Consumer Search and Prices in the Automobile Market*

José Luis Moraga-González†
Zsolt Sándor‡
Matthijs R. Wildenbeest§

March 2022

Abstract

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. We characterize the search problem as a standard discrete choice problem and propose a parametric search cost distribution that generates closed-form expressions for the probability of purchasing a product. We propose a method to estimate the model that supplements aggregate product data with individual-specific data which allows for the separate identification of search costs and preferences. We estimate the model using data from the automobile industry and find that search costs have non-trivial implications for elasticities and markups. We study the effects of exclusive dealing regulation and find that firms benefit at the expense of consumers, who face higher search costs and higher prices than would be the case if multi-brand dealerships were used.

Keywords: consumer search, differentiated products, demand estimation, automobiles, exclusive dealing

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, Frank Verboven, as well as three anonymous referees and the editor 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, CEPR, and the Public-Private Sector Research Center at IESE (Barcelona).

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

§Indiana University, Kelley School of Business and CEPR, 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 et al., 1995, 2004; Nevo, 2001; Petrin, 2002) has proceeded by assuming that consumers have perfect information about all the products available in the market. In markets like those referred to above, the full information assumption is, arguably, unrealistic. 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 et al. (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. Because consumers choose to visit distinct sellers even if they have similar preferences, substitution patterns across products are not only driven by product differentiation but also by the variation in consumer information sets generated by costly search.[2]

We develop our search model in Section 2. In our model, consumers search for differentiated products. We assume consumers have prior information on some of the characteristics of the products, but have to search to figure out whether a product is a good match. Weitzman (1979) has shown that in such a setting the optimal solution to the search problem is to rank sellers in terms of reservation utilities, to visit them in descending reservation utility order, and to stop search when the highest observed utility is above the reservation utility of the next option to be searched. When the number of alternatives is large, applying Weitzman's solution becomes intractable because the number of ways in which a consumer may end up buying a particular product increases factorially in the number of alternatives. For example, with just three alternatives, there are eleven distinct search paths a consumer may follow before deciding to buy a given product, while with ten alternatives there are close to nine million search paths.[3] To address this dimensionality problem, we

[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 et al. (2012) find that over 75 percent of consumers visited only one online bookstore before buying a book online, whereas De los Santos et al. (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 et al., 1997; Scott Morton et al., 2011), and the retail market for illicit drugs (Galenianos and Gavazza, 2017).

[2]Sovinsky Goeree (2008) provides an alternative way to obtain variation in information sets in a BLP-type framework by linking consideration sets to advertising.

[3]Let A, B, and C be the three alternatives. Conditional on buying a specific alternative (let's say A), the eleven different search paths (in order of search) are then A, AB, AC, BA, CA, ABC, ACB, BAC, BCA, CAB, and CBA. More generally, if there are n alternatives, then the number of search paths for each alternative is given by $(n-1)!\sum_{k=0}^{n-1}(n-k)/k!$ (see footnote 4 of Choi et al., 2018).

leverage recent findings from the theoretical search literature that make it possible to compute the purchase 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. More specifically, following Armstrong (2017) and Choi et al. (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 derive purchase probabilities for general search cost distributions, but, because these expressions are generally not closed form and therefore difficult to calculate, we propose a parametric specification for the search cost distribution that generates closed-form expressions. We do this by solving the model backwards: starting from a Type I Extreme Value distribution for the minimum of the reservation value and realized utility, we derive the unique parametric specification for the search cost distribution that rationalizes this distributional assumption. We refer to this search cost distribution as the Gumbel preserving search cost distribution.

We discuss identification and estimation of the model in Section 3. Because both search costs and preferences enter the choice probabilities additively, search costs and preferences are not identified without imposing additional restrictions. One way to achieve identification of search costs is by assuming search costs only depend on variables that are excluded from the utility specification. This approach has the advantage that, conditional on having appropriate instruments, the model can be estimated using aggregate data but a disadvantage is that only variation in search costs is identified and not the level of search costs. Using insights from search theory we propose an alternative strategy that is based on using individual-level data on search behavior. Because the reservation values that guide search decisions respond differently to changes in utility covariates than to changes in search cost covariates, variation in observed search and purchase decisions allows us to separately identify the effect of a covariate on utility and on search costs. Intuitively, 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, alternatively, by relatively high search costs. Without observing search decisions we cannot distinguish between the two. However, search data reveals how much search was done by consumers who end up not buying—if these consumers searched a lot it must be that search costs are low and the utility for the outside option is high; if they did not search that much, it must be that search costs are high.

