms, consumer heterogeneity in both preferences and search costs, and, like in Choi, Dai, and Kim (2018) and Haan, Moraga-González, and Petrikaitė (2018), for price deviations to be observable before searching.

Some recent empirical research on consumer search behavior has focused on developing techniques to estimate search costs using aggregate market data. Hong and Shum (2006) develop a structural method to retrieve information on search costs for homogeneous products using only price data. Moraga-González and Wildenbeest (2008) extend the approach of Hong and Shum (2006) to the case of oligopoly and present a

[3] In the US, Seattle-based startup *Tred* allows consumers to test-drive and buy cars without having to visit the dealership.

maximum likelihood estimator. Hortaçsu and Syverson (2004) study a search model where search frictions coexist with vertical product differentiation. Our paper contributes to this line of work by incorporating consumer search into the BLP framework.

A number of recent papers present related models of search and employ micro- or aggregate-level data on search behavior to estimate preferences as well as the costs of searching. Although several of these papers also focus on search for a good product fit (Kim, Albuquerque, and Bronnenberg, 2010; Kim, Albuquerque, and Bronnenberg, 2017), most of these papers assume that consumers are searching for prices (De los Santos, Hortaçsu, and Wildenbeest, 2012; Seiler, 2013; Honka, 2014; Koulayev, 2014; Pires, 2016; Honka, Hortaçsu, and Vitorino, 2017; Dinerstein, Einav, Levin, and Sundaresan, 2018). An important difference between these papers and ours is that they do not model unobserved product characteristics and hence they do not allow for price endogeneity.[4]

A further distinction can be made according to whether consumers are assumed to search sequentially or non-sequentially. A computational advantage of non-sequential search (De los Santos, Hortaçsu, and Wildenbeest, 2012; Honka, 2014; Moraga-González, Sándor, and Wildenbeest, 2015; Murry and Zhou, 2018) is that consumers' search decisions are determined before any search activity takes place, and therefore do not depend on realized search outcomes. This allows one to formulate the search and purchase decision as a two-stage problem in which the consumer selects products in the first stage, and then makes a purchase decision from products that appear in this choice set. A complicating factor is that without restrictions on the number of choice sets, there is a dimensionality problem, and the literature has focused on various ways to deal with this when estimating such models.[5] Note that non-sequential search models are related to "consideration set formation" models in the marketing literature, which relate consideration set heterogeneity to advertising or search frictions (see, for example, Roberts and Lattin, 1991; Mehta, Rajiv, and Srinivasan, 2003). Abaluck and Adams (2018) show that certain consideration set formation models can be identified without using information on consideration sets by using differences in cross-derivatives. However, the class of models they consider rules out consideration set probabilities that depend on the characteristics of all products, which means that their results do not directly apply to traditional search models such as our model.

In sequential search models search decisions do depend on search outcomes, which complicates the estimation of such models and typically leads to high-dimensional integrals for choice and search probabilities. While this may be manageable in applications with a small number of products or search decisions (as in Koulayev, 2014), the approach we use in this paper reduces this dimensionality problem by integrating out different search paths that lead to a purchase decision and is therefore useful for larger choice sets. Kim, Albuquerque, and Bronnenberg (2017) propose an alternative method that avoids the use of high-dimensional integrals and estimate their probit choice model by maximum likelihood using view rank and sales rank data for camcorders sold at Amazon.com. Another alternative approach is put forward in Jolivet and Turon

[4]Honka, Hortaçsu, and Vitorino (2017) use a control function approach to address *advertising* endogeneity in a three-stage structural model (consisting of awareness, consideration, and choice).

[5]For instance, Honka (2014) invokes assumptions that allow her to use the Marginal Improvement Algorithm of Chade and Smith (2006) to limit the number of choice sets, whereas an earlier version of this paper (Moraga-González, Sándor, and Wildenbeest, 2015) achieved tractability by adding a choice-set specific error term that is Type I Extreme Value (minimum) distributed to the costs of searching a group of alternatives.

(2018), who derive a set of tractable inequalities from Weitzman's optimal sequential search algorithm that can be used to set-identify demand-side parameter distributions; they estimate their model using individual purchase data for CDs sold at a French e-commerce platform. Note that both Kim, Albuquerque, and Bronnenberg (2017) and Jolivet and Turon (2018) do not explicitly deal with price endogeneity.

Our paper also fits into a broader literature that estimates demand for automobiles, which includes BLP, Goldberg (1995), Petrin (2002), and Berry, Levinsohn, and Pakes (2004).[6] As in Petrin (2002) and Berry, Levinsohn, and Pakes (2004) we use a combination of micro and aggregate data. Our estimation procedure is most similar to Berry, Levinsohn, and Pakes (2004), but instead of using moments based on second-choice data we use moments based on search behavior of individuals. Recent papers in this literature have studied car dealership locations and how this affects consumer demand and competition. For instance, Albuquerque and Bronnenberg (2012) use transaction level data as well as detailed data on the location of consumers and car dealers to estimate a model of supply and demand and find that consumers have a strong disutility for travel. In a related paper, Nurski and Verboven (2016) focus on dealer networks to study whether the exclusive contracts often used in the European car market act as barrier to entry. The most important difference between these papers and our paper is that they assume consumers have perfect information about all the alternatives in the market. This means that distance from a consumer to a car dealer is interpreted as a transportation cost, i.e., distance is treated as a product characteristic that enters directly in the utility function. In contrast, in our paper distance enters as a search cost shifter *in addition to* entering as a utility shifter and as such generates variation in the subsets of cars sampled by consumers. In Section 3 we discuss how sequential search theory allows for the separate identification of the effects of distance on utility and search costs when using data on search behavior. We compare the two approaches when estimating the model in Section 4 and show that the elasticity estimates and markups from the search cost model are quite different from those obtained from the transportation cost model.

2 Economic Model

2.1 Utility and demand

We consider a market where there are J different products (indexed $j = 1, 2, \dots, J$) sold by F different firms (or sellers) (indexed $f = 1, 2, \dots, F$). We shall denote the set of products by $\mathcal{J}$ and the set of firms by $\mathcal{F}$. The utility consumer i derives from product j is given by:

$$u_{ij} = \alpha_i p_j + \mathbf{x_j'} (\boldsymbol{\beta} + \mathbf{V_i} \sigma) + \xi_j + \varepsilon_{ij}, \tag{1}$$

where α_i is a consumer-specific price coefficient, the variable p_j denotes the price of product j and the vector $(\mathbf{x_j}, \xi_j, \varepsilon_{ij})$ describes different product attributes from which the consumer derives utility. As usual, $\mathbf{x_j}$ includes a 1 in order to allow for a constant term in the utility function. We assume that the consumer

[6]See Murry and Schneider (2016) for an overview of studies on the economics of retail markets for new and used cars.

observes the product attributes contained in $\mathbf{x_j}$ and ξ_j without searching, which in the case of cars could include characteristics such as horsepower, weight, transmission type, ABS, air-conditioning, and number of gears. Information on these characteristics are readily available from, for instance, the Internet, specialized magazines, and consumer reports. The variable ε_{ij}, which is assumed to be independently and identically Type I Extreme Value (TIEV) distributed across consumers and products, is a match parameter and measures the "fit" between consumer i and product j. We assume that ε_{ij} captures "search-like" product attributes, that is, characteristics that can only be ascertained upon close inspection and interaction with the product, like comfortability, spaciousness, engine noisiness, and gearbox smoothness in the case of cars. We assume that the econometrician observes the product attributes contained in $\mathbf{x_j}$ but cannot observe those in ξ_j and ε_{ij}. The variable ξ_j is often interpreted as (unobserved) quality, and, since quality is likely to be correlated with the price of a product, this will lead to the usual price endogeneity problem.

Consumers differ in the way they value price and product characteristics. The parameter α_i and the expression $(\boldsymbol{\beta} + \mathbf{V_i}\sigma)$ capture consumer heterogeneity in tastes for price and product attributes. Here σ is a parameter and $\mathbf{V_i}$ is a diagonal matrix that contains either demographic characteristics or standard normal draws on its main diagonal such that the first component corresponds to the first component of $\mathbf{x_j}$, the second to the second component of $\mathbf{x_j}$, and so on. Following Petrin (2002), we allow the price coefficient α_i to vary across income groups, i.e.,

$$\alpha_i = \begin{cases} \dfrac{\alpha_{(1)}}{y_i} & \text{for } y_i < \overline{y}; \\ \dfrac{\alpha_{(2)}}{y_i} & \text{for } y_i \geq \overline{y}, \end{cases} \tag{2}$$

where $\alpha_{(1)}$ and $\alpha_{(2)}$ are deterministic parameters, y_i is the yearly income of consumer i, and $\overline{y}$ is a chosen income bound.

The utility from not buying any of the products is

$$u_{i0} = \varepsilon_{i0}.$$

Therefore, we regard product $j = 0$ as the "outside" option; this includes the utility derived from not purchasing a new product. We allow for multi-product firms: firm $f \in \mathcal{F}$ supplies a subset $\mathcal{G}_f \subset \mathcal{J}$ of all products. In the car industry dealers typically sell disjoint sets of products, so we shall assume that $\mathcal{G}_f \cap \mathcal{G}_g = \varnothing$ for any $f \neq g,\ f, g \in \mathcal{F}$.

We assume consumers must search to find out the exact utility they derive from each of the products available. To be more specific, we assume that before searching a consumer i knows (i) the location of each seller and the subset of product available at each store, (ii) product characteristics p_j, $\mathbf{x_j}$ and ξ_j for each product j, (iii) the distribution F of match values ε_{ij}, and (iv) the utility of her outside option ε_{i0}. We thus regard the search of a consumer i as a process by which she discovers the exact values of the matching parameter ε_{ij} upon visiting a store that sells product j.[7]

[7] In general, one can distinguish between store search and brand search. In our model consumers search among different brands. The difference with store search is that the same brand may be sold by several stores, which is not the case in the Dutch car market.

Consumers search sequentially with costless recall, i.e., they determine after each visit to a store whether to buy any of the inspected products so far, to opt for the outside option, or to continue searching. Let c_{if} denote the search cost of consumer i for visiting firm f. We assume that search costs vary across consumers and firms.[8] Let F_{if}^c be the cumulative distribution of consumer i's cost of searching firm f, with corresponding density f_{if}^c. We allow the search cost distribution to have full support, although, as we explain in Section 2.3, only the non-negative part affects search behavior (see also footnote 13).

2.2 Optimal sequential search

The utility function in equation (1) can be rewritten as

$$u_{ij} = \delta_{ij} + \varepsilon_{ij},$$

where δ_{ij} is the mean utility consumer i derives from product j. As explained above, in this expression the consumer knows δ_{ij}, but has to search to discover ε_{ij}. The distribution of match values ε_{ij}, which is the same for all consumers and products, is given by $F(z) = \exp(-\exp(-z))$.

Since we allow for consumer-specific taste parameters, the distribution of consumer i's utility u_{ij} from a given product j differs across consumers. This leads to the usual aggregation problem we need to deal with. Since the utility shock ε_{ij} is an IID draw from a TIEV distribution, the utility distribution for product j faced by consumer i is

$$F_{ij}(z) = F(z - \delta_{ij}) = \exp\left(-\exp\left(\delta_{ij} - z\right)\right),$$

that is, the distribution of u_{ij} is Gumbel with location parameter δ_{ij} and scale parameter 1.[9]

Since search happens at the firm level, indexed by f, it is useful to define the random variable U_{if} as the highest utility consumer i gets from the products sold by firm f, i.e.,

$$U_{if} = \max_{j \in G_f} \{u_{ij}\}.$$

Denoting the distribution of U_{if} by F_{if}, we get

$$F_{if}(z) = \Pr[U_{if} \leq z] = \prod_{j \in G_f} F_{ij}(z) = \prod_{j \in G_f} F(z - \delta_{ij}) = F(z - \delta_{if}), \tag{3}$$

[8] We allow for the possibility that a consumer has zero search cost for one or more products; if a consumer has zero search cost for a specific product this means that she knows the match utility she derives from the product in question ex-ante. According to our data, many consumers visit dealers for the purpose of inspecting and test driving cars. Some consumers might know a car's "fit" ex-ante, for example because a family member, friend, or neighbor owns that specific car, or because they may have rented it while on vacations somewhere. We model these consumers as having zero search costs for the specific car in question. Having zero search costs may also explain not visiting any dealership. Since we allow search costs to be consumer-firm specific, having zero search costs for one product does not imply zero search costs for all products. Also note that by letting search costs be consumer-firm specific, the model is more flexible in rationalizing choice sets than a setting in which search costs are only consumer specific.

[9] Throughout the paper, when we refer to the Gumbel distribution, we mean the distribution with CDF $\exp(-\exp(-(z-\mu)/\beta))$, where μ is a location parameter and β is a scale parameter. When we refer to the TIEV distribution we mean the distribution with CDF $\exp(-\exp(-z))$ (sometimes referred to as the standard Gumbel distribution). The difference between the two distributions is that the TIEV distribution has location parameter normalized to zero and scale parameter normalized to 1, whereas these parameters can vary for the Gumbel distribution.

where $\delta_{if} = \log\left(\sum_{h \in G_f} \exp(\delta_{ih})\right)$.

Having determined the distribution of the maximum utility a consumer i can get at a firm f, we are now ready to describe consumer i's optimal search strategy. Following Weitzman (1979) we first define the expected gains to consumer i from searching for a product at firm f when the best utility the consumer has found so far is r:

$$H_{if}(r) \equiv \int_r^\infty (z - r) dF_{if}(z). \tag{4}$$

If consumer i's expected gains are higher than the cost c_{if} she has to incur to search the products of firm f, then she should pay a visit to firm f. Correspondingly, we define the so-called *reservation value* r_{if} as the solution to equation

$$H_{if}(r) - c_{if} = 0 \tag{5}$$

in r. Notice that H_{if} is decreasing and strictly convex so equation (5) has in general a unique solution. Therefore

$$r_{if} = H_{if}^{-1}(c_{if}).$$

Note that r_{if} is a scalar, and that for each consumer i there is one such scalar for every firm f.

Weitzman (1979) demonstrates that the optimal search strategy for a consumer i consists of visiting sellers in descending order of reservation values r_{if} and stopping search as soon as the best option encountered so far (which includes the outside option) gives a higher utility than the reservation value of the next option to be searched. The following result decomposes the reservation value into a utility component and a search cost component:

Lemma 1 *Under the assumption that ε_{ij} is IID TIEV-distributed, consumer i's reservation value for firm f can be written as*

$$r_{if} = \delta_{if} + H_0^{-1}(c_{if}),$$

where

$$H_0(r) \equiv \int_r^\infty (z - r) dF(z).$$

The proof is in Appendix A. This lemma shows that there are two sources of variation in the consumer reservation values: they vary because utility distributions differ across sellers and because the costs of searching distinct sellers also differ. Note that because ε_{ij} is TIEV distributed, the result in Lemma 1 implies

$$H_0(r) = \gamma - r + \int_{\exp(-r)}^\infty \frac{\exp(-t)}{t} dt,$$

where γ is the Euler constant and the integral is the exponential integral.

2.3 Buying probabilities and market shares

Because we allow for both consumer and firm heterogeneity, Weitzman's (1979) solution is extremely hard to implement in our setting. For example, with just three options, there are eleven different search paths a consumer can follow before purchasing from a specific seller. As the number of sellers grows, the number of search paths increases factorially. To solve this problem, we next utilize a recent finding by Armstrong (2017) and Choi, Dai, and Kim (2018) which consists of a methodology for the computation of the purchase decisions without having to take into account the myriad of search paths consumers may possibly follow.[10]

For every seller f, let us then define the random variable

$$w_{if} = \min\{r_{if}, U_{if}\} \equiv \min\left\{r_{if}, \max_{j\in G_f}\{u_{ij}\}\right\}. \tag{6}$$

Armstrong (2017) and Choi, Dai, and Kim (2018) show that the solution to the sequential search problem (searching across firms in descending order of reservation values and stopping and buying the best of the observed products when its realized utility is higher than the next highest reservation value) is equivalent to picking the firm with the highest w_{if} from all the firms and choosing the product with the highest utility from that firm. Accordingly, the probability that buyer i buys product j is

$$s_{ij} = P_{ij|f}P_{if}, \tag{7}$$

where

$$P_{ij|f} = \Pr\left(u_{ij} \geq \max_{h\in G_f} u_{ih}\right) \tag{8}$$

and

$$P_{if} = \Pr\left(w_{if} \geq \max_{g\in\{0\}\cup\mathcal{F}} w_{ig}\right). \tag{9}$$

Here $P_{ij|f}$ denotes the probability of picking product j out of the G_f products of firm f while P_{if} is the probability of buying from firm f. Note that in P_{if} the symbol $\max_g$ includes consideration of the outside alternative as well. Also note that the outside option does not appear in $P_{ij|f}$ since a consumer will never buy the outside option conditional on buying from firm f.

Because computing $P_{ij|f}$ is standard (see below in Section 2.4), we focus now on the computation of the probability P_{if}. The distribution of $w_{if} = \min\{r_{if}, U_{if}\}$ can be obtained by computing the CDF of the minimum of two independent random variables.[11] This means that

$$\begin{aligned} F^w_{if}(z) &= 1-\left(1-F^r_{if}(z)\right)\left(1-F_{if}(z)\right); \\ &= F^r_{if}(z)(1-F_{if}(z)) + F_{if}(z), \end{aligned} \tag{10}$$

where F^w_{if} and F^r_{if} are the CDF's of w_{if} and r_{if}, respectively; recall that $F_{if}(z)$ is the CDF of U_{if}, the

[10] See also Armstrong and Vickers (2015) for an earlier account of the fact that sequential search models produce demands consistent with discrete choice.

[11] Specifically, if $Z = \min\{X, Y\}$ with X, Y independent, then $F_Z(z) = 1-(1-F_X(z))(1-F_Y(z))$.

maximum utility of all products sold at firm f, which has been specified above in equation (3).

To obtain the distribution of the reservation values, we can use equation (4):

$$F_{if}^{r}(z) = \Pr(r_{if} < z) = \Pr[H_{if}(r_{if}) > H_{if}(z)] = \Pr[c_{if} > H_{if}(z)] = 1 - F_{if}^{c}(H_{if}(z)).$$

Substituting this into equation (10) gives

$$F_{if}^{w}(z) = 1 - F_{if}^{c}(H_{if}(z))(1 - F_{if}(z)). \tag{11}$$

Equation (11) provides a relationship between the search cost distribution and the distribution of the w's and, because the gains from search H_{if} can only be positive, shows that even if we allow F_{if}^{c} to have negative support, only the distribution for positive values matter. This means that the part of the search cost distribution that has negative support behaves like an atom at zero.

To obtain the probability that consumer i buys from firm f, we can use that the w_{ig}'s in equation (9) are independent. We therefore take the product of each F_{ig}^{w} evaluated at w_{if} to get the CDF over all $g \neq f$, which, after integrating out w_{if}, gives this expression for P_{if}:

$$P_{if} = \int \left(\prod_{g \neq f} F_{ig}^{w}(z) \right) f_{if}^{w}(z)\, dz. \tag{12}$$

Finally, the unconditional choice probability can be obtained from s_{ij} in equation (7) by integrating out the consumer-specific variables. Denoting by τ_i the vector of all consumer-specific random variables in s_{ij}, the probability that product j is purchased is the integral

$$s_j = \int s_{ij} dF_{\tau}(\tau_i), \tag{13}$$

where F_{τ} is the CDF of τ_i.

2.4 Computation and distributional assumptions

There are two difficulties in the computation of the market shares in equation (13). The first difficulty is the computation of the buying probabilities s_{ij} in equation (7). For this, we need to calculate the probabilities $P_{ij|f}$ and P_{if}, which are given in equations (8) and (12), respectively. Since ε_{ij} is an IID draw from a TIEV distribution, $P_{ij|f}$ in equation (8) has the familiar closed form:

$$P_{ij|f} = \frac{\exp(\delta_{ij})}{\sum_{h \in G_f} \exp(\delta_{ih})} = \frac{\exp(\delta_{ij})}{\exp(\delta_{if})}.$$

There is no closed-form solution for the probability P_{if} in equation (12) for arbitrary search cost distributions, even when assuming ε_{ij} follows a TIEV distribution. Nevertheless, we can compute P_{if} by first plugging equation (4) into the distribution of w given in equation (11), deriving the density f_{if}^{w}, and then performing the integration in equation (12) numerically.

The second difficulty is the computation of the integral in equation (13). Such an integral cannot be computed analytically but, following BLP and most of the subsequent literature, can be estimated by Monte Carlo methods by drawing the demographic characteristics and random coefficients of, say, N consumers, and then computing $\hat{s}_{ij}$ for each consumer $i = 1, \dots, N$. The Monte Carlo estimator of s_j is then taken as the sample mean $\hat{s}_j = \frac{1}{N}\sum_{i=1}^{N} \hat{s}_{ij}$.

These two difficulties together make the estimation of the demand model somewhat slow because in every iteration the two integrals in equations (12) and (13) need to be computed. Although we can cope with these issues (see our estimates for normally distributed search costs in Section 4.3), it is nevertheless desirable to reduce the computational complexity of our model for further work and future applications.

In what follows we propose an alternative to numerical integration of equation (12) that significantly speeds up the estimation of the model. The idea is to use a search cost distribution for which we obtain a closed-form expression for equation (12). In that case, the computation of the buying probabilities and the first order conditions for profits maximization become much easier and estimation of the model becomes more amenable.

Proposition 1 *For search costs that are distributed according to CDF*

$$F_{if}^{c}(c) = \frac{1 - \exp\left(-\exp\left(-H_0^{-1}(c) - \mu_{if}\right)\right)}{1 - \exp(-\exp(-H_0^{-1}(c)))}, \tag{14}$$

where μ_{if} is a consumer-firm specific location parameter of the search cost distribution, the CDF of w_{if} is given by a Gumbel distribution with location parameter $\delta_{if} - \mu_{if}$, i.e.,

$$F_{if}^{w}(z) = \exp\big(-\exp\left(-(z - (\delta_{if} - \mu_{if}))\right)\big). \tag{15}$$

The proof of this proposition can be found in Appendix B, and builds on the idea that according to equation (11) there is a one-to-one relationship between the search cost distribution and the distribution of the random variable w that determines the visiting probabilities.

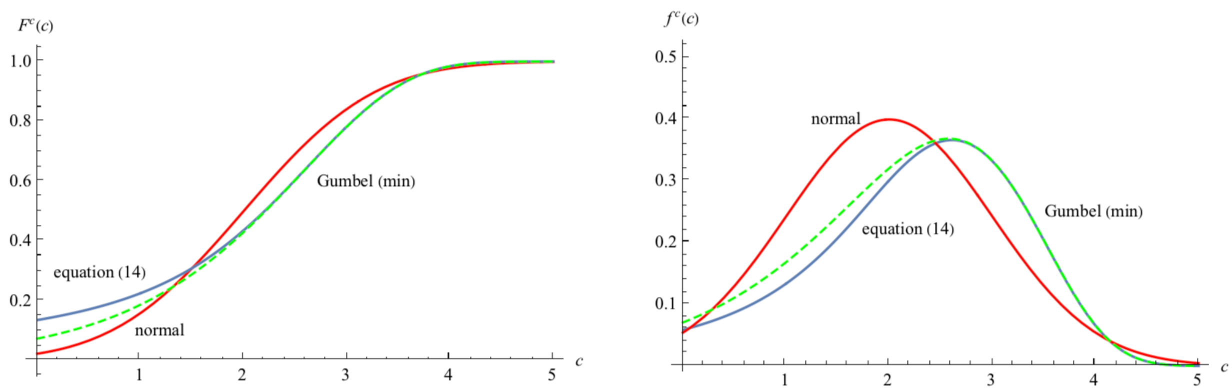
Given equation (15), calculation of the probability of buying from firm f in equation (12) is straightforward, i.e.,

$$P_{if} = \frac{\exp\left(\delta_{if} - \mu_{if}\right)}{1 + \sum_{g=1}^{F} \exp\left(\delta_{ig} - \mu_{ig}\right)}.$$

The probability of buying product j is $s_{ij} = P_{if} P_{ij|f}$ and simplifies to

$$\begin{aligned} s_{ij} &= \frac{\exp\left(\delta_{if} - \mu_{if}\right)}{1 + \sum_{g=1}^{F} \exp\left(\delta_{ig} - \mu_{ig}\right)} \times \frac{\exp\left[\delta_{ij}\right]}{\sum_{h \in G_f} \exp\left[\delta_{ih}\right]} \\ &= \frac{\exp\left(\delta_{ij} - \mu_{if}\right)}{1 + \sum_{g=1}^{F} \exp\left(\delta_{ig} - \mu_{ig}\right)} = \frac{\exp\left(\delta_{ij} - \mu_{if}\right)}{1 + \sum_{g=1}^{F} \sum_{h \in G_g} \exp\left(\delta_{ih} - \mu_{ih}\right)} \\ &= \frac{\exp\left(\delta_{ij} - \mu_{if}\right)}{1 + \sum_{k=1}^{J} \exp\left(\delta_{ik} - \mu_{ig}\right)}, \end{aligned} \tag{16}$$

where in the last line g denotes the firm that produces product k. Note that the numerator in the second expression contains δ_{ij} and not δ_{if}. An advantage of the closed-form expression for the buying probabilities in equation (16) is that it makes the estimation of the model of similar difficulty as most standard discrete choice models of demand.



(a) Search cost CDF (b) Search cost PDF

Figure 1: Search cost distribution for $\mu_{if} = 2$

Notice that a positive μ_{if} is necessary for $F^c_{if}(c)$ to be a proper distribution, which can be achieved by letting μ_{if} be a function of search cost shifters according to a log-exp function form, i.e.,

$$\mu_{if} = \log\left[1 + \exp\left(\mathbf{t}'_{\mathbf{if}}(\lambda + \mathbf{W_i}\nu)\right)\right], \tag{17}$$

where the vector $\mathbf{t_{if}}$ includes search cost shifters that are consumer and/or firm specific, such as the distance from the household to the seller and the household's income, λ and ν are search cost parameters, and $\mathbf{W_i}$ is a diagonal matrix that contains random draws from the standard normal distribution in its first entry as well as other demographic characteristics. Given this, it is straightforward to verify that the search cost distribution in equation (14) is increasing in c and takes value 1 when c approaches infinity. Moreover, it has an atom at zero, that is $F^c_{if}(0) = \exp(-\mu_{if})$, which conveniently allows for a fraction of consumers to know their match values with the products sold by firm f ex-ante; as μ_{if} increases, this share becomes smaller and becomes negligible as μ_{if} grows large.[12] Finally, we observe that, on the positive support, the distribution given by equation (14) has a shape relatively similar to the normal distribution. To see this, we plot in Figure 1 the search cost distribution in equation (14) and the corresponding density for $\mu_{if} = 2$. Also shown (in red) is a normal distribution and density with mean $\mu_{if} = 2$ and variance set to 1. Note that the two distributions are relatively similar on the positive real line.[13] The dashed green curves in Figure 1 are for a

[12]The estimated probability of the atom according to our estimates presented in Section 4.3 is rather small: it has mean 0.020 (with standard deviation 0.101) and median 0.000 across consumers and dealers.

[13]Note that in our model there is no loss of generality by allowing consumers to have negative search costs for some products. This is because if there were consumers with a negative cost of searching a particular firm, they would behave exactly in the same way as consumers with zero search cost. As a result, without loss of generality, we can allow for search cost distributions with full support, such as the normal distribution.

Gumbel distribution (for the minimum) with location parameter $\mu_{if} = 2 + \gamma = 2.577$ and scale parameter set to one.[14] This particular distribution is useful as a comparison because it is closed form and has a shape similar to the search cost distribution given by equation (14).[15]

2.5 Search probabilities

As we explain in Section 3 when discussing estimation and identification, we use individual search data in addition to aggregate data on market shares. In particular, we employ data on the share of consumers not searching beyond the outside option as well as the share of consumers searching only one time. We now compute those probabilities.

The probability that a consumer does not search beyond the outside option can be calculated as follows. Take a consumer i whose outside option is u_{i0}. This consumer chooses to refrain from searching with probability

$$v_{i0} = \Pr[u_{i0} \geq \max_{k \neq 0} \{r_{ik}\}] = \int_{-\infty}^{\infty} \int_{-\infty}^{u_{i0}} dF_i^{r_{max}}(z) dF(u_{i0}) = \int_{-\infty}^{\infty} F_i^{r_{max}}(z) f(z) dz, \tag{20}$$

where $F_i^{r_{max}}(z)$ stands for the distribution of the $\max_{k \neq 0} \{r_{ik}\}$ and is given by

$$F_i^{r_{max}}(z) = \prod_{k \neq 0} F_{ik}^r(z) = \prod_{k \neq 0} \left(1 - F_{ik}^c(H_{if}(z))\right).$$

The distribution of reservation values F_{if}^r can be derived by solving equation (10) for F_{if}^r, i.e.,

$$F_{if}^r(z) = \frac{F_{if}^w(z) - F_{if}(z)}{1 - F_{if}(z)}.$$

Using $F_{if}^w(z)$ as specified in equation (15), we get

$$F_{if}^r(z) = \frac{\exp\left(-\exp\left(\delta_{if} - \mu_{if} - z\right)\right) - \exp(-\exp(\delta_{if} - z))}{1 - \exp(-\exp(\delta_{if} - z))}. \tag{21}$$

The share of consumers not searching at all is obtained by integrating over all consumers, i.e.,

$$q_0 = \int v_{i0} dF_\tau(\tau_i), \tag{22}$$

[14] The Gumbel distribution for the minimum is the mirror image of the Gumbel distribution. It is used to model the *minimum* of a number of draws from various distributions and has CDF $1 - \exp(-\exp((z - \mu)/\beta))$.

[15] For large values of x, the exponential integral $\int_x^\infty \frac{\exp(-t)}{t} dt$ that appears in $H_0(r)$ is approximately zero, which means $x = H_0^{-1}(c) = c - \gamma$, where γ is the Euler constant. We therefore get the following closed-form expression for $F_{if}^c(c)$:

$$F_{if}^c(c) = \frac{1 - \exp(-\exp(c - \gamma - \mu_{if}))}{1 - \exp(-\exp(c - \gamma))}. \tag{18}$$

For large c the denominator goes to 1, which means that for large μ_{if} we can approximate this by a Gumbel (minimum) distribution, i.e.,

$$F_{if}^c(c) = 1 - \exp(-\exp(c - (\mu_{if} + \gamma))). \tag{19}$$

where the integral is taken over the non-observable characteristics (random coefficients) of all consumers. This integral can be estimated by Monte Carlo by drawing from the distributions of demographics and random coefficients as well as the utility distribution of the outside option.[16]

We now compute the probability a consumer i searches only one time. For this we need that the outside option is not good enough and that the best of the outside option and the first searched option is good enough to stop, i.e.,

$$v_{i1} = \sum_f \Pr\left[u_{i0} < \max_{k\neq 0}\{r_{ik}\} \text{ and } \max\{u_{i0}, u_{if}\} > \max_{k\neq 0,f}\{r_{ik}\} \text{ and } r_{if} > \max_{k\neq 0,f}\{r_{ik}\}\right].$$

The sum across f appears because the first searched option can be any of the dealerships. We can simplify this to (see Appendix C for details):

$$\begin{aligned} v_{i1} \quad = \quad & \sum_f \Pr\left[u_{i0} < r_{if} \text{ and } \max_{k\neq 0,f} r_{ik} < u_{i0}\right] \\ & + \sum_f \Pr\left[u_{i0} < \max_{k\neq 0,f} r_{ik} \text{ and } \max_{k\neq 0,f} r_{ik} < u_{if} \text{ and } \max_{k\neq 0,f} r_{ik} < r_{if}\right]. \end{aligned}$$

This probability can be written as

$$v_{i1} = \sum_f \int_{-\infty}^{\infty} \left(F_{if}^{w}(y) - F_{if}^{r}(y)\right) F_{i,-f}^{r}(y) f(y) dy + \sum_f \int_{-\infty}^{\infty} F_{i,-f}^{r}(x) F(x) f_{if}^{w}(x) dx. \tag{23}$$

where $F_{i,-f}^r$ and $f_{i,-f}^r$ are the CDF and PDF of $\max_{k\neq 0,f}\{r_{ik}\}$.

The share of consumers searching only one time is then given by

$$q_1 = \int v_{i1} dF_\tau(\tau_i). \tag{24}$$

Again, this probability can be computed by Monte Carlo by jointly sampling from the taste distributions as well as the utility distribution of the outside option (for the first integral of equation (23)) and the distribution of w_{if} (for the second integral).[17]

2.6 Supply side

Although we do not include the supply side when estimating the model, we do need to specify a supply side model for several of the counterfactual simulations that are part of Section 4.6. We assume firms maximize their profits by setting prices, taking into account prices and attributes of competing products as well as the locations of all sellers. Let $\mathbf{p}$ denote the vector of Nash equilibrium prices. Assuming a pure strategy

[16] Specifically, let V be a TIEV draw (obtained as $-\log(-\log U)$ with U uniform on $(0,1)$). A draw from the density of the outside option is then V.

[17] More specifically, let V be a TIEV draw (obtained as $-\log(-\log U)$ with U uniform on $(0,1)$), which we can use directly for the first integral of equation (23). For the second integral of equation (23), from the draws of the random coefficients we compute $\delta_{if} - \mu_{if}$. A draw with density f_{if}^w is then $\delta_{if} - \mu_{if} + V$.

equilibrium exists, any product j should have a price that satisfies the first order condition

$$s_j(p) + \sum_{r \in \mathcal{G}_f} (p_r - mc_r) \frac{\partial s_r(p)}{\partial p_j} = 0.$$

To obtain the price-cost markups for each product we can rewrite the first order conditions as

$$\mathbf{p} - \mathbf{mc} = \boldsymbol{\Delta}(\mathbf{p})^{-1} \mathbf{s}(\mathbf{p}), \tag{25}$$

where the element of $\boldsymbol{\Delta}(\mathbf{p})$ in row j column r is denoted by Δ_{jr} and

$$\Delta_{jr} = \begin{cases} -\dfrac{\partial s_r}{\partial p_j}, & \text{if } r \text{ and } j \text{ are produced by the same firm;} \\ 0, & \text{otherwise.} \end{cases}$$

The derivation of the partial derivatives of the market shares with respect to price is straightforward in case of Gumbel distributed w's.[18] Details on the derivation of the market share derivatives with respect to price for general search cost distributions are provided in Appendix D. Note that for market share derivatives it matters whether consumers observe deviation prices before or after search. In our application, because consumers can easily observe list prices while being at home, we adopt the assumption that consumers observe (deviation) prices before they start searching. Notice that this assumption differs from most of the literature on consumer search for differentiated products (Wolinsky, 1986; Anderson and Renault, 1999) and, as demonstrated in recent work (see, e.g., Armstrong and Zhou, 2011; Haan, Moraga-González, and Petrikaitė, 2018), this assumption has implications for the behavior of prices and search costs.[19]

3 Estimation

Our estimation procedure closely resembles Berry, Levinsohn, and Pakes (2004), who use a combination of micro and aggregate data to estimate a differentiated products demand system. However, whereas Berry, Levinsohn, and Pakes use moments based on second-choice data, we use moments based on search and purchase data, which allows us to explicitly take consumers' search behavior into account. We first give a brief outline of the estimation procedure and then discuss the calculation of the moments we are using for estimation in more detail. We finish with a discussion of the identification of the key parameters in the model.

[18] For the case of Gumbel distributed w's, the own-price derivative $\partial s_j / \partial p_j = \int (\partial s_{ij} / \partial p_j) dF_\tau(\tau_i)$, where $\partial s_{ij} / \partial p_j = \alpha_i s_{ij} (1 - s_{ij})$. The derivative $\partial s_j / \partial p_k = \int (\partial s_{ij} / \partial p_k) dF_\tau(\tau_i)$, where $\partial s_{ij} / \partial p_k = -\alpha_i s_{ij} s_{ik}$.

[19] In a standard search model a firm chooses its price to maximize the payoff from the consumers who visit. By changing the price a firm thus affects the selling probability, but not the visiting probability. In contrast, when prices are observed from home like in our model, changing the price affects both the visiting and the buying probability. While in most standard models prices increase in search costs, in models where prices are observable before search, prices can be decreasing in search costs. We will return to this point later in the paper when we study the equilibrium effects of moving towards a full information equilibrium.

3.1 Outline of the estimation procedure

Our goal is to estimate the parameters of the model, which include the utility parameters α, β, and σ, as well as the parameters of the search cost distribution λ and ν. We use aggregate data on sales, prices, and other product characteristics, combined with micro data on purchases and search behavior. Following Berry, Levinsohn, and Pakes (2004), the nonlinear parameters of a (full information) model can be estimated using micro data by allowing for choice-specific constants, which capture the linear part of utility δ (see also Goolsbee and Petrin, 2004; Train and Winston, 2007). The estimated δ's can then be used in a second step to estimate the linear utility parameters. Alternatively, by imposing additional restrictions on the joint distribution of $(\xi, \mathbf{x})$, the linear and nonlinear parameters of the model can be estimated jointly. In what follows, we use the two-step procedure. Even though we loose efficiency by not imposing restrictions on the joint distribution of $(\xi, \mathbf{x})$ when estimating the nonlinear parameters, the main advantage of using the two-step procedure is that we do not loose consistency if the restrictions are not correct. This is especially important in our setting because the unobserved quality variable ξ is not only likely to be correlated with prices, but also potentially correlated with distances from consumers to sellers. This means that the joint estimation of the search parameters and the linear utility parameters requires an instrument for distance, which may be difficult to obtain in practice.[20]

We estimate the model by generalized method of moments (GMM) using two sets of moments. The first set of moments relates demographic information to buying decisions and is useful for estimating the nonlinear utility parameters of the model. The second set of moments relates demographic information to search decisions, which allows us to estimate the search parameters of the model. We estimate δ by matching the observed market shares to the model's predicted market shares. For this we use the result in Berry (1994) and BLP that δ can be computed as the unique fixed point of a contraction mapping. The fact that this mapping is also a contraction in our search model follows from the fact that the first order derivatives of the market shares with respect to the unobserved characteristics have the same form as in BLP (see Appendix E for more details).[21]

3.2 Moments and GMM estimation

In our application we use individual-level survey data that provides information on each respondent's latest car purchase, as well as their search behavior related to that car purchase. We use moments that are based on these survey responses as well as corresponding model predictions—following Berry, Levinsohn, and Pakes (2004) we let the GMM estimation routine select the nonlinear parameters of the model such that the difference between the predicted probabilities and the observed probabilities in the survey data is minimized.

[20] In an earlier version of this paper we did jointly estimate the linear and nonlinear parameters of the model using variables that relate to the cost of operating a dealership (variation in property vales and local taxes) as instruments for distance from consumers to dealerships.

[21] We use the SQUAREM algorithm (Varadhan and Roland, 2008) instead of standard contraction iterations, which is found to be faster and more robust than the standard BLP contraction (Reynaerts, Varadhan, and Nash, 2012).

To relate demographic information to buying decisions, we follow Petrin (2002) in using moments based on the following type of conditional expectations:

$$E\left[\mathbb{1}\{a_i \in T\} | q_i \in R_k\right], \quad k = 1, 2 \text{ where } \mathcal{R} = \{R_1, R_2\}. \tag{26}$$

In this expression, $\mathbb{1}\{a_i \in T\}$ is an indicator for the event that consumer i makes choice $a_i \in \{0, 1, \ldots, J\}$ from a certain group of products T, and q_i is a generic demographic characteristic, which is partitioned in two subsets according to $\mathcal{R}$. For instance, we use two moments that match the model's predicted average probability of buying a new car conditional on income level to the survey data, i.e.,

$$E[\mathbb{1}\{i \text{ purchases new vehicle}\} \mid \{y_i < \overline{y}\}],$$
$$E[\mathbb{1}\{i \text{ purchases new vehicle}\} \mid \{y_i \geq \overline{y}\}],$$

where $\mathbb{1}\{i \text{ purchases new vehicle}\}$ is an indicator for the event that consumer i purchases a new vehicle and $\{y_i < \overline{y}\}$ and $\{y_i \geq \overline{y}\}$ correspond to the events that consumer i is in the low or high income group, respectively. Other moments of this type are listed in Table 6 of Section 4.3 and include the probability of purchasing an MPV or SUV conditional on family size, the probability of buying a family car conditional on the presence of children in the household, the probability of buying a luxury brand conditional on income, and the probability of buying a new vehicle conditional on the number of dealerships nearby.

The type of moments we use to relate demographic information to search decisions is similar to the type given in equation (26), but with the number of searches as the choice variable instead of the type of car bought. Specifically, these moments match to the survey data the model's predicted average probability of searching once as well as searching more than once conditional on income level, the number of dealers nearby, whether the head of household is senior, and whether there are kids in the household.

We use GMM to estimate the nonlinear parameters of the model. The GMM estimator of the nonlinear parameters θ_2 is

$$\hat{\theta}_2 = \arg\min_{\theta_2} g\left(\theta_2\right)' \Xi g\left(\theta_2\right),$$

where $g(\theta_2)$ is the column vector of moments and Ξ is a weighting matrix. In Appendix G we discuss several alternatives for Ξ; however, in our application we use $\Xi = I_M$, where I_M is the identity matrix of dimension M, which is the number of moments.

In a second step, assuming θ_2 known, we can obtain θ_1 as the linear IV estimator

$$\widehat{\theta}_1 = \left(\mathbf{x}'\mathbf{z}\Xi\mathbf{z}'\mathbf{x}\right)^{-1} \mathbf{x}'\mathbf{z}\Xi\mathbf{z}'\delta,$$

where $\mathbf{z}$ is a matrix of exogenous instruments. Note that when calculating the standard errors for θ_1, we take into account that θ_2 is estimated.

3.3 Identification

Identification of the model is based on the assumption that we observe aggregate market share data and a large number of individual-level search and purchase data. It is important to point out, however, that if the search cost and utility specifications do not contain common covariates then the model can be identified without search data. In this case the main identification argument is that the individual-level purchase data identify the mean utility values (i.e., the components of δ) and the nonlinear parameters (i.e., the price/income coefficients $\alpha_{(1)}$, $\alpha_{(2)}$, the random coefficients σ, and the search cost parameters λ and ν) based on parametric random coefficient discrete choice identification arguments (see also Berry, Levinsohn, and Pakes, 2004).

Since the number of mean utilities is relatively large, it is not practical to estimate them directly. Therefore, following Berry, Levinsohn, and Pakes (2004), we compute them as a function of the nonlinear parameters by using the observed market shares, as described in Section 3.1. Given the estimates of δ, the identification of the linear utility parameters is straightforward. Following the literature, our main identification assumption is that the demand unobservables ξ_j are mean independent of a set of exogenous instruments Z, i.e., the conditional moment restrictions

$$E[\xi_j|Z] = 0, \quad j = 1, \ldots, J$$

hold. Since ξ_j captures unobserved quality differences between products, it is likely to be positively correlated with price, which leads to the usual endogeneity problem that arises when estimating aggregate demand models. To deal with this endogeneity problem, we follow the literature and use instrumental variables. Identification when prices are endogenous requires the use of appropriate instruments. Following Reynaert and Verboven (2014), we use the vector of predicted prices of a regression of price on cost shifters as an instrument for price (see also Gandhi and Houde, 2016). These are estimates of $E\left[p_j|w_j\right]$, which is the optimal instrument under the assumption that prices are equal to marginal costs, where w_j is the cost shifter of product j.

The moment conditions we use for estimation of the model are tightly linked to the nonlinear parameters we estimate. The first set of moment conditions relates purchase probabilities to demographics. We use several moment conditions for the interaction of household demographics and product attributes as well as for the random coefficients that we include in the utility. For instance, in our application we include moment conditions for the interaction of price and income to identify the nonlinear price coefficients and use moment conditions for the interaction of larger car types (MPV/SUV and family car) and household composition (family size and dummy for children) to identify corresponding utility parameters as well as a random coefficient on size. The second set of moment conditions relates search probabilities to demographics. Here we use a separate moment condition for each search cost covariate.

A practically relevant issue regarding identification is that there may potentially be common covariates in utility and search cost, whose effects should be separately identified. Unfortunately these effects cannot be separately identified without search data in the version of the model in which the w's are Gumbel distributed, that is, when the distribution of the w's is given by equation (15). Here we show through formal arguments that when search data is available, as in our application, the effects of common covariates (for instance, when distance enters both the search cost and utility specification) are separately identified. In order to do so, we use the buying probability as well as the probability of not searching. We can argue that, based on parametric identification arguments for the random coefficient logit model, the mean utilities, the nonlinear parameters, and the total effects of the variables involved in $\delta_{ij} - \mu_{if}$ can be identified based on the individual-level purchase data. For example, if the distance from consumer i to seller f appears linearly both in δ_{ij} and μ_{if} with coefficients β_δ and β_μ then we can identify the difference $(\beta_\delta - \beta_\mu)$ but not β_δ and β_μ separately. Intuitively, we know that the decision from which seller to buy depends on both the maximum utility from the products of the seller and the search cost. Therefore, using data on firm choice we can only identify the total effects from a combination of these quantities. Consider now the share of consumers not searching, which is given by equation (22), and can be written as[22]

$$q_{i0} = \int \int \prod_{k \neq 0} F^r_{ik}(z) f(z) \; dz \; dF_\tau(\tau_i).$$

Since the distribution of reservation utilities can be written as

$$F^r_{ij}(z) = \frac{\exp(-\exp(\delta_{ij} - \mu_{if} - z)) - \exp(-\exp(\delta_{ij} - z))}{1 - \exp(-\exp(\delta_{ij} - z))},$$

the expression $\prod_{k \neq 0,f} F^r_{ik}(x)$ in the integral contains both $\delta_{ij} - \mu_{if}$ and δ_{ij} for $f = 1, \ldots, F$. As explained above, the expressions $\delta_{ij} - \mu_{if}$ are identified from individual-level purchase data, so in the share q_{i0} the only unknowns left are the utility parameters involved in the δ_{ij}'s. These can be identified based on parametric random coefficient discrete choice model identification arguments, because in the survey data we can observe the share of consumers not searching conditional on various demographic characteristics. Intuitively, search probabilities carry information on the consumer's preferences for products beyond the information provided by the buying choices, and this yields separate identification of the two sets of parameters. In our application, we include moments that relate the probability of searching at least twice, which is defined as $1 - q_{i0} - q_{i1}$, which allows us to separately identify the distance variable in the search cost specification from the distance variable in the utility specification.

Finally we briefly consider the general model. In this case there is no closed form expression for the distribution of reservation values, which makes it difficult to separate the total effects in a way similar to the case of Gumbel distributed w's. However, we believe that in this case it is also valid that different types of

[22] An alternative is to consider probabilities of searching a given seller but not buying any product, which have a similar structure. Specifically, the probability that consumer i searches only firm f and opts for the outside alternative is

$$v_{if0} = \int_{-\infty}^{\infty} \left(1 - F^r_{if}(x)\right) \prod_{k \neq 0,f} F^r_{ik}(x) \, F_{if}(x) \, f(x) \, dx.$$

Since these probabilities vary with the dealers they offer more variation for identification. Nevertheless, our data contain relatively few such observations.

probabilities can identify different combinations of the maximum utility and search cost parameters, which eventually yields separate identification of the parameters in a way similar to solving a system of nonlinear equations. The intuition for this is similar to that of the Gumbel-distributed w case: the reservation values that guide search decisions respond differently to changes in search costs than to changes in utility, which, given that the combined effect of utility and search cost variables on purchases is identified using purchase data only, allows us to separate the combined effect into a utility part and a search part.

4 Application

In this section we apply our model to the Dutch market for new cars. In 2008 approximately 500,000 new passenger cars were sold in the Netherlands, which makes it the sixth biggest car market in Europe in terms of sales. The top selling make is Volkswagen, which in 2008 had a market share of 9.2 percent, followed by Ford (8.7 percent), Opel (8.3 percent), Peugeot (8.2 percent), and Toyota (8.0 percent). The most popular car models tend to be small in size—in 2008, the two top selling models were the Peugeot 207 and Opel Corsa (both are in the so-called supermini class) followed by the Volkswagen Golf (small family car class).

In Section 4.1 we discuss the aggregate and individual level data we use for estimation of the model. In Section 4.2 we use the individual level survey data to provide some background information on how consumers search in this market. In this section we also provide some reduced-form evidence that search behavior is related to demographics such as income, family size, age, and distances to dealerships. In the remainder of Section 4, we report the results for estimation of our search model, robustness results, as well as the results of several counterfactual simulations.

4.1 Data

We use a combination of aggregate and individual level data to estimate the model. Our aggregate data consists of prices, sales, physical characteristics, and locations of dealers of virtually all cars sold in the Netherlands between 2003 and 2008. We include a model in a given year if more than fifty cars have been sold during that year; this means "exotic" car brands like Rolls-Royce, Bentley, Ferrari, and Maserati are excluded. This leaves us with a total of 320 different models that were sold during this period—in any given year about 230 different models. We treat each model-year combination as one observation, which results in a total of 1,382 observations.

The data on product characteristics are obtained from *Autoweek Carbase*, which is an online database of prices and specifications of all cars sold in the Netherlands from the early eighties until now.[23] Characteristics include horsepower, number of cylinders, maximum speed, fuel efficiency, weight, size, and dummy variables for whether the car's standard equipment includes air-conditioning, power steering, cruise control, and a board computer. Transaction prices are not available, so all prices are listed (post-tax) prices, normalized to 2006 euros using the Consumer Price Index.[24]

[23] See autoweek.nl/carbase.

[24] The tax when buying a new car in the Netherlands consists of a sales tax as well as an additional automobile tax. The

In practice, prices for cars are typically determined by bargaining between the car sales person and the consumer. Because we only observe list prices, bargaining is difficult to incorporate in our model; instead we assume that the list price is the price the consumer pays. Not taking bargaining into account affects the analysis in two ways. First of all, if bargaining is important, consumers may visit multiple dealerships in order to get a better deal. Ignoring bargaining therefore means that we are implicitly assuming that the main reason for visiting dealerships is to learn more about the products. Secondly, bargaining affects the prices consumers pay, which means bargaining also affects the gains from search. However, unless the gains from bargaining are very different across brands, this is not very likely to affect the results. Moreover, most car dealerships in the Netherlands have very few cars in stock due to space limitations, which means there is typically less incentive for car dealerships to offer discounts on the list price in comparison to for instance the United States, where it is more common to have a large number of cars in stock. Also note that an alternative interpretation of our search model is that consumers are searching and bargaining for prices and ε_{ij} relates to the difference between the list price and the actual prices consumers pay at the store. This is consistent with a pure characteristics interpretation of the model (as in Berry and Pakes, 2007), with heterogeneity in preferences coming from the random coefficients only.

We have supplemented the aggregate data with several macroeconomic variables, including the number of households and average gasoline prices, as reported by Statistics Netherlands. The total number of households allows us to construct aggregate market shares (calculated as sales divided by the number of households), while average gasoline prices are used to construct our kilometers per euro (KP€) variable, which is calculated as kilometers per liter (KPL) divided by the price of gasoline per liter.

We define a firm as all brands owned by the same company. We use information on the ownership structures from 2007 to determine which car brands are part of the same parent company—the 39 different brands in our sample are owned by 16 different companies. For instance, in 2007 Ford Motor Company owned Ford, Jaguar, Land Rover, Mazda, and Volvo.

Table 1 gives the sales weighted means for the main variables we use in our analysis. The number of models has increased from 213 in 2003 to 241 in 2008. Sales were lowest in 2005 and peaked in 2007. Prices have been going up mostly in real terms, although 2008 saw a sharp decrease, due to a tax cut in 2008 (see also footnote 24) and possibly as a result of the onset of the most recent recession. The share of European cars sold shows a downward trend, mainly to the benefit of cars that originate from East Asia. The ratio of horsepower to weight has been increasing steadily. The share of cars with cruise control as standard equipment increased in the first half of the sampling period, but then decreased somewhat. Cars have become more fuel efficient during the sampling period. Nevertheless, as shown in the KP€ column of Table 1, fuel efficiency has not increased enough to offset rising gasoline prices—the number of kilometers that can be traveled for one euro has decreased over the sample period. The share of luxury brand cars, where luxury

sales tax (BTW) in the period 2003-2008 was 19 percent. The automobile tax (BPM) was 45.2 percent of the pre-tax price during most of the sampling period, but was lowered to 42.3 percent in February 2008. The automobile tax paid also depends on whether the car uses diesel or gasoline (gasoline users deduct €1,540 from the pre-tax price of a car before applying the automobile tax (€1,442 during most 2008), while diesel users add €328 (€308 in 2008)). Moreover, from July 2006 on there are additional additions or deductions to the pre-tax price that are based on the energy efficiency of the car and whether the car is a hybrid or not.

Table 1: Summary statistics

Year	No. of Models	Sales	Price	European	HP/ Weight	Size	Cruise Control	KPL	KP€	Luxury Brand
2003	213	481,913	19,562	0.762	0.787	7.153	0.229	14.480	12.497	0.081
2004	228	476,581	19,950	0.749	0.788	7.184	0.308	14.696	11.737	0.080
2005	233	457,897	20,540	0.727	0.794	7.270	0.301	14.861	10.987	0.096
2006	231	475,636	20,367	0.715	0.804	7.271	0.308	15.120	10.707	0.092
2007	236	495,091	20,509	0.712	0.810	7.330	0.281	15.112	10.356	0.093
2008	241	489,584	18,613	0.714	0.813	7.271	0.293	15.813	10.290	0.095
All	1,382	479,450	19,916	0.730	0.799	7.247	0.286	15.018	11.091	0.089

Notes: Prices are in 2006 euros. All variables are sales weighted means, except for the number of models and sales.

brands include Audi, BMW, Cadillac, Jaguar, Land Rover, Lexus, Mercedes-Benz, and Porsche, increased in 2005 to 9.6 percent of total sales, and has stayed relatively constant throughout the rest of the sample.

In addition to car characteristics we use information on the location of car dealerships and combine this with geographic data on where people reside to construct a matrix of distances between households and the different car dealerships. These distances are later used to proxy the cost of visiting a dealership to learn all product characteristics of a vehicle. We also use data on the distribution of household characteristics as search cost covariates.

Our demographic and socioeconomic data on households are obtained from Statistics Netherlands. These data are available at various levels of regional disaggregation (neighborhoods, districts, city councils, counties, and provinces). Since the purpose of our study is to estimate the importance of search costs, we choose to work at the highest level of regional disaggregation, that is, at the neighborhood level. This permits us to proxy the costs of traveling to the different car dealers rather accurately. Statistics Netherlands provides a considerable amount of useful demographic and socioeconomic data at this level of disaggregation.

For every neighborhood, the demographic data include the number of inhabitants and their distribution by age groups, the number of households, the average household size, the proportion of single-person households, and the proportion of households with children. We only include neighborhoods with a strictly positive number of inhabitants, which leaves us with a total of 11,122 neighborhoods for 2007.[25] Most neighborhoods are relatively small; the mean number of inhabitants is 1,471.

In addition to demographic data we have information on the exact location of each neighborhood on the map of the Netherlands. Using a geographical software package we use this information to construct a proxy for the distance that needs to be travelled when visiting a car dealership. To be able to do this, for every brand we have first obtained the addresses of all its dealerships in the Netherlands. For instance, Saab has a total of 37 dealers in the Netherlands. Since we have the exact addresses of the 37 dealerships of Saab, for every neighborhood, we can compute the Euclidean distance from the center of the neighborhood to the

[25] There are 284 neighborhoods for which the number of inhabitants is zero. These are neighborhoods that tend to be located in industrial areas, ports, and remote rural areas. There are a few neighborhoods for which we miss some of the relevant variables. To complete the data set we proceed by using information obtained at lower levels of disaggregation (districts or city councils).

closest Saab dealer. We do this for all car manufacturers and obtain a matrix of 11,122 by 39 containing the minimum distances from the center of a neighborhood to a car dealer. There is a lot of variation in the distances to the closest dealer of each brand across neighborhoods. Volvo, arguably a brand similar to Saab but with very different sales, has a total of 114 dealerships in the whole country—clearly on average the minimum distance to a Volvo dealer is much smaller than the minimum distance to a Saab dealer. A similar picture arises for other brands. For instance, Audi has 161 dealers, whereas BMW has only 57, even though both brands are active in the luxury segment of the market and in this case both have similar sales figures. Table 2 gives some descriptive statistics for the distances to the nearest dealer for all the car brands in our data. Opel is the most accessible: almost 79% of all households live within 5 kilometer from an Opel dealer. Porsche has the lowest percentage of households within 5 kilometer: only 6.4% of households is within easy reach.

Table 2: Descriptive statistics for distances

Brand	Number of dealerships	Number of cars sold in 2008	Weighted average distance	Percentage of households within 5 km
Alfa Romeo	75	3,050	7.96	42.1
Audi	161	16,738	4.68	68.0
BMW	57	15,170	8.35	38.5
Cadillac	15	198	18.96	14.3
Chevrolet/Daewoo	137	7,421	5.00	64.9
Chrysler/Dodge	32	2,589	12.51	25.4
Citroën	162	24,139	4.40	69.9
Dacia	98	4,549	7.25	55.3
Daihatsu	99	9,186	6.23	52.9
Fiat	142	21,010	4.86	66.6
Ford	233	42,504	3.66	78.1
Honda	65	8,479	7.99	39.6
Hyundai	138	17,433	5.08	63.2
Jaguar	16	752	18.32	14.6
Jeep	43	784	10.13	36.2
Kia	115	12,236	5.74	54.7
Lancia	50	761	11.12	34.5
Land Rover	20	1,421	14.43	16.9
Lexus	13	1,044	19.55	16.4
Mazda	121	7,582	5.57	57.7
Mercedes-Benz	83	10,446	6.58	48.0
Mini	37	3,417	11.41	29.2
Mitsubishi	108	7,805	5.57	53.9
Nissan	114	10,259	6.02	51.9
Opel	233	40,405	3.55	78.7
Peugeot	187	40,250	4.14	72.6
Porsche	8	531	25.70	6.4
Renault	196	37,526	4.18	70.9
Saab	37	1,938	11.46	24.3
Seat	127	13,061	6.08	57.3
Skoda	97	9,461	6.18	50.5
Smart	21	952	13.65	20.1
Subaru	45	1,422	10.96	27.7
Suzuki	124	14,547	5.00	60.2
Toyota	141	38,997	4.69	66.6
Volkswagen	188	45,034	4.05	74.1
Volvo	114	16,487	5.33	61.3

Notes: Averages are weighted by number of households in each neighborhood.

Our last dataset is obtained from two separate surveys that were administered by TNS NIPO (tns-nipo.com), a Dutch survey agency, in 2010 and 2011. As part of their ongoing investigation (named "De Nederlandse Automobilist") on the characteristics and behavior of Dutch motorists, over 1,200 car drivers are surveyed every year. These drivers are part of TNS NIPObase, which is a panel of around 200,000 respondents. The dataset contains 2,530 observations—1,297 for the survey carried out in 2010 and 1,233 for the 2011 survey. Our data consists of a subset of the questions in the survey and focuses on two aspects of consumer decision making: product orientation and the purchase decision. Each observation corresponds to a single respondent. All questions in the survey relate to the car that is owned by the respondent at the time of questioning. We have information about the make and model of that car, as well as the year in which the car was bought. We also know whether the car they bought was used or new. In addition, the respondents answered questions that provide useful information on how consumers search in this market. In particular, respondents reported the brands of the dealerships they visited before buying the car, and for which brands they made a test drive at the dealer.[26] Respondents also reported how many different dealerships they visited of the same brand as the brand they purchased, and the maximum number of different dealerships they visited on the same day.[27] Finally, the respondents answered questions about their household income, household size, age, whether there are children living in the household, and zip code. We use these data for the calculation of the moments that are used to estimate the nonlinear parameters of the model. We exclude respondents for which we do not observe income or a zip code, which leaves us with 2,024 observations. For the moments that are used to identify the nonlinear parameters of the model, we focus on new car purchases in 2008 only—we assume that all respondents that did not buy a new car in 2008 went for the outside option, which includes not buying a car and buying a used car. According to data from the survey, slightly over 7 percent of the respondents bought a new car in 2008, which equals the share of households in the Netherlands that bought a new car in 2008.[28]

4.2 Search behavior

Before moving to the estimation of our search model in the next section, in this section we use the survey data described above to provide some background information on search behavior in the Dutch car market. Figure 2 gives a histogram of the number of dealers of different brands visited (conditional on searching) by respondents who bought a new car. We focus on purchases between 2003 and 2008 only, since that period overlaps with the aggregate data we will be using for the main analysis, giving us a total of 1,250 respondents who bought a car between those years, of which 540 were new cars.

The average number of dealers of different brands visited for new car purchases is 2, which is slightly

[26] The specific questions that were asked are: "For which of the following brands did you visit a dealer?" and "For which of the following brands did you make a test drive at the dealer?"

[27] Approximately two-thirds of respondents who bought a new car and visited dealerships of distinct brands did not make those visits on the same day, which is consistent with a sequential search strategy.

[28] The survey data is from 2010 and 2011, and given that the survey is about the last car bought, it is likely that purchases of new cars in earlier years are underrepresented (if a consumer bought her last car in 2008, she may have bought a car in 2004 as well; the problem is that the 2004 purchase will not be in the survey data). For construction of the moments it is important that the probability of buying a new car from the survey data reflects the aggregate probability of buying a new car, and we found this to be the case for purchases in 2008 only, which corresponds to the most recent year in our market share data.

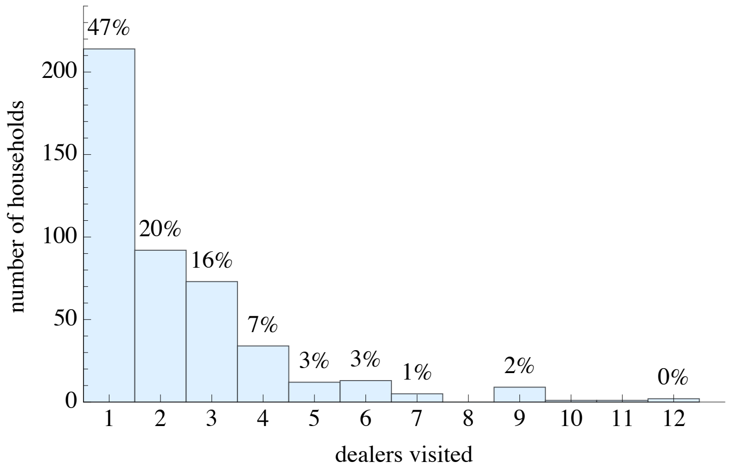


Figure 2: Histogram number of dealers visited for new car purchases

below the average of 3 dealers found by Moorthy, Ratchford, and Talukdar (1997) for the United States in the early nineties. Although 47 percent of respondents who actively searched visited at least one dealer of a single brand only, the distribution is positively skewed with some consumers visiting dealers of as many as 12 different brands.[29]

Consumers visit car dealerships for various reasons, such as learning more about the characteristics of cars or to bargain. A limitation of our study is that we do not observe data that helps discern among these potential reasons for visiting various dealers of the same brand. For instance, we cannot tell from our data to what extent consumers visit dealerships to bargain over prices. However, we do know from the survey that in 45 percent of the dealer visits a test drive was involved. Moreover, among those who visited one or more dealers to shop for a new car, over 75 percent made at least one test drive at one of the visited dealerships. Because test drives are typically done to learn about car characteristics that can hardly be learnt otherwise (including whether the car is a good fit), the survey data is consistent with a search model in which an important reason for visiting car dealers is to learn more about the product.[30]

Besides information on dealer visits, the survey data contains demographic information such as zip code, household income, family size, age, and household composition. To obtain a better insight into what explains the differences in search behavior across the respondents, we run several regressions. We first use the information on dealer visits from the survey to investigate what determines the number of dealer visits. Column (A) of Table 3 gives the results of an ordered probit regression in which we explain the number of

[29] Approximately 16 percent of respondents who bought a new car claim they have not visited any dealers. A relatively large proportion of the non-visits are (company) car leases—although only 18 percent of new car purchases in the survey data are company car leases, they represent over 25 percent of the non-visits. Other possible explanations for purchases occurring without any dealer visit are online car purchases, and parallel imports. Buying a new car online is only possible in the Netherlands since 2006, when the online car dealer nieuweautokopen.nl started operating.

[30] Close to 47 percent of the consumers who searched for and bought a new car visited only one dealer of a given brand, making it unlikely that these consumers were visiting dealers for price shopping. Around 36 percent visited 2 or 3 dealers of the same brand. While price shopping could explain this behavior, there are other potential explanations for this observation. For example, consumers visit several dealers of the same brand because not all dealers have cars available for test driving for the models they are interested in. Though most consumers in the Netherlands order new cars which means that they have to wait for delivery, some consumers value immediacy and visit various dealers to learn what cars are on inventory.

dealer visits by the log of household income, a dummy for whether there are kids living in the household, a dummy for whether the partner of the head of the household is 65 years or older, and a dummy for whether the respondent purchased a new car. In addition we include year fixed effects as well as fixed effects for the make that was ultimately bought. As shown in the table, the log income coefficient is positive and highly significant. Even though this suggests that higher income leads to more search, this does not necessarily mean that higher income respondents have lower search costs because more wealthy consumers also tend to buy more expensive cars and the benefits from search may be higher for this type of cars. Although only significant at the ten percent level, having children in the household reduces the number of searches, while being older increases the number of searches. The new car dummy indicates that people visit more dealers when buying a new car than when buying a used car, which may reflect the fact that it is less common to buy a used car at a car dealer than a new car. In specification (B) we focus on new car purchases only. Although this does not change the income coefficient much, the effect of children in the household is now twice as large, whereas the senior dummy is no longer significantly different from zero.

Table 3: Dealer visits

	Number of dealer visits				Probability of dealer visit			
	(A)		(B)		(C)		(D)	
Variable	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
Coefficients								
log(income)	0.215	(0.067)***	0.203	(0.099)**	0.085	(0.024)***	0.067	(0.032)**
kids	-0.154	(0.083)*	-0.303	(0.140)**	-0.081	(0.031)***	-0.115	(0.048)**
senior	0.185	(0.093)**	0.079	(0.125)	0.082	(0.033)**	0.043	(0.042)
new car	0.602	(0.074)***			0.246	(0.027)***		
distance					-0.070	(0.020)***	-0.079	(0.028)***
Pseudo R^2	0.052		0.041		0.102		0.098	
# obs	1,013		442		35,385		15,028	

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Data is for 2003-2008. All specifications include year and make dummies. Specifications (B) and (D) use data for new car purchases only. For some respondents income is missing, so the number of observations used to estimate specifications (A) and (B) is less than the number of respondents buying a car (1,250) and those buying a new car (540). Distance is measured in units of 10 kilometers.

To see how the physical distance from a respondent to a dealer location affects decisions on whether or not to visit a dealer, in specifications (C) and (D) we regress an indicator for whether a dealer is visited by a respondent on the same set of covariates as before, as well as the Euclidian distance between the centroid of the zip code where the respondent resides and the nearest dealer of each of the car brands in our data. Specification (C) takes all car purchases into account, whereas in specification (D) we only focus on new car purchases. The effects of income, kids, and senior are similar to the results for the ordered probit regressions: income is positively related to a dealer visit, children negatively, and senior positively, although the effect for the latter disappears when we condition on new car purchases. In both specifications distance has a negative impact on the probability of visiting a dealer and is highly significant. This finding is consistent with the fact that, according to the survey, 41 percent of new car buyers responded that distance was a factor they took into account when determining which dealers to visit. That distance matters is also reported in related work. Albuquerque and Bronnenberg (2012), using individual car transaction data in the San Diego metropolitan area between 2004 and 2006, find that consumers have a strong disutility for travel when buying a car. Similarly, Nurski and Verboven (2016) find that dealer proximity is an important determinant of demand for automobiles in Belgium. While these papers interpret distance as a transportation cost parameter that directly lowers utility, we treat it also as a variable that increases the cost of searching cars and creates limited and heterogeneous information.

The above findings are consistent with search frictions playing a role in the car market. Not only is there substantial heterogeneity across respondents in how many dealers they visit, but also consumers tend to visit car dealers to learn more about the characteristics of the cars they sell. Moreover, demographics such as income and location seem to play a role in the decision which dealers to visit.

4.3 Estimation results

In this section we report the estimation results for the search model. We also report results for the full information model, so we can see how taking into account search frictions affects the estimates of demand parameters and markups. For the estimation of our main specification we use the procedure outlined in Section 3, which, by using moments that are based on individual-level search and purchase data from the survey, facilitates identification of the random coefficients as well the parameters of the search cost distribution. As discussed in Section 3.3, using particular search moments allows us to separately identify common covariates. Moreover, by using a two-step procedure that separates the estimation of the mean utility parameters from the nonlinear parameters, estimation of the nonlinear parameters does not rely on how we specify the linear part of the utility function, which means that correlation between search cost shifters and the unobserved characteristic ξ will not affect estimation of the search cost parameters.

Conditional logit model

We first show results for the conditional logit model, which we estimate using aggregate data only. An advantage of the logit model is that it allows us to explore the effects of search frictions in a very simple setting, which is particularly useful for studying how the model behaves under different distributional assumptions for the search cost distribution. However, by using only aggregate data to estimate the model, nonparametric identification of common covariates is not achievable, so any variable that is used as a search cost shifter cannot be separately identified from the same variable entering the utility specification as well. An additional problem is that it is difficult to deal with correlation between the unobserved characteristic ξ and search cost shifters. This can become an issue if location decisions are based on the level of unobserved quality, so that distance is potentially correlated with the unobserved characteristic. This may bias estimation of the distance parameter in the search cost specification if not properly instrumented for.

Table 4 gives the parameter estimates for the simplest version of the model, the conditional logit model. We use a simplified version of equation (1)—we only allow for a single price coefficient and do not allow for any random coefficients, so the indirect utility function is given by

$$u_{ij} = \alpha p_j + \mathbf{x_j}'\beta + \xi_j + \varepsilon_{ij}. \tag{27}$$

As car attributes we use a constant, horsepower per weight (HP/weight), a dummy for whether the brand is non-European, a cruise control dummy, fuel efficiency (km/euro), size, and a dummy for whether the car is a luxury brand. We also include car-segment dummies, where car models are classified as small, family, luxury, sports, MPV, or SUV.[31] As a benchmark case, we first present the demand estimates for a model without search frictions. The results in column (A) are obtained by regressing $\log(s_j) - \log(s_0)$ on product characteristics and prices using an instrumental variables (IV) approach to control for possible correlation between unobserved characteristics and price. Following most of the previous literature, we assume the car characteristics to be exogenous. As an instrument for price we use the predicted price from a regression of price on a constant, log(HP/weight), a non-European dummy, cruise control, log(km/liter), log(size), a luxury brand dummy, and segment dummies.

All parameter estimates in column (A) of Table 4 have the expected sign. The results indicate that cars produced by non-European firms yield negative marginal utility, which means cars produced by European firms (e.g., Peugeot/Citroën, Fiat, Volkswagen, etc.) have a higher mean consumer valuation than cars produced by non-European firms (e.g., Toyota, Honda, etc.). Luxury brands generate more utility than non-luxury brands. Size, a higher mileage per euro, and cruise control as standard equipment all affect the consumers' mean utility in a positive way. Finally, the price coefficient is large and very precisely estimated, which results in relatively large average own-price elasticity estimates in absolute value.

In the last three columns of Table 4 we present the demand estimates using our consumer search model. Here we only use distance as a search cost shifter—we relate the search cost parameter λ to distances from the centroid of a neighborhood to the nearest dealers. Even in the simple conditional logit framework, once we include search frictions, there is no longer a closed-form solution for the market share equations, so we proceed by simulating them. Specifically, we randomly draw 2,209 neighborhoods from the demographic data, where each neighborhood is weighted by number of inhabitants.[32] Next we use the distances to the nearest dealer for each of the brands in our sample to simulate search behavior for the 2,209 selected "consumers." We estimate the model by GMM, using the same instrument for price as when estimating specification (A), and use $(Z'Z)^{-1}$ as the weighting matrix, where Z contains the instruments. The results shown in column (B) of Table 4 are obtained using the search cost distribution specified in equation (14), and show that search costs are positively related to distance and significantly different from zero at the one percent level. Most

[31] The classification we use is based on the Euro NCAP Class vehicle classification. The largest class in terms of sales-weighted market share in the period 2003-2008 is the supermini class with a market share of 0.347, followed by the small family car class (0.214), the large family car class (0.176), and the small MPV class (0.148). In our analysis we combine the small and large family car classes into a single family car class (combined sales-weighted market share of 0.390 during 2003-2008), and combine the small and large MPV classes (market share of 0.172), as well as the small and large off-road 4x4 classes into a single SUV class (market share of 0.054). The market shares of cars in the remaining classes are 0.028 for luxury cars and 0.007 for sports cars.

[32] These draws are in fact a certain type of quasi-random draws constructed from a $(0, 2, 47)$-net in base 47, which contains $47^2 = 2{,}209$ draws (see Sándor and András, 2004).

Table 4: Estimation results conditional logit model

	Full information		Search					
	IV Logit Demand (A)		GMM/IV Logit Demand (B)		GMM/IV Logit Demand (C)		GMM/IV Logit Demand (D)	
Variable	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
Preference parameters								
constant	-25.237	(3.169)***	-22.103	(2.918)***	-21.424	(2.902)***	-21.893	(2.912)***
HP/weight	7.639	(1.913)***	6.784	(1.723)***	6.755	(1.718)***	6.775	(1.722)***
non-European	-1.623	(0.282)***	-1.456	(0.256)***	-1.442	(0.255)***	-1.452	(0.256)***
cruise control	0.716	(0.206)***	0.496	(0.192)***	0.493	(0.191)***	0.495	(0.191)***
fuel efficiency	2.566	(0.517)***	2.459	(0.456)***	2.449	(0.454)***	2.456	(0.455)***
size	19.328	(3.950)***	16.546	(3.609)***	16.480	(3.598)***	16.527	(3.606)***
luxury brand	2.915	(0.662)***	2.717	(0.591)***	2.712	(0.589)***	2.715	(0.590)***
price	-0.257	(0.057)***	-0.220	(0.052)***	-0.219	(0.052)***	-0.220	(0.052)***
Search cost parameters								
distance	—		0.142	(0.045)***	0.087	(0.025)***	0.152	(0.046)***
Avg. own-price elast.	-7.480		-6.396		-6.374		-6.390	
Search cost distr.	–		Eq.(14)		Normal		Gumbel (min)	

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. The number of observations is 1,382. Standard errors are in parenthesis. The number of simulated consumers used for the estimation of specifications (B)-(D) is 2,209. Instrument for price is obtained using the fitted price from a regression of price on a constant, log(HP/weight), non-European, cruise control, log(km/liter), log(size), luxury brand, and segment dummies. We include average distance to dealerships as an instrument. The Cragg-Donald Wald F statistic for specification (A) is 24.95 with p-value 0.00.

utility parameters decrease in magnitude in comparison to the full information estimates. This is the case for the price parameter as well, which results in a lower average absolute own-price elasticity in the search model than in the full information model.

An advantage of the search cost distribution we use in specification (B) is that it gives closed-form expressions for the buying probabilities. However, the model can be estimated using different search cost distributions by using numerical integration to obtain the probability that a consumer buys from firm f, as in equation (12). In specification (C) we use this approach and use a normal distribution for search costs. Although the effect of distance in the search cost specification is smaller when using a normal distribution, the estimates for the utility parameters are very similar and do not seem to depend much on the difference in parametric specification.[33] As shown in specification (D), the estimated utility parameters are also very similar when using a Gumbel (minimum) distribution. Note that a Gumbel (minimum) distribution is very similar in shape to the search cost distribution specified in equation (14) (see also Figure 1), which results in an estimated search cost parameter that is of similar magnitude. So regardless of whether we use a normal search cost distribution, a Gumbel (minimum) distribution, or the one based on equation (14), a comparison of the estimation results with search to those without search shows that the price coefficient goes down in

[33]The difference arises because the variance of search costs is different across specifications. The standard deviation of the normal distribution of search costs we use in specification (C) is normalized to one—when the standard deviation is normalized to 1.6 instead, the estimate for distance goes up to 0.142.

absolute value, which suggests that at least in the conditional logit model, ignoring search frictions may result in an overestimation of consumer price sensitivity. We will come back to this when discussing the results for the complete model below.

Complete model

To estimate the complete model we supplement the aggregate data that was used to estimate the conditional logit model with search and purchase data from the individual-level survey data. The demand side estimates for the complete model are based on the utility function in equation (1) and the search cost distribution that is derived from Gumbel distributed w's, as specified in equation (14). We use the same car attributes as those shown in Table 4 for the estimation of the simplified model. As before, we estimate the mean marginal utility of each of these attributes, and also allow the marginal utility for some of the attributes to differ across consumers by estimating a variance term for these attributes. Specifically, for the constant, size, and luxury brand dummy we use a standard normal draw for the corresponding component of the diagonal matrix V_i. For family car and MPV the corresponding components of the diagonal matrix are kids $\times\ \upsilon_{3i}$ and log(family size) $\times\ \upsilon_{3i}$, where kids is a dummy for whether there are children present in the household and υ_{3i} is a $\chi^2(3)$-distributed draw truncated at 95%. An advantage of using this distribution is that it is bounded and skewed toward positive taste (see also Petrin, 2002).

We allow for heterogeneity in the price parameter in accordance with equation (2). This means that in addition to normalizing prices by household income, we allow α to differ according to income groups (as in Petrin, 2002). A simulated consumer's household income, which is used for obtaining the mean utilities, is randomly drawn from a log-normal distribution with scale parameter 0.28 (which is estimated outside the model) and neighborhood-specific location parameter such that the mean (after-tax) household income level in the neighborhood where the simulated consumer resides matches the neighborhood data from Statistics Netherlands. The income bound $\overline{y}$ we use corresponds to a household income (after tax) of €25,000.[34] The senior dummy is obtained from the neighborhood-specific percentage of households with a head of household older than 65 years, i.e., the senior dummy equals 1 if that percentage is larger than a uniform draw on $(0, 1)$ and zero otherwise. Similarly, the kids dummy is obtained from the neighborhood-specific percentage of households with kids, i.e., the kids dummy equals 1 if that percentage is larger than a uniform draw on $(0, 1)$ and zero otherwise.

We let search costs interact with a rich set of demographics. Specifically, in addition to distance we let search costs depend on the logarithm of household income as well as a dummies for whether the head of household is senior and whether there are kids living in the household. We also interact these variables with distance. In addition, we estimate the mean and standard deviation of a normal distributed random search cost constant.

The estimation results for the search model are presented in the first column of Table 5.[35] The nonlinear

[34]For the choice of income bound we are constrained by the income bins used in the survey. The chosen bound approximately equals a household income of €30,200 before taxes, which corresponds to one of the cutoffs used to create bins in the survey data.

[35]Although not reported, we have estimated versions of the model with more random coefficients. Even though most of the

Table 5: Estimation results complete model

	Search (A)		Full Information (B)	
Variable	Coeff.	Std. Err.	Coeff.	Std. Err.
Price coefficients (price/income)				
income less than 25k	-3.856	(0.271)***	-2.548	(0.082)***
income more than 25k	-5.253	(0.526)***	-3.117	(0.153)***
Base coefficients				
constant	-10.951	(2.080)***	-19.715	(2.406)***
HP/weight	3.990	(1.263)***	5.252	(1.429)***
non-European	-0.918	(0.187)***	-1.227	(0.210)***
cruise control	0.143	(0.134)	0.432	(0.154)***
fuel efficiency	2.089	(0.343)***	2.311	(0.384)***
size	10.024	(2.594)***	13.437	(3.056)***
luxury brand	1.582	(0.449)***	2.012	(0.495)***
price	-0.004	(0.041)	-0.102	(0.043)***
Random coefficients				
constant	0.553	(0.716)	0.255	(0.526)
size	1.755	(0.655)***	1.039	(0.513)**
luxury brand	1.328	(0.031)***	0.800	(0.017)***
family car × kids	0.000	(0.013)	0.150	(0.004)***
MPV/SUV × log(family size)	0.049	(0.011)***	0.158	(0.005)***
distance	-0.580	(0.040)***	-0.156	(0.012)***
Search cost parameters				
distance	0.821	(0.181)***	—	
log(income)	-0.806	(0.264)***	—	
× distance	-0.178	(0.053)***	—	
kids	-2.393	(0.153)***	—	
× distance	0.538	(0.069)***	—	
senior	-0.135	(0.224)	—	
× distance	-0.188	(0.040)***	—	
constant (mean)	11.883	(0.957)***	—	
constant (st.dev.)	4.552	(0.093)***	—	
Objective function	0.045		0.048	
Objective function (non-search)	0.037		0.048	
Average own-price elasticity	-3.940		-5.321	

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. The number of observations is 1,382. The number of random draws used for the aggregate data is 2,209. The number of random draws used for the micro data is 25 per respondent. Standard errors are in parenthesis. The linear part of utility includes car segment fixed effects. Instrument for price is obtained using the fitted price from a regression of price on a constant, log(HP/weight), non-European, cruise control, log(km/liter), log(size), luxury brand, and segment dummies.

parameters, which include the nonlinear price coefficients, the random coefficients, and the search cost parameters, are obtained using the GMM approach outlined in Section 3. The base coefficients are obtained in a second step by regressing mean utility on car attributes. We allow the linear price coefficient to be correlated with ξ_j, and use two-stage least squares to estimate the linear parameters of the model. As

estimated base coefficients change somewhat as a result of allowing for more random coefficients, the price and search cost parameter estimates appear robust to changes in the number of random coefficients. We prefer a specification with less random coefficients since this increases the precision of the estimates. Moreover, with a full set of random coefficients it is very difficult to do counterfactual exercises due to numerical issues when solving for the price equilibrium (see also Skrainka, 2012).

before, we use the fitted price from a regression of price on cost shifters as an instrument for price.

Both nonlinear price coefficients are statistically significant at high significance levels. Although we allow prices to enter consumers' mean utility, the estimated linear price coefficient is close to zero and not significant. All else equal, the positive HP/weight parameters suggests consumers prefer more powerful cars. The estimated coefficient of the non-European dummy is negative and highly significant, which indicates consumers on average prefer cars produced by European firms instead of cars produced by non-European firms. As expected, consumers put a positive value on cruise control being a standard option. The mean parameter estimate for fuel efficiency is positive and precisely estimated. The large estimate of the standard deviation parameter for size indicates that consumers differ in their preference for large cars, although the positive estimate for the corresponding base parameter suggests consumers on average prefer larger cars. The mean parameter estimate for the luxury brand dummy is positive and precisely estimated, although the relatively large estimate of the corresponding standard deviation parameter indicates that there is substantial heterogeneity in consumers' marginal valuation for luxury brands. The estimate corresponding to the interaction parameter of MPV/SUV and log(family size) indicates that larger households put positive value on larger cars such as MPVs and SUVs. However, households with kids do not put additional value on family cars in comparison to other demographic groups in our data.

In addition to serving as a search cost shifter, the distance from a consumer to the nearest dealer of a brand could also be part of the utility function. One way in which it could enter utility is because of service: if consumers prefer to have their cars serviced at the dealer, a car's indirect utility is likely to be directly affected by distance to the closest dealership. To control for this, we have also added distance from the consumer to the nearest dealer to the utility function. According to the estimates, distance has a negative effect on utility, which may capture the negative effect on utility of service visits to dealers located farther away from the consumer.

Most search cost parameter estimates are highly significant. The estimates indicate that search costs are positively related to distance, but that the effect is smaller for senior households. Both income and the kids dummy are negatively related to search costs, although the kids dummy has a positive effect when interacted with distance. Note that the average distance to a dealer is slightly over 7 kilometers, which means that the combined marginal effect of kids on search costs is positive for most people, but negative for income and senior. This is consistent with the reduced form results discussed in the previous section. The estimates for the distribution of the random search cost constant indicate that there is a lot of variation in constant search costs. Moreover, the estimated mean is relatively high, which is consistent with a high proportion of consumers not searching in any given year.

The last column of Table 5 gives parameter estimates in case we assume consumers have full information, as in BLP. The results in column (B) of this table are obtained using the same preference parameters as in specification (A) of the table, although we can no longer use the search related moments. The estimated price coefficients point to a significantly higher marginal disutility of price in comparison to the estimates for the search model.

Table 6: Fit moments

	Survey	Search (A)	Full info (B)
Purchase related probabilities			
$E[\mathbb{1}\{i \text{ purchases an MPV or SUV}\} \mid \{i \text{ has a family size } \leq 3\}]$	0.012	0.012	0.011
$E[\mathbb{1}\{i \text{ purchases an MPV or SUV}\} \mid \{i \text{ has a family size } > 3\}]$	0.021	0.020	0.020
$E[\mathbb{1}\{i \text{ purchases a family car}\} \mid \{i \text{ has children in the household}\}]$	0.036	0.034	0.034
$E[\mathbb{1}\{i \text{ purchases a family car}\} \mid \{i \text{ has no children in the household}\}]$	0.026	0.021	0.020
$E[\mathbb{1}\{i \text{ purchases a luxury brand}\} \mid \{y_i < \overline{y}\}]$	0.003	0.001	0.002
$E[\mathbb{1}\{i \text{ purchases a luxury brand}\} \mid \{y_i \geq \overline{y}\}]$	0.006	0.008	0.008
$E[\mathbb{1}\{i \text{ purchases a new vehicle}\} \mid \{y_i < \overline{y}\}]$	0.038	0.040	0.040
$E[\mathbb{1}\{i \text{ purchases a new vehicle}\} \mid \{y_i \geq \overline{y}\}]$	0.085	0.085	0.086
$E[\mathbb{1}\{i \text{ purchases a new vehicle}\} \mid \{\# \text{ of dealers within 10km of } i < 15\}]$	0.034	0.034	0.034
$E[\mathbb{1}\{i \text{ purchases a new vehicle}\} \mid \{\# \text{ of dealers within 10km of } i \geq 15\}]$	0.069	0.069	0.070
Search related probabilities			
$E[\mathbb{1}\{i \text{ searches once}\} \mid \{y_i < \overline{y}\}]$	0.041	0.039	
$E[\mathbb{1}\{i \text{ searches once}\} \mid \{y_i \geq \overline{y}\}]$	0.057	0.058	
$E[\mathbb{1}\{i \text{ searches once}\} \mid \{\# \text{ of dealers within 10km of } i < 15\}]$	0.034	0.036	
$E[\mathbb{1}\{i \text{ searches once}\} \mid \{\# \text{ of dealers within 10km of } i \geq 15\}]$	0.052	0.051	
$E[\mathbb{1}\{i \text{ searches once}\} \mid \{i \text{ is senior}\}]$	0.048	0.049	
$E[\mathbb{1}\{i \text{ searches once}\} \mid \{i \text{ is not a senior}\}]$	0.050	0.050	
$E[\mathbb{1}\{i \text{ searches once}\} \mid \{i \text{ has children in the household}\}]$	0.062	0.062	
$E[\mathbb{1}\{i \text{ searches once}\} \mid \{i \text{ has no children in the household}\}]$	0.045	0.045	
$E[\mathbb{1}\{i \text{ searches at least twice}\} \mid \{y_i < \overline{y}\}]$	0.055	0.056	
$E[\mathbb{1}\{i \text{ searches at least twice}\} \mid \{y_i \geq \overline{y}\}]$	0.089	0.089	
$E[\mathbb{1}\{i \text{ searches at least twice}\} \mid \{\# \text{ of dealers within 10km of } i < 15\}]$	0.042	0.042	
$E[\mathbb{1}\{i \text{ searches at least twice}\} \mid \{\# \text{ of dealers within 10km of } i \geq 15\}]$	0.079	0.079	
$E[\mathbb{1}\{i \text{ searches at least twice}\} \mid \{i \text{ is senior}\}]$	0.087	0.087	
$E[\mathbb{1}\{i \text{ searches at least twice}\} \mid \{i \text{ is not a senior}\}]$	0.072	0.072	
$E[\mathbb{1}\{i \text{ searches at least twice}\} \mid \{i \text{ has children in the household}\}]$	0.084	0.084	
$E[\mathbb{1}\{i \text{ searches at least twice}\} \mid \{i \text{ has no children in the household}\}]$	0.070	0.070	

Table 6 gives the estimated probabilities used for the moments as well as those from the survey data. Both specifications are able to match the probabilities from the survey data relatively well—most estimated probabilities used as moments are well within 0.2 percent point of the corresponding probabilities from the survey data.

4.4 Demand elasticities and markups

Panel A of Table 7 gives demand elasticity estimates for a selection of car models sold in 2008 for both the search model (using the estimates in column (A) of Table 5) and the full information model (using the estimates in column (B) of the table). For all models, demand is estimated to be more inelastic in the search model than in the full information model. This means that assuming consumers have full information, while in reality they do not, will lead to an overestimation of price sensitivity for most car models. This is consistent with theory: as shown by Haan, Moraga-González, and Petrikaitė (2018) and Choi, Dai, and Kim (2018), the sequential search model leads to prices that are lower than in the full information model. If the true data generating process for prices is governed by the search model, fitting a full information model will overestimate the absolute value of the own-price elasticities to be able to accommodate observed prices.

Table 7: Demand elasticity estimates

	Toyota Auris	Renault Megane	Opel Zafira	VW Passat	Ford Mondeo	Audi A4	BMW 3-series	Mercedes C-class
PANEL A: PERCENTAGE CHANGE IN SALES								
Search								
Toyota Auris	-2.8099	0.0149	0.0195	0.0306	0.0379	0.0271	0.0261	0.0194
Renault Megane	0.0123	-2.9395	0.0176	0.0295	0.0387	0.0272	0.0228	0.0140
Opel Zafira	0.0101	0.0110	-3.4951	0.0363	0.0610	0.0320	0.0268	0.0154
Volkswagen Passat	0.0099	0.0115	0.0228	-3.6007	0.0500	0.0755	0.0307	0.0188
Ford Mondeo	0.0100	0.0124	0.0313	0.0409	-3.6450	0.0401	0.0257	0.0161
Audi A4	0.0082	0.0100	0.0189	0.0710	0.0461	-4.2754	0.0506	0.0714
BMW 3-series	0.0089	0.0094	0.0177	0.0323	0.0331	0.0566	-4.2084	0.0565
Mercedes C-class	0.0087	0.0076	0.0134	0.0260	0.0272	0.1051	0.0743	-4.0651
Full information								
Toyota Auris	-3.3399	0.0042	0.0066	0.0113	0.0143	0.0121	0.0109	0.0079
Renault Megane	0.0035	-3.6696	0.0069	0.0121	0.0151	0.0130	0.0109	0.0082
Opel Zafira	0.0034	0.0043	-4.4039	0.0122	0.0153	0.0134	0.0115	0.0084
Volkswagen Passat	0.0037	0.0047	0.0076	-4.6476	0.0170	0.0156	0.0135	0.0096
Ford Mondeo	0.0038	0.0048	0.0078	0.0139	-4.8321	0.0150	0.0132	0.0095
Audi A4	0.0037	0.0048	0.0079	0.0147	0.0173	-5.7093	0.0292	0.0190
BMW 3-series	0.0037	0.0045	0.0076	0.0142	0.0170	0.0327	-5.7425	0.0234
Mercedes C-class	0.0035	0.0044	0.0073	0.0132	0.0160	0.0280	0.0307	-6.0945
PANEL B: ABSOLUTE CHANGE IN SALES								
Search								
Toyota Auris	-110.4835	0.5857	0.7673	1.2046	1.4922	1.0644	1.0263	0.7647
Renault Megane	0.5294	-126.6935	0.7576	1.2694	1.6692	1.1731	0.9822	0.6052
Opel Zafira	0.5670	0.6193	-197.1573	2.0484	3.4431	1.8038	1.5119	0.8708
Volkswagen Passat	0.8416	0.9810	1.9366	-305.9133	4.2505	6.4105	2.6063	1.5956
Ford Mondeo	0.9970	1.2337	3.1132	4.0650	-362.3460	3.9855	2.5550	1.5975
Audi A4	0.5905	0.7199	1.3543	5.0908	3.3094	-306.7607	3.6286	5.1197
BMW 3-series	0.5603	0.5932	1.1171	2.0368	2.0878	3.5708	-265.2968	3.5602
Mercedes C-class	0.3953	0.3461	0.6092	1.1807	1.2360	4.7703	3.3710	-184.5576
Full information								
Toyota Auris	-131.3253	0.1656	0.2579	0.4463	0.5604	0.4752	0.4274	0.3113
Renault Megane	0.1497	-158.1576	0.2967	0.5200	0.6505	0.5624	0.4713	0.3527
Opel Zafira	0.1906	0.2425	-248.4212	0.6868	0.8630	0.7573	0.6488	0.4766
Volkswagen Passat	0.3118	0.4019	0.6493	-394.8622	1.4480	1.3246	1.1445	0.8126
Ford Mondeo	0.3744	0.4808	0.7803	1.3848	-480.3601	1.4939	1.3117	0.9417
Audi A4	0.2637	0.3452	0.5686	1.0519	1.2405	-409.6441	2.0953	1.3654
BMW 3-series	0.2334	0.2847	0.4794	0.8944	1.0718	2.0619	-362.0064	1.4734
Mercedes C-class	0.1609	0.2017	0.3334	0.6013	0.7286	1.2722	1.3951	-276.6906

Notes: Demand elasticities are calculated for 2008. Panel A gives the percentage change in market share of model i with a one percent change in the price of model j, where i indexes rows and j columns. Panel B gives the change in sales of model i with a one percent change in the price of model j. Elasticities for the search model are calculated using estimates from specification (A) in Table 5; those for the full information model are based on specification (B) in Table 5.

Table 8: Markups and variable profits

	price	pre-tax price	sales	Search markup over MC	Search percentage markup	Search variable profit	Full information markup over MC	Full information percentage markup	Full information variable profit
Toyota Auris	16,857	11,345	3,932	4,207	37.08	16.54	3,165	27.89	12.44
Renault Megane	18,647	12,048	4,310	4,772	39.61	20.57	3,191	26.48	13.75
Opel Zafira	22,812	15,037	5,641	4,571	30.40	25.78	3,251	21.62	18.34
Volkswagen Passat	24,129	15,447	8,496	4,871	31.54	41.39	3,296	21.34	28.00
Ford Mondeo	25,230	16,536	9,941	4,922	29.76	48.93	3,300	19.96	32.81
Audi A4	30,383	19,324	7,175	5,352	27.70	38.40	3,391	17.55	24.33
BMW 3-series	30,875	20,481	6,304	4,822	23.54	30.40	3,309	16.16	20.86
Mercedes C-class	32,609	20,878	4,540	5,427	25.99	24.64	3,337	15.98	15.15

Notes: Prices and markup over MC are in Euros. Variable profit is in Euro×1 mln. Markups and profits for the search model are calculated using estimates from specification (A) in Table 5; those for the full information model are based on specification (B) in Table 5. Percentage markup is calculated as $(p_j^* - mc_j)/p_j^*$, where p_j^* is the pre-tax price of car j. Variable profit is calculated as $q_j \cdot (p_j^* - mc_j)$, where q_j is the sales of car j.

The cross-price elasticities show the opposite pattern: the percentage change in market share as a result of a percent increase in price of a rival model is in the majority of the cases larger in the search model than in the full information model. In the search model, consumers with a relatively high buying probability for a specific model are likely to have a relatively low search cost constant, which results in higher purchase probabilities for other car models as well, explaining the higher cross-price elasticities. Moreover, in our search model consumers are relatively more likely to switch to a brand that is located nearby than in the full information model. In order to see this in our results, we computed the change in sales due to a one percent increase in price of a rival car model; see panel B of Table 7. A comparison of the three German luxury cars reveals that, irrespective of their market shares, substitution is more likely towards the brand with more dealerships. For example, a one percent price increase for Audi A4 implies a sales increase of 3.57 units for BMW 3-series (which has a market share of 1.3 percent of total sales and 57 dealerships) and of 4.77 units for the Mercedes C-class (market share of 0.9 percent and 83 dealerships). In the full information model the effect of distance on utility is not sufficiently large to yield a similar phenomenon: a price increase of Audi A4 implies a sales increase of 2.06 units for the BMW 3-series and of 1.27 units for the Mercedes C-class, which seems to be a consequence of the relative magnitude of the market shares. In conclusion, in the search model substitution patterns, in addition to car characteristics, also appear to be influenced by search costs and proximity of car dealerships; this phenomenon cannot be captured by the full information model.

Table 8 compares the estimated markups between the search model and the full information model. Consistent with the elasticity patterns reported in Table 7, estimated markups in the search model are higher for all the cars. The estimated average percentage markup across all models in 2008 is 34 percent for the search model versus 24 percent for the full information model. These numbers are not unrealistic: BLP report an average ratio of markup to retail price of 24 percent for their main specification and they note that it is "not extraordinarily high" (BLP, p. 883), whereas Petrin (2002) finds an average markup of 17 percent for the model with moments. Goldberg (1995), on the other hand, obtains higher markup estimates:

Table 9: Estimation results alternative search cost distributions

Variable	(A) Coeff.	(A) Std. Err.	(B) Coeff.	(B) Std. Err.
Price coefficients (price/income)				
income less than 25k	-2.806	(0.118)***	-3.655	(0.212)***
income more than 25k	-3.777	(0.247)***	-5.156	(0.337)***
Base coefficients				
constant	-11.518	(1.967)***	-11.397	(2.064)***
HP/weight	3.643	(1.197)***	3.864	(1.229)***
non-European	-0.736	(0.175)***	-0.852	(0.180)***
cruise control	0.186	(0.127)	0.162	(0.133)
fuel efficiency	1.915	(0.317)***	1.992	(0.328)***
size	9.915	(2.438)***	10.356	(2.533)***
luxury brand	1.566	(0.414)***	1.547	(0.429)***
price	-0.035	(0.036)	-0.006	(0.038)
Random coefficients				
constant	0.054	(0.321)	0.549	(0.261)**
size	0.383	(0.329)	0.660	(0.445)
luxury brand	0.845	(0.012)***	1.177	(0.040)***
family car × kids	0.043	(0.006)***	0.000	(0.006)
MPV/SUV × log(family size)	0.086	(0.004)***	0.054	(0.007)***
distance	-0.212	(0.027)***	-0.412	(0.055)***
Search cost parameters				
distance	0.568	(0.070)***	0.843	(0.117)***
log(income)	0.044	(0.083)	-0.698	(0.166)***
× distance	-0.149	(0.019)***	-0.199	(0.037)***
kids	-0.972	(0.041)***	-3.224	(0.226)***
× distance	0.157	(0.013)***	0.874	(0.098)***
senior	-0.421	(0.042)***	-0.512	(0.193)**
× distance	-0.021	(0.008)***	-0.109	(0.038)***
constant (mean)	4.642	(0.255)***	11.741	(0.500)***
constant (st.dev.)	2.203	(0.042)***	4.546	(0.044)***
Objective function	0.056		0.049	
Objective function (non-search)	0.046		0.039	
Average own-price elasticity	-3.767		-3.875	
Search cost distribution	Normal		Gumbel (min)	

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. The number of observations is 1,382. The number of random draws used for the aggregate data is 2,209. The number of random draws used for the micro data is 25 per respondent. Standard errors are in parenthesis. The linear part of utility includes car segment fixed effects. Instrument for price is obtained using the fitted price from a regression of price on a constant, log(HP/weight), non-European, cruise control, log(km/liter), log(size), luxury brand, and segment dummies.

wholesale price markups of on average 38 percent, which implies even larger retail price markups. In both the search model and the full information model, the Ford Mondeo is the most profitable model among the ones listed in Table 8—the most profitable car model among all cars in the sample is the Peugeot 207.

4.5 Robustness

Table 9 shows the parameter estimates using alternative specifications for the search cost distribution. In specification (A) we re-estimate the main specification but use a normal search cost distribution instead of the one given by equation (14), which is based on a Gumbel distribution for w.[36] Most of the estimated utility parameters are relatively similar to those for the main specification reported in column (A) of Table 5. However, even though the price coefficient in the linear part of utility is more negative, the price coefficients that depend on income are less negative, and as a result average demand is slightly more inelastic in comparison to the main specification. Another difference between the results in column (A) and those for the main specification is that the parameters of the search cost distribution are uniformly lower than those for the main specification, which can be explained by the larger variance of the search cost distribution given by equation (14) (see also footnote 33). Column (B) gives estimates when using a Gumbel (minimum) search cost distribution—both utility and search cost parameters are similar to those for the main specification.

4.6 Counterfactuals

In this section we study the effects of three changes in the primitives of the model, using the search model estimates reported in column (A) of Table 5. For all counterfactual simulations we use the supply side model that is discussed in Section 2.6. First, we look at what happens to equilibrium prices when the costs of visiting dealers change. Some dealers have recently started to bring cars to buyers for test driving at home or work.[37] We model this situation as if the cost of transportation to the dealership went down all the way to zero. Secondly, we make a comparison between prices in the search model and prices in the full information model by using the estimates for the search model to simulate a full information equilibrium. Finally, multi-brand firms sometimes choose to retail their products using different dealership networks. A well-known case of this in the automobile industry is the 1998 merger between Daimler-Benz and Chrysler whose retail networks largely remained separate. Business analysts have suggested that this likely hindered Chrysler's market penetration in Europe (see also Finkelstein, 2002). We explore what happens to prices and profits when the Renault-Nissan Alliance, one of the multi-brand firms that does not yet sell all its cars in all its dealerships, starts doing so for some of its brands.

Change in search costs

To see how prices change if search costs for all consumers and dealers decrease simultaneously, we take the estimates reported in column (A) of Table 5 and simulate equilibrium prices and market shares when setting the distance-related part of each consumer's search cost equal to a specific percentage of the estimated distance-related search cost. The average decrease in price when moving to zero distance-related search costs is 1 percent. We find that the closer a dealer is (on average) to consumers, the more prices decrease as a result of a reduction in distance-related search costs. For instance, average price reductions exceed 2.5 percent for some of the brands with many dealership locations, such as Peugeot and Renault. On the

[36]To estimate the model using a normal search cost distribution, P_{if} in equation (12) needs to be integrated numerically, which is time consuming: one function evaluation takes about 200 seconds on a Late 2015 27 inch iMac with 4 GHz Intel Core i7 processor, whereas only 10 seconds when using the search cost distribution given by equation (14).

[37]See http://www.edmunds.com/car-news/phil-long-dealership-group-sells-and-services-cars-and-will-travel.html as well as footnote 3 in the introduction.

Table 10: Simulated prices for different levels distance-related search costs

	percentage of distance-related search costs			full info
	100%	50%	0%	($\lambda = -\infty$)
Price				
Toyota Auris	16,857	16,830	16,808	16,816
Renault Megane	18,647	18,428	18,166	17,852
Opel Zafira	22,812	22,718	22,582	22,565
Volkswagen Passat	24,129	23,957	23,839	23,911
Ford Mondeo	25,230	25,099	24,992	24,976
Audi A4	30,383	30,116	29,946	29,851
BMW 3-series	30,875	30,756	30,628	30,668
Mercedes C-class	32,609	32,401	32,147	30,963
Average	29,214	29,097	28,981	28,873
Share not searching	0.8779	0.8472	0.7957	—
Average percentage markup	33.54	33.25	32.97	32.62

Notes: Prices are in Euros. The prices shown are simulated prices when search costs are x percent of the distance-related part of the estimated search costs from specification (A) in Table 5, where x is between 0 and 100 percent. The last column gives simulated prices for the full information model.

other hand, brands with relatively few dealership locations, such as Jeep, Land Rover, and Smart see much smaller price decreases, or even price increases. This could be explained by the fact that brands that have very few dealers tend to have consumers who have very strong brand preferences so a decrease in search costs has less impact on these brands. Moreover, we observe that the price effect is largest for relatively small and inexpensive models such as the Peugeot 107, Chevrolet Matiz, and Renault Twingo, while we see relatively low price decreases or even price increases for large and expensive car models such as the Land Rover Defender and BMW 7-series. Table 10 summarizes the effects on prices for the same car models as in Table 8.

In Table 10 we also report prices when using the estimates to simulate the pricing equilibrium under the full information model.[38] A comparison of the average price under full information with actual prices suggests that prices are on average €341 higher because of search frictions.[39] As shown in the last column of the table, for several models the simulated prices under full information exceed the simulated prices for the search model when the distance-related part of search costs is set to zero percent. The non-monotonic relationship we find between prices and search costs when moving from the search model to the full information model deserves an explanation. As noted before, in our model an increase in search costs has three different effects. First, as can be seen in Table 10, as search costs increase the share of consumers not searching goes up. If

[38]The full information model can be obtained by setting search costs to zero in the search model. In the Gumbel distributed w case this means that μ_{if} has to be equal to zero, which requires that $\lambda \to -\infty$. This implies that the simulated prices under the full information model are not necessarily similar to those when all search costs *parameters* are set to zero.

[39]In a related study, Murry and Zhou (2018) use individual-level transaction data for new cars to quantify how geographical concentration among car dealers affects competition and search behavior and find that search frictions lead to an average price increase of $333.

the consumers who remain in the market are the more elastic ones, as demonstrated in Moraga-González, Sándor, and Wildenbeest (2017), firms get an incentive to lower their prices. Second, as it is standard in search models, higher search costs give firms enhanced market power over the consumers who visit, and thereby firms have an incentive to raise their prices. Finally, because we have a model in which prices are observed before search, higher search costs increase competition for visits, and this puts downward pressure on prices (see also footnote 19). The latter effect is because consumers in our model determine whether to continue searching or not by making a tradeoff between the expected utility and the search costs of visiting an additional dealer. Since the former depends on the observed prices of the cars, a lower price for a specific car model makes it more likely that a consumer visits the dealer selling that car, which leads to a negative relation between search costs and prices. Which of the three effects will dominate depends on the level of search costs as well as the level of competition. Both of these are determined at the car model level, which explains why we find that simulated prices can be increasing for some car models when search costs go down, while decreasing for other cars.

Selling different makes at same dealership

Most of the 16 firms in our data own several brands. For instance, the Volkswagen Group owns Audi, Seat, Skoda, and Volkswagen, and the Renault-Nissan Alliance owns Renault, Nissan, and Dacia. Whereas firms typically sell their brands in separate dealerships, in a setting where search frictions are important it might be beneficial to sell multiple brands at the same location.[40] Although in the Netherlands most manufacturers sell their brand at separate, single-brand dealerships, some manufacturers let some (or all) of their brands be sold under one roof. For instance, Audi and Volkswagen are typically sold at the same dealership, just as BMW and Mini (both part of the BMW Group).

To study the effects of retailing different brands within a single dealership on prices, market shares, and profits, we let all Renault dealerships sell Dacia as well, while at the same time we eliminate the locations of the Dacia dealerships. In terms of the model this means that by visiting this combined Renault-Dacia dealer, consumers observe relevant characteristics for all Renault and Dacia models. Since Renault has the denser dealership network of the two brands, the number of dealers selling Dacia cars doubles from 98 to 196, which means that average search costs for Dacia go down considerably.

Table 11 reports the simulated sales, market shares, prices, and variable profits of the major brands in our dataset after prices have been re-optimized following the change in Renault's dealership network. Although most manufacturers see a modest decline in sales following the change, sales for Dacia cars go up by 3,156 units, which is an increase of approximately 69 percent, despite an average price increase of over six percent. Variable profits tell a similar story: they slightly decline for most manufacturers, while the Renault-Nissan Alliance profits go up by roughly 10.6 million. Notice that while the dealership merger greatly benefits the sales and variable profits of the Dacia brand, they hurt the sales and profits of the Renault brand. After the

[40] Moraga-González and Petrikaitė (2013) show that a firm that puts on display all its products unfolds the economies of search associated to one-stop shopping, which makes the firm more attractive for consumers and tends to increase profitability. However, a firm that stocks more products together increases competition with the rival firms and this tends to lower profits. Which of these effects dominates depends on the magnitude of search costs.

Table 11: Sales, market shares, and prices for change in Renault-Nissan network

	sales (units)		market shares (%)		prices (€)		profits (mln €)	
Group	before	after	before	after	before	after	before	after
BMW	18,587	18,543	3.80	3.78	50,527	50,527	90.67	90.47
Daimler-Chrysler	14,771	14,748	3.02	3.01	42,581	42,581	75.25	75.14
Fiat	24,821	24,749	5.07	5.04	18,594	18,595	95.00	94.71
Ford	68,746	68,565	14.04	13.97	38,692	38,689	322.42	321.53
Fuji	1,422	1,417	0.29	0.29	23,560	23,561	5.65	5.63
General Motors	49,962	49,848	10.20	10.16	24,096	24,097	217.67	217.22
Honda	8,479	8,464	1.73	1.72	23,017	23,018	34.57	34.51
Hyundai	17,433	17,360	3.56	3.54	16,521	16,519	68.68	68.37
Kia	12,236	12,219	2.50	2.49	19,100	19,101	49.43	49.37
Mitsubishi	7,805	7,768	1.59	1.58	27,808	27,808	29.35	29.21
Porsche	531	529	0.11	0.11	76,783	76,781	3.43	3.41
PSA Peugeot Citroen	64,389	64,200	13.15	13.08	21,520	21,521	288.44	287.65
Renault-Nissan	52,334	54,565	10.69	11.12	21,088	21,132	232.64	243.21
Dacia	4,549	7,705	0.93	1.57	7,960	8,452	17.71	32.37
Nissan	10,259	10,205	2.10	2.08	25,497	25,515	43.08	42.97
Renault	37,526	36,654	7.66	7.47	21,068	21,038	171.86	167.87
Suzuki	14,547	14,518	2.97	2.96	12,477	12,475	55.37	55.23
Toyota	49,227	49,100	10.05	10.01	30,018	30,014	204.23	203.57
Volkswagen	84,294	84,120	17.22	17.14	27,945	27,945	397.50	396.70
Total	489,584	490,713	100.00	100.00	29,214	29,217	2,170.28	2,175.93

Notes: Profits exclude fixed costs. Results are obtained using the estimates from specification (A) in Table 5.

dealership merger, Renault faces more competition from Dacia, and some consumers who otherwise would have bought a Renault car, now buy a Dacia car. As a result, Dacia's variable profits increase by close to 15 million, whereas Renault's profits actually decrease by approximately 4 million.

Notice that in this calculation we have ignored changes in the fixed costs associated to this business reorganization, as well as any possible changes in brand preferences for Renault or Dacia following the changes. Dacia is considered to have a lower brand value than Renault, so it could happen that consumers' perception of the value of a Renault car drops when it is sold in the same dealership as a Dacia car. To analyze this in more detail, we calculate how much mean utility for a Renault car has to drop such that a dealership merger would no longer be profitable: if Renault's mean utility goes down by more than 0.082, variable profits for the Renault-Nissan Alliance will go down following the change in dealership network. Using the average estimated price coefficient, this corresponds to a decrease in value of roughly €496.

5 Conclusions

In many markets consumers have imperfect information about the utility they get from the various alternatives available and have to engage in costly search to find out which products they prefer most. While the theoretical consumer search literature is well established, much less work exists trying to estimate demand for environments in which consumer search is important. This paper has contributed to the literature by presenting and estimating a discrete choice model of demand with optimal sequential consumer search. While doing so, we have allowed for unobserved product heterogeneity as in BLP, which sets our paper apart from recent contributions on the theme.

In our model consumers are initially unaware of whether a given product is a good match or not. Consumers search sequentially for differentiated products, taking into account their preferences for the various alternatives as well as the costs of searching them. The optimal strategy is to rank alternatives according to reservation utilities, visit the alternatives starting from the alternative with the highest reservation utility, and to stop searching when the highest observed utility exceeds the reservation utility of the next best alternative. To solve the model, we use recent findings from consumer search theory that re-characterize the search problem as a standard discrete choice problem (Armstrong, 2017; Choi, Dai, and Kim, 2018). Although the model can be estimated for arbitrary search cost distributions, we have shown that for specific assumptions on the search cost distribution we can obtain a closed-form expression for a consumer's choice probability, which dramatically speeds up the estimation and has no major impact on the estimates of demand, elasticities, and markups. We have provided various approaches to estimate the model and have applied them to the Dutch market for automobiles. We use distances from consumers to the nearest dealer of a specific brand as well as household characteristics reflecting the opportunity cost of time to specify consumer search costs. Even though, conditional on having appropriate instruments, versions of the model can be estimated using only aggregate data such as market shares, product characteristics, and consumer and dealer locations, we have supplemented the data with a survey on actual dealer visits for a large number of respondents, which allows us to add more search cost covariates, strengthen the identification, and improve the precision of the estimates.

The survey reveals that consumers conduct a rather limited amount of search before buying. Moreover, a great deal of the searches involves test-driving a car. Our estimation results have shown that search costs are both significant and economically meaningful. Assuming, instead, that search frictions are negligible and consumers have full information results in higher (absolute) own-price elasticity estimates, as well as lower estimated percentage markups. According to our counterfactual estimates, the price of the average car is approximately €341 higher than in the absence of search frictions.

In line with recent theoretical work, we have argued that the effects of lower search costs in a market are potentially ambiguous. On the one hand, lower search costs result in more search and thereby lead to stronger pressure on firms to cut prices. On the other hand, lower search costs make it easier for a firm to enter the search set of a consumer, which weakens the incentives of firms to cut prices. Finally, higher search costs push some inelastic consumers out of the market which changes the overall elasticity of demand. In our application we have found that prices can go up for some car models when moving from a search model to a full information model.

Finally, we have investigated the effects of changes in the way car manufacturers use their dealership networks to retail their cars. Intuition suggests that the effect of retailing more car brands within a dealership on prices and profits is likely to be ambiguous. On the one hand, if a firm offers more cars at a dealership, this dealership becomes more attractive for consumers because of the implied economies of search associated with one-stop shopping. This demand effect tends to increase prices and variable profits. However, if a firm chooses to offer more cars at its dealerships then rival firms need to be more aggressive if they wish to enter consumer search sets. This competition effect tends to lower prices and decrease variable profits. For the case of the Dutch market, we have found that if Renault started to sell Dacia cars in all its dealerships, then the decrease in sales of Renault cars would be more than offset by the increase in sales of Dacia cars as well as the higher prices for Dacia cars, and as a result the variable profits of the Renault-Nissan Alliance go up.

One of the limitations of our model is that we are mostly ignoring the vertical structure of the market. For instance, our model does not take bargaining between the car dealership and the consumer into account. We only observe list prices in our data and our survey data only allows us to observe whether a consumer visited a dealership of a specific brand, which makes it difficult to infer whether the consumer visited a dealer to learn more about the characteristics of a car or to bargain for a better price. We leave it to future work to focus more on this aspect of the market.[41] Another interesting vertical aspect of this market which we leave for future work is how the network of dealerships is determined, especially when search frictions are important.

[41]In recent work, D'Haultfœuille, Durrmeyer, and Février (2018) develop a full information BLP-type model that includes bargaining and can be estimated using aggregate data even when transaction prices are not observed.

References

Abaluck, Jason and Abi Adams: "What Do Consumers Consider Before They Choose? Identification from Asymmetric Demand Responses," Mimeo, 2018.

Albuquerque, Paulo and Bart J. Bronnenberg: "Measuring the Impact of Negative Demand Shocks on Car Dealer Networks," *Marketing Science* 31, 4–23, 2012.

Anderson, Simon P. and Régis Renault: "Pricing, Product Diversity, and Search Costs: A Bertrand-Chamberlin-Diamond Model," *RAND Journal of Economics* 30, 719–735, 1999.

Armstrong, Timothy B.: "Large Market Asymptotics for Differentiated Product Demand Estimators with Economic Models of Supply," *Econometrica* 84, 1961–1980, 2016.

Armstrong, Mark: "Ordered Consumer Search," *Journal of European Economic Association* 15, 989–1024, 2017.

Armstrong, Mark and John Vickers: "Which Demand Systems Can Be Generated by Discrete Choice?" *Journal of Economic Theory* 158-A, 293–307, 2015.

Armstrong, Mark and Jidong Zhou: "Paying for Prominence," *Economic Journal* 121, F368–F395, 2011.

Bar-Isaac, Heski, Guillermo Caruana, and Vicente Cuñat: "Search, Design and Market Structure," *American Economic Review* 102, 1140–1160, 2012.

Berry, Steven: "Estimating Discrete Choice Models of Product Differentiation," *RAND Journal of Economics* 25, 242–262, 1994.

Berry, Steven, James Levinsohn, and Ariel Pakes: "Automobile Prices in Market Equilibrium," *Econometrica* 63, 841–890, 1995.

Berry, Steven, James Levinsohn, and Ariel Pakes: "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," *Journal of Political Economy* 112, 68–105, 2004.

Berry, Steven and Ariel Pakes: "The Pure Characteristics Demand Model," *International Economic Review* 48, 1193–1225, 2007.

Burdett, Kenneth and Kenneth L. Judd: "Equilibrium Price Dispersion," *Econometrica* 51, 955–969, 1983.

Chade, Hector and Lones Smith: "Simultaneous Search," *Econometrica* 74, 1293–1307, 2006.

Choi, Michael, Anovia Yifan Dai, and Kyungmin Kim: "Consumer Search and Price Competition," *Econometrica* 86, 1257–1281, 2018.

De los Santos, Babur, Ali Hortaçsu, and Matthijs R. Wildenbeest: "Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior." *American Economic Review* 102, 2955–2980, 2012.

De los Santos, Babur, Ali Hortaçsu, and Matthijs R. Wildenbeest: "Search with Learning for Differentiated Products: Evidence from E-Commerce," *Journal of Business & Economic Statistics* 35, 626–641, 2017.

D'Haultfœuille, Xavier, Isis Durrmeyer, and Philippe Février: "Automobile Prices in Market Equilibrium with Unobserved Price Discrimination," forthcoming in the *Review of Economic Studies*, 2018.

Dinerstein, Michael, Liran Einav, Jonathan Levin, and Neel Sundaresan: "Consumer Price Search and Platform Design in Internet Commerce," *American Economic Review* 108, 1820–1859, 2018.

Finkelstein, Sydney: "The DaimlerChrysler Merger," Tuck School of Business at Dartmouth Case #1-0071, 2002.

Galenianos, Manolis and Alessandro Gavazza: "A Structural Model of the Retail Market for Illicit Drugs," *American Economic Review* 107, 858–896, 2017.

Gandhi, Amit and Jean-François Houde: "Measuring Substitution Patterns in Differentiated Products Industries," Mimeo, 2016.

Goldberg, Pinelopi Koujianou: "Product Differentiation and Oligopoly in International Markets: The Case of the U.S. Automobile Industry," *Econometrica* 63, 891–951, 1995.

Goolsbee, Austan and Amil Petrin: "The Consumer Gains from Direct Broadcast Satellites and the Competition with Cable TV," *Econometrica* 72, 351–381, 2004.

Haan, Marco, José Luis Moraga-González, and Vaiva Petrikaitė: "A Model of Directed Consumer Search," *International Journal of Industrial Organization* 61, 223–255, 2018.

Hong, Han and Matthew Shum: "Using Price Distributions to Estimate Search Costs," *RAND Journal of Economics* 37, 257–275, 2006.

Honka, Elisabeth: "Quantifying Search and Switching Costs in the U.S. Auto Insurance Industry," *RAND Journal of Economics* 45, 847–884, 2014.

Honka, Elisabeth, Ali Hortacşu, and Maria Ana Vitorino: "Advertising, Consumer Awareness, and Choice: Evidence from the U.S. Banking Industry," *RAND Journal of Economics* 48, 611–646, 2017.

Hortaçsu, Ali and Chad Syverson: "Product Differentiation, Search Costs, and Competition in the Mutual Fund Industry: A Case Study of S&P 500 Index Funds," *Quarterly Journal of Economics* 119, 403–456, 2004.

Jolivet, Grégory and Hélène Turon: "Consumer Search Costs and Preferences on the Internet," forthcoming in the *Review of Economic Studies*, 2018.

Kim, Jun B., Paulo Albuquerque, and Bart J. Bronnenberg: "Online Demand Under Limited Consumer Search," *Marketing Science* 29, 1001–1023, 2010.

Kim, Jun B., Paulo Albuquerque, and Bart J. Bronnenberg: "The Probit Choice Model Under Sequential Search with an Application to Online Retailing," *Management Science* 63, 3911–3929, 2017.

Koulayev, Sergei: "Search for Differentiated Products: Identification and Estimation," *RAND Journal of Economics* 45, 553–575, 2014.

Mehta, Nitin, Surendra Rajiv, and Kannan Srinivasan: "Price Uncertainty and Consumer Search: A Structural Model of Consideration Set Formation," *Marketing Science* 22, 58–84, 2003.

Moorthy, Sridhar, Brian T. Ratchford and Debabrata Talukdar: "Consumer Information Search Revisited: Theory and Empirical Analysis," *Journal of Consumer Resesarch* 23, 263–277, 1997.

Moraga-González, José Luis and Matthijs R. Wildenbeest: "Maximum Likelihood Estimation of Search Costs," *European Economic Review* 52, 820–848, 2008.

Moraga-González, José Luis, and Vaiva Petrikaitė: "Search Costs, Demand-Side Economies and the Incentives to Merge under Bertrand Competition," *RAND Journal of Economics* 44, 391–424, 2013.

Moraga-González, José Luis, Zsolt Sándor, and Matthijs R. Wildenbeest: "Consumer Search and Prices in the Automobile Market," Mimeo, 2015.

Moraga-González, José Luis, Zsolt Sándor, and Matthijs R. Wildenbeest: "Prices and Heterogeneous Search Costs,"*RAND Journal of Economics* 48, 125–146, 2017.

Murry, Charles and Henry S. Schneider: "The Economics of Retail Markets for New and Used Cars," in the *Handbook on the Economics of Retail and Distribution*, edited by Emek Basker. Edward Elgar Publishers, 2016.

Murry, Charles and Yiyi Zhou: "Consumer Search and Automobile Dealer Co-Location," Mimeo, 2018.

Nevo, Aviv: "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica* 69, 307–342, 2001.

Nurski, Laura and Frank Verboven: "Exclusive Dealing as a Barrier to Entry? Evidence from Automobiles," *Review of Economic Studies* 83, 1156–1188, 2016.

Petrin, Amil: "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy* 110, 705–729, 2002.

Pires, Tiago: "Costly Search and Consideration Sets in Storable Goods Markets," *Quantitative Marketing and Economics* 14, 157–193, 2016.

Reinganum, Jennifer F.: "A Simple Model of Equilibrium Price Dispersion," *Journal of Political Economy* 87, 851–858, 1979.

Reynaert, Mathias and Frank Verboven: "Improving the Performance of Random Coefficients Demand Models: The Role of Optimal Instruments," *Journal of Econometrics* 179, 83–98, 2014.

Reynaerts, Jo, Ravi Varadhan, and John C. Nash: "Enhancing the Convergence Properties of the BLP (1995) Contraction Mapping," Vives Discussion Paper # 35, 2012.

Roberts, John H. and James M. Lattin: "Development and Testing of a Model of Consideration Set Composition," *Journal of Marketing Research* 28, 429–440, 1991.

Sándor, Zsolt and Péter András: "Alternative Sampling Methods for Estimating Multivariate Normal Probabilities," *Journal of Econometrics* 120, 207–234, 2004.

Scott Morton, Fiona, Jorge Silva-Risso, and Florian Zettelmeyer: "What Matters in a Price Negotiation: Evidence from the U.S. Auto Retailing Industry," *Quantitative Marketing and Economics* 9, 365–402, 2011.

Seiler, Stephan: "The Impact of Search Costs on Consumer Behavior: A Dynamic Approach," *Quantitative Marketing and Economics* 11, 155–203, 2013.

Skrainka, Benjamin S.: "A Large Scale Study of the Small Sample Performance of Random Coefficient Models of Demand," Mimeo, 2012.

Sovinsky Goeree, Michelle: "Limited Information and Advertising in the U.S. Personal Computer Industry," *Econometrica* 76, 1017–1074, 2008.

Stahl, Dale O.: "Oligopolistic Pricing with Sequential Consumer Search," *American Economic Review* 79, 700–712, 1989.

Stigler, George: "The Economics of Information," *Journal of Political Economy* 69, 213–225, 1961.

Train, Kenneth E. and Clifford Winston: "Vehicle Choice Behavior and the Declining Market Share of U.S. Automakers," *International Economic Review* 48, 1469–1496, 2007.

Varadhan, Ravi and Christophe Roland: "Simple and Globally Convergent Methods for Accelerating the Convergence of Any EM Algorithm," *Scandinavian Journal of Statistics* 35, 335–353, 2008.

Weitzman, Martin L.: "Optimal Search for the Best Alternative," *Econometrica* 47, 641–654, 1979.

Wolinsky, Asher: "True Monopolistic Competition as a Result of Imperfect Competition," *Quarterly Journal of Economics* 101, 493–511, 1986.

Wolinsky, Asher: "Procurement via Sequential Search," *Journal of Political Economy* 113, 785–810, 2005.

APPENDIX

A Proof of Lemma 1

Rewrite H_{if} as follows:

$$H_{if}(r) = \int_r^\infty (z - r) dF(z - \delta_{if}).$$

Using the change of variables $t = z - \delta_{if}$, we get

$$H_{if}(r) = \int_{r-\delta_{if}}^\infty (t - (r - \delta_{if})) dF(t).$$

Notice that the right-hand side of this expression is just $H_0(r - \delta_{if})$. Now, recall that r_{if} solves $H_{if}(r) = c_{if}$, so that $r_{if} = H_{if}^{-1}(c_{if})$. Because $H_{if}(r) = H_0(r - \delta_{if})$, then $r_{if} - \delta_{if} = H_0^{-1}(c_{if})$ and the result follows. ■

B Proof of Proposition 1

According to equation (11) there is a one-to-one relationship between the search cost distribution and the distribution of the random variable w that determines the visiting probabilities, i.e.,

$$F_{if}^c(H_{if}(z)) = \frac{1 - F_{if}^w(z)}{1 - F_{if}(z)}. \tag{A28}$$

This relationship is extremely useful because it suggests that, for a given distribution of the maximum utility at a seller F_{if}, we can choose an appropriate distribution for the w's for which a search cost distribution exists that rationalizes it according to equation (A28). Specifically, use the change of variables $c = H_{if}(z)$ in equation (A28) to obtain

$$F_{if}^c(c) = \frac{1 - F_{if}^w(H_{if}^{-1}(c))}{1 - F_{if}(H_{if}^{-1}(c))} = \frac{1 - F_{if}^w(\delta_{if} + H_0^{-1}(c))}{1 - F_{if}(\delta_{if} + H_0^{-1}(c))} = \frac{1 - F_{if}^w(\delta_{if} + H_0^{-1}(c))}{1 - F(H_0^{-1}(c))}, \tag{A29}$$

where we have used Lemma 1 to derive the second equality. Denote by μ_{if} the location parameter of the search cost distribution, which contains the search cost covariates. To obtain a closed-form expression for the buying probabilities it is convenient to assume that the w_{if}'s follow a Gumbel distribution. Moreover, we want to make sure that the search cost distribution in equation (A29) is a function of μ_{if} but at the same time does not depend on δ_{if}, which can be achieved by assuming that the location parameter of F_{if}^w is $\delta_{if} - \mu_{if}$, i.e.,

$$F_{if}^w(z) = \exp\left(-\exp\left(-(z - (\delta_{if} - \mu_{if}))\right)\right).$$

Evaluating this expression at $\delta_{if}+H_0^{-1}(c)$ gives $F_{if}^w = \exp(-\exp(-(H_0^{-1}+\mu_{if})))$, which equals $F(H_0^{-1}+\mu_{if})$. Substituting this into equation (A29) gives

$$F_{if}^c(c) = \frac{1 - F(H_0^{-1}(c) + \mu_{if})}{1 - F(H_0^{-1}(c))}.$$

Using that $\partial H_0^{-1}(c)/\partial c = -1/(1 - F(H_0^{-1}(c)))$, the corresponding search cost density is given by

$$f_{if}^c(c) = \frac{f\left(H_0^{-1}(c) + \mu_{if}\right) - F_{if}^c(c) \cdot f\left(H_0^{-1}(c)\right)}{\left(1 - F(H_0^{-1}(c))\right)^2},$$

where $f(\cdot)$ is the density of match values ε_{ij}. Because we have taken the utility shock distribution F to be the TIEV distribution we get:

$$F_{if}^c(c) = \frac{1 - \exp\left(-\exp\left(-H_0^{-1}(c) - \mu_{if}\right)\right)}{1 - \exp(-\exp(-H_0^{-1}(c)))}.$$

■

C Derivation of the Probability of Searching Once

We first note that:

$$\begin{aligned}
v_{i1} &= \sum_f \Pr\left[u_{i0} < \max_{k\neq 0}\{r_{ik}\} \text{ and } \max\{u_{i0}, u_{if}\} > \max_{k\neq 0,f}\{r_{ik}\} \text{ and } r_{if} > \max_{k\neq 0,f}\{r_{ik}\}\right]. \\
&= \sum_f \Pr\left[u_{i0} > u_{if} \text{ and } r_{if} > u_{i0} > \max_{k\neq 0,f}\{r_{ik}\} \text{ and } r_{if} > \max_{k\neq 0,f}\{r_{ik}\}\right] \\
&\quad + \sum_f \Pr\left[u_{i0} < r_{if} \text{ and } u_{if} > u_{i0} \text{ and } u_{if} > \max_{k\neq 0,f}\{r_{ik}\} \text{ and } r_{if} > \max_{k\neq 0,f}\{r_{ik}\}\right] \\
&= \sum_f \Pr\left[u_{i0} > u_{if} \text{ and } r_{if} > u_{i0} > \max_{k\neq 0,f}\{r_{ik}\}\right] \\
&\quad + \sum_f \Pr\left[u_{i0} < r_{if} \text{ and } u_{if} > u_{i0} \text{ and } u_{if} > \max_{k\neq 0,f}\{r_{ik}\} \text{ and } r_{if} > \max_{k\neq 0,f}\{r_{ik}\}\right] \\
&= \sum_f \Pr\left[u_{i0} > u_{if} \text{ and } r_{if} > u_{i0} > \max_{k\neq 0,f}\{r_{ik}\}\right] \\
&\quad + \sum_f \Pr\left[\max_{k\neq 0,f}\{r_{ik}\} < u_{i0} < r_{if} \text{ and } u_{if} > u_{i0} \text{ and } r_{if} > \max_{k\neq 0,f}\{r_{ik}\}\right] \\
&\quad + \sum_f \Pr\left[u_{i0} < \max_{k\neq 0,f}\{r_{ik}\} \text{ and } u_{if} > \max_{k\neq 0,f}\{r_{ik}\} \text{ and } r_{if} > \max_{k\neq 0,f}\{r_{ik}\}\right].
\end{aligned}$$

The sum across f appears because the first searched option can be any of the dealerships. We can simplify this to

$$
\begin{aligned}
v_{i1} &= \sum_f \Pr\left[u_{i0} < r_{if} \text{ and } \max_{k\neq 0,f} r_{ik} < u_{i0}\right] \\
&\quad + \sum_f \Pr\left[u_{i0} < \max_{k\neq 0,f} r_{ik} \text{ and } \max_{k\neq 0,f} r_{ik} < u_{if} \text{ and } \max_{k\neq 0,f} r_{ik} < r_{if}\right].
\end{aligned}
$$

This probability can be written as

$$
\begin{aligned}
v_{i1} &= \sum_f \int_{-\infty}^{\infty} \left(1 - F^r_{if}(y)\right) F^r_{i,-f}(y) f(y) dy \\
&\quad + \sum_f \int_{-\infty}^{\infty} F(x)\left(1 - F_{if}(x)\right)\left(1 - F^r_{if}(x)\right) f^r_{i,-f}(x) dx,
\end{aligned}
$$

where $F^r_{i,-f}$ and $f^r_{i,-f}$ are the CDF and PDF of $\max_{k\neq 0,f} r_{ik}$. Note that by equation (11),

$$
\int_{-\infty}^{\infty} F(x)\left(1 - F_{if}(x)\right)\left(1 - F^r_{if}(x)\right) f^r_{i,-f}(x) dx = \int_{-\infty}^{\infty} F(x)\left(1 - F^w_{if}(x)\right) f^r_{i,-f}(x) dx,
$$

and by integration by parts

$$
\begin{aligned}
\int_{-\infty}^{\infty} F(x)\left(1 - F^w_{if}(x)\right) f^r_{i,-f}(x) dx &= \int_{-\infty}^{\infty} F(x)\left(1 - F^w_{if}(x)\right) dF^r_{i,-f}(x) \\
&= -\int_{-\infty}^{\infty} F^r_{i,-f}(x) d\left[F(x)\left(1 - F^w_{if}(x)\right)\right] \\
&= \int_{-\infty}^{\infty} F^r_{i,-f}(x) F(x) f^w_{if}(x) dx - \int_{-\infty}^{\infty} \left(1 - F^w_{if}(x)\right) F^r_{i,-f}(x) f(x) dx,
\end{aligned}
$$

where we use

$$
d\left[F(x)\left(1 - F^w_{if}(x)\right)\right] = f(x)\left(1 - F^w_{if}(x)\right) - F(x) f^w_{if}(x).
$$

We finally get that

$$
v_{i1} = \sum_f \int_{-\infty}^{\infty} \left(F^w_{if}(y) - F^r_{if}(y)\right) F^r_{i,-f}(y) f(y) dy + \sum_f \int_{-\infty}^{\infty} F^r_{i,-f}(x) F(x) f^w_{if}(x) dx. \tag{A30}
$$

where $F^r_{i,-f}$ and $f^r_{i,-f}$ are the CDF and PDF of $\max_{k\neq 0,f}\{r_{ik}\}$.

The share of consumers searching only one time is then given by

$$
q_1 = \int v_{i1} dF_\tau(\tau_i).
$$

where τ_i contain the random coefficients and the Gumbel variables. ■

D Market Share Derivatives

Own-price derivatives

The market share derivative with respect to price is given by

$$\begin{aligned}\frac{\partial s_j}{\partial p_j} &= \int \frac{\partial s_{ij}}{\partial p_j} dF_\tau(\tau_i);\\ &= \int -\alpha_i \frac{\partial s_{ij}}{\partial \delta_{ij}} dF_\tau(\tau_i);\end{aligned}$$

The derivative of the buying probability s_{ij} with respect to δ_{ij} is

$$\begin{aligned}\frac{\partial s_{ij}}{\partial \delta_{ij}} &= \frac{\partial P_{if}}{\partial \delta_{if}} \frac{\partial \delta_{if}}{\partial \delta_{ij}} P_{ij|f} + P_{if} \frac{\partial P_{ij|f}}{\partial \delta_{ij}};\\ &= \frac{\partial P_{if}}{\partial \delta_{if}} P_{ij|f}^2 + s_{ij}(1 - P_{ij|f});\end{aligned}$$

The derivative of P_{if} with respect to δ_{if} is given by

$$\frac{\partial P_{if}}{\partial \delta_{if}} = \int \left(\prod_{g \neq f} F_{ig}^w(z) \right) \frac{\partial f_{if}^w(z)}{\partial \delta_{if}} dz.$$

The density of w is given by

$$f_{if}^w(z) = f_{if}^c(H_{if}(z))(1 - F_{if}(z))^2 + F_{if}^c(H_{if}(z)) f_{if}(z).$$

Using $\partial H_{if}(z)/\partial \delta_{if} = 1 - F_{if}(z)$, the derivative of f_{if}^w with respect to δ_{if} is given by

$$\frac{\partial f_{if}^w}{\partial \delta_{if}} = f_{if}'^c(1 - F_{if})^3 + 3 f_{if}^c(1 - F_{if}) f_{if} + F_{if}^c f_{if}(1 - \exp(\delta_{if} - z)),$$

where f_{if}^c and F_{if}^c are evaluated at $H_{if}(z)$, and f_{if} and F_{if} are evaluated at z. Note that when the search cost distribution is normal then $f_{if}'^c = f_{if}^c \cdot (\mu_c - H_{if}(z))$, where μ_c is the mean parameter of the search cost distribution.

Cross-price derivatives

The market share derivative of product j with respect to the price of product k is given by

$$\begin{aligned}\frac{\partial s_j}{\partial p_k} &= \int \frac{\partial s_{ij}}{\partial p_k} dF_\tau(\tau_i);\\ &= \int -\alpha_i \frac{\partial s_{ij}}{\partial \delta_{ik}} dF_\tau(\tau_i);\end{aligned}$$

If product k is sold by firm f, the derivative of the buying probability s_{ij} with respect to δ_{ik} is

$$\begin{aligned}\frac{\partial s_{ij}}{\partial \delta_{ik}} &= \frac{\partial P_{if}}{\partial \delta_{if}} \frac{\partial \delta_{if}}{\partial \delta_{ik}} P_{ij|f} + P_{if} \frac{\partial P_{ij|f}}{\partial \delta_{ik}}; \\ &= \frac{\partial P_{if}}{\partial \delta_{if}} P_{ik|f} P_{ij|f} - s_{ij} P_{ik|f}.\end{aligned}$$

If product k is sold by another firm, the derivative of the buying probability s_{ij} with respect to δ_{ik} is

$$\begin{aligned}\frac{\partial s_{ij}}{\partial \delta_{ik}} &= \frac{\partial P_{if}}{\partial \delta_{ih}} \frac{\partial \delta_{ih}}{\partial \delta_{ik}} P_{ij|f}; \\ &= \frac{\partial P_{if}}{\partial \delta_{ih}} P_{ik|h} P_{ij|f}.\end{aligned}$$

The derivative of P_{if} with respect to δ_{ih} is given by

$$\frac{\partial P_{if}}{\partial \delta_{ih}} = \int \left(\prod_{g \neq f,h} F_{ig}^{w}(z) \right) \frac{\partial F_{ih}^{w}(z)}{\partial \delta_{ih}} f_{if}^{w}(z) dz.$$

The derivative of F_{ih}^{w} with respect to δ_{ih} is given by

$$\begin{aligned}\frac{\partial F_{ih}^{w}}{\partial \delta_{ih}} &= -f_{ih}^{c}(1 - F_{ih})^2 - F_{ih}^{c} f_{ih}(z); \\ &= -f_{ih}^{w}.\end{aligned}$$

E Contraction Mapping

Contraction Theorem (BLP). Let $f : \mathbb{R}^J \to \mathbb{R}^J$ be defined as

$$f_j(\xi) = \xi_j + \log s_j - \log \sigma_j(\xi), \quad j = 1, \ldots, J,$$

where $s = (s_1, \ldots, s_J)$ is the vector of observed market shares and suppose that the market share vector $\sigma(\xi)$ as a function of $\xi = (\xi_1, \ldots, \xi_J) \in \mathbb{R}^J$ satisfies the following conditions.

1. σ is continuously differentiable in ξ and

$$\frac{\partial \sigma_j}{\partial \xi_j}(\xi) \leq \sigma_j(\xi), \quad \frac{\partial \sigma_j}{\partial \xi_k}(\xi) < 0 \text{ for any } j,\ k \neq j \text{ and } \xi \in \mathbb{R}^J,$$

   (the former is equivalent to the fact that the function $\overline{\sigma}_j : \mathbb{R}^J \to \mathbb{R}$, $\overline{\sigma}_j(\xi) = \sigma_j(\xi) \exp(-\xi_j)$ is decreasing in ξ_j) and

$$\sum_{k=1}^{J} \frac{\partial \sigma_j}{\partial \xi_k}(\xi) > 0 \text{ for any } \xi \in \mathbb{R}^J.$$

2. The share of the outside alternative $\sigma_0(\xi) = 1 - \sum_{j=1}^{J} \sigma_j(\xi)$ is decreasing in all its arguments and it

satisfies that for any j and $x \in \mathbb{R}$ the limit

$$\lim_{\xi_{-j} \to -\infty} \sigma_0 \left(\xi_1, \ldots, \xi_{j-1}, x, \xi_{j+1}, \ldots, \xi_J\right) \equiv \widetilde{\sigma}_0^j \left(x\right)$$

is finite and the function $\widetilde{\sigma}_0^j : \mathbb{R} \to \mathbb{R}$ obtained as the limit satisfies that

$$\lim_{x \to -\infty} \widetilde{\sigma}_0^j \left(x\right) = 1 \quad \text{and} \quad \lim_{x \to \infty} \widetilde{\sigma}_0^j \left(x\right) = 0,$$

where $\xi_{-j} \to -\infty$ means that $\xi_1 \to -\infty, \ldots, \xi_{j-1} \to -\infty, \xi_{j+1} \to -\infty, \ldots, \xi_J \to -\infty$.

3. The function $\overline{\sigma}_j$ defined in Condition 1 satisfies

$$\lim_{\xi \to -\infty} \overline{\sigma}_j \left(\xi\right) > 0.$$

Then there are values $\underline{\xi}, \overline{\xi} \in \mathbb{R}$ such that the function $\overline{f} : \left[\underline{\xi}, \overline{\xi}\right]^J \to \mathbb{R}^J$ defined by $\overline{f}_j \left(\xi\right) = \min\left[\overline{\xi}, f_j\left(\xi\right)\right]$ has the property that $\overline{f}\left(\left[\underline{\xi}, \overline{\xi}\right]^J\right) \subseteq \left[\underline{\xi}, \overline{\xi}\right]^J$, is a contraction with modulus less than 1 with respect to the sup norm $\|(x_1, \ldots, x_J)\| = \max_j |x_j|$, and, in addition, f has no fixed point outside $\left[\underline{\xi}, \overline{\xi}\right]^J$.

Here we verify that conditions 1, 2, and 3 of this theorem are satisfied for σ equal to the market share vector function $s = (s_1, \ldots, s_J)$, where $s_j = \int s_{ij} dF_\tau(\tau_i)$, $j = 1, \ldots, J$, in the case specified in Section 2.4. Note that equation (16) implies that

$$s_{i0} = \frac{1}{1 + \sum_{k=1}^J \exp\left(\delta_{ik} - \mu_{ig}\right)} \tag{A31}$$

and the derivatives

$$\begin{aligned}
\frac{\partial s_{ij}}{\partial \xi_j} &= \left(1 - s_{ij}\right) s_{ij}, \quad \frac{\partial s_{ij}}{\partial \xi_k} = -s_{ik} s_{ij} \quad \text{for} \quad j \neq k, \quad \frac{\partial s_{i0}}{\partial \xi_j} = -s_{ij} s_{i0}, \\
\frac{\partial s_j}{\partial \xi_j} &= \int \left(1 - s_{ij}\right) s_{ij} dF_\tau(\tau_i), \quad \frac{\partial s_j}{\partial \xi_k} = -\int s_{ij} s_{ik} dF_\tau(\tau_i), \quad \frac{\partial s_0}{\partial \xi_j} = -\int s_{ij} s_{i0} dF_\tau(\tau_i).
\end{aligned} \tag{A32}$$

Condition 1. Clearly the market share vector s is continuously differentiable in ξ. We can see that $\frac{\partial s_j}{\partial \xi_j} \leq s_j$ holds because $\frac{\partial s_j}{\partial \xi_j} - s_j = -\int s_{ij}^2 dF_\tau(\tau_i) \leq 0$. The inequality $\frac{\partial s_j}{\partial \xi_k} < 0$ holds obviously. The third inequality, $\sum_{k=1}^J \frac{\partial s_j}{\partial \xi_k} > 0$ follows by observing that

$$\sum_{k=1}^J \frac{\partial s_j}{\partial \xi_k} = \sum_{k=1}^J \frac{\partial s_k}{\partial \xi_j} = -\frac{\partial s_0}{\partial \xi_j},$$

which is positive by equation (A32).

Condition 2. The fact that the share of the outside alternative $s_0 = 1 - \sum_{j=1}^J s_j$ is decreasing in all its arguments follows from equation (A32). Next we compute the limit

$$\lim_{\xi_{-j}\to-\infty} s_0\left(\xi_1,\ldots,\xi_{j-1},x,\xi_{j+1},\ldots,\xi_J\right)\equiv \widetilde{s}_0^{\,j}\left(x\right).$$

From equation (A31) we see that

$$\widetilde{s}_0^{\,j}\left(x\right)=\frac{1}{1+\exp\left(\delta_{ij}\left(x\right)-\mu_{ig}\right)},$$

where $\delta_{ij}\left(x\right)$ denotes the expression δ_{ij} where ξ_j is replaced by x. From this it is straightforward to obtain that $\lim_{x\to-\infty}\widetilde{s}_0^{\,j}\left(x\right)=1$ and $\lim_{x\to\infty}\widetilde{s}_0^{\,j}\left(x\right)=0$.

Condition 3. We show that $\lim_{\xi\to-\infty} s_j\left(\xi\right)\exp\left(-\xi_j\right)>0$. We have

$$\lim_{\xi\to-\infty} s_j\exp\left(-\xi_j\right)=\int\lim_{\xi\to-\infty} s_{ij}\exp\left(-\xi_j\right)dF_\tau(\tau_i).$$

Further, from equation (16) we have

$$s_{ij}\exp\left(-\xi_j\right)=\frac{\exp\left(\alpha_i p_j+x_j'\beta_i-\mu_{if}\right)}{1+\sum_{k=1}^{J}\exp\left(\delta_{ik}-\mu_{ig}\right)},$$

so the numerator does not depend on ξ_j for any $j=1,\ldots,J$. Therefore,

$$\lim_{\xi\to-\infty} s_j\exp\left(-\xi_j\right)=\exp\left(\alpha_i p_j+x_j'\beta_i-\mu_{if}\right),$$

which is strictly positive. In conclusion, the contraction property holds.

F Calculation of the Moments

In order to describe the computation of the moments it is useful to introduce some notation. Suppose that we observe the demographic characteristics and purchase decisions of N consumers. Let $i\in\{1,\ldots,N\}$ and for simplicity maintain the notation v_i for the vector of consumer i's unobserved and observed characteristics (i.e., $V_i=\operatorname{diag}\left(v_i\right)$). Denote by $a_i\in\{0,1,\ldots,J\}$ the choice of i, y_i the income of i and r_i a discrete demographic characteristic of i out of the vector d_i of all demographic characteristics, like age or family size. In order to be general we use a generic demographic characteristic q_i for either y_i or r_i, and let $\mathcal{R}$ be a partitioning of the possible values of q_i into a few (two or three) subsets. Let T denote a certain group of products, like family car. We consider moments based on the following type of conditional expectations:

$$E\left[1\left(a_i\in T\right)|q_i\in R_k\right],\quad k=1,2\ \text{ where }\ \mathcal{R}=\left\{R_1,R_2\right\}.$$

For specific examples we refer to Section 3.

Let $R\in\mathcal{R}$. Since we observe the choice of each consumer $i\in\{1,\ldots,N\}$, the moments boil down to aggregation of the choice a_{iT} of i regarding T over those consumers i whose demographic characteristic satisfies that $q_i \in R$, where

$$a_{iT} = \begin{cases} 1 & \text{if } \; a_i \in T \\ 0 & \text{otherwise.} \end{cases}$$

Note that a_{iT} is a Bernoulli random variable with success probability

$$s_{iT}(\theta) = P(a_{iT} = 1|p, x, \theta, \xi, v_i) = P(a_i \in T|p, x, \theta, \xi, v_i) = \sum_{j \in T} P(a_i = j|p, x, \theta, \xi, v_i) = \sum_{j \in T} s_{ij}$$

and independent across i conditional on (p, x, θ, ξ, v_i). Therefore,

$$\begin{aligned} E[a_{iT}|p, x, \theta, \xi, v_i] &= s_{iT}(\theta) \\ \operatorname{var}[a_{iT}|p, x, \theta, \xi, v_i] &= s_{iT}(\theta)(1 - s_{iT}(\theta)). \end{aligned} \tag{A33}$$

Aggregation over i s.t. $q_i \in R$ yields the moment

$$g_q^R(\theta) = \frac{1}{N_R} \sum_{\substack{i=1 \\ q_i \in R}}^{N} (a_{iT} - s_{iT}(\theta)), \tag{A34}$$

where N_R is the number of consumers i for which $q_i \in R$. This is just the sample counterpart of the moment condition

$$E[a_{iT} - s_{iT}(\theta)|p, x, \theta, \xi, v_i] = 0$$

over the sample of consumers i with $q_i \in R$.

G The Weighting Matrix

The optimal choice of Ξ is a matrix proportional to

$$\left[\operatorname{Var}\left(g\left(\theta^0\right) \mid Z\right)\right]^{-1} = \left(E\left[g\left(\theta^0\right) g\left(\theta^0\right)' \mid Z\right]\right)^{-1},$$

where θ^0 is the true value of the parameter vector θ. Since several micromoments depend on the demand unobserved characteristics ξ, this weighting matrix is not a block diagonal in general. Several components of this matrix can be computed exactly; for example, the variances of the moments (see equation (A33)). However, several other components cannot be computed exactly. Therefore, in order to obtain a positive definite weighting matrix, we propose this matrix to be approximated as $\Xi = V_N(\hat{\theta})$, where $\hat{\theta}$ is a previously obtained consistent estimator of θ^0 and $V_N\left(\hat{\theta}\right)$ is a diagonal matrix whose main diagonal contains the variances of the moments computed as in equation (A33). An alternative choice for Ξ is $\Xi = I_M$, where I_M is the identity matrix of dimension M, which is the number of moments.