We propose a method to estimate the model using aggregate data on characteristics and market shares, as well as individual-level data that contains information on purchases and related search behavior. Following Berry et al.'s (2004) two-step procedure, we use the aggregate data to obtain mean utilities of the products. These mean utilities and the individual-level 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 purchase and search probabilities. In a second step we estimate the mean utility parameters from the estimated mean utilities, using an instrumental variables approach to deal with price endogeneity. As pointed out by Berry et al. (2004), an advantage of using the two-step procedure is that instruments are not needed when estimating the nonlinear parameters of the model, which makes estimation of our model less susceptible to misspecification bias.

In Section 4 we apply our model to the Dutch market for new cars. The automobile market is precisely a market in which advertisements, reports in specialized magazines, television programs, and the Internet convey much but not 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. Our data, which supplements the usual aggregate product-level data with survey data that contains information on purchases and searches, 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.[4] We interpret these two facts as being consistent with our search model. Section 4 also discusses the estimation results for our model. Our search cost estimates are precisely estimated 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 cross-distance elasticity estimates indicate that consumers are relatively more likely to substitute towards similar brands when a brand's dealerships are located further away from consumers. Moreover, due to search costs consumers are far more likely to switch within brand than across brands—in a full information model it does not matter for substitution patterns whether models are sold together or not, so a similar pattern does not arise. We conclude that accounting for costly search and its effects on generating heterogeneity in consumer choice sets is important for explaining variability in purchase patterns.

Our model follows most of the consumer search literature in assuming that deviations from equilibrium prices are not observed before searching (e.g., Diamond, 1971; Burdett and Judd, 1983; Wolinsky, 1986). Since consumers' search decisions are based on expected equilibrium prices, this implies that firms are tempted to hold up visiting consumers by raising prices. This is reminiscent of the mechanism behind the *Diamond paradox*, which in case of homogeneous products may lead to a collapse of the market (Diamond, 1971). Although product differentiation keeps the market from collapsing in our model, in Section 5 we show that this Diamond-type hold-up problem nevertheless has important implications for counterfactuals in which dealership networks change. Specifically, we study the role of exclusive dealing, which is the prevalent form of automobile distribution in the European Union and allows manufacturers to put restrictions on how many brands are offered in dealerships. We first show that manufacturers face little incentive to create multi-brand dealerships that sell cars of brands that are part of the same business group—in this case substitution towards other brands owned by the same firm re-enforces the Diamond-type incentives to hold up visiting consumers, which results in substantial price increases that ultimately lower profits. In contrast, when a brand unilaterally adds cars of rival brands to its dealership locations, substitution towards competing brands weakens the hold-up incentive enough to lead to a reduction in prices, with higher profits as a result. Although this creates an incentive for individual firms to start selling competing brands in their dealerships,

[4] 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 69 percent made at least one test drive at one of the visited dealerships.

we show that if the industry collectively adopts these arrangements, industry profits go down substantially. Exclusive dealing regulation could prevent this from happening, with firms benefitting at the expense of consumers who face higher search costs and prices than would be the case with multi-brand dealerships.

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. Early contributions focused on the effects of costly search in homogeneous product markets (see for instance Burdett and Judd, 1983; Stahl, 1989), whereas more recent work has focused on costly search in markets with product differentiation (see Wolinsky, 1986; Anderson and Renault, 1999; Bar-Isaac et al., 2012). Our search model is most closely related to the framework of Wolinsky (1986) but we allow for asymmetric multi-product firms and consumer heterogeneity in preferences and search costs.

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 (Kim et al., 2010, 2017; De los Santos et al., 2012; Seiler, 2013; Honka, 2014; Koulayev, 2014; Pires, 2016; Honka et al., 2017; Seiler and Pinna, 2017; Dinerstein et al., 2018; Jiang et al., 2021). 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.[5] Most of these papers assume search is sequential.[6] In sequential search models search decisions 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 et al. (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 (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.

Our paper also fits into a broader literature that estimates demand for automobiles, which includes BLP, Goldberg (1995), Petrin (2002), and Berry et al. (2004).[7] As in Petrin (2002) and Berry et al. (2004) we use a combination of micro and aggregate data. Our estimation procedure is most similar to Berry et al. (2004), but instead of using moments based on second-choice data we use moments based on purchases and search behavior of individuals. Recent papers in this literature have studied car dealership locations and

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

[6]A computational advantage of non-sequential search (De los Santos et al., 2012; Honka, 2014; Moraga-González et al., 2015; Murry and Zhou, 2020) 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.

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

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, that is, 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. This is important if consumers prefer to have their cars serviced at the dealer, so distance would enter the model through search costs as well as through the indirect utility function. 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. Moreover, unlike in a full information model, pricing decisions in our search model are affected by within-dealer substitution patterns across car models. This has implications for our main counterfactual in which we look at the competitive effects of exclusive dealing: although our results for the case in which rival brands are added confirm Nurski and Verboven's (2016) main finding that there is a collective incentive for exclusive dealing regulation despite unilateral incentives to sell multiple brands, we find the effects on prices, profits, and welfare to be much larger in our search 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 indirect utility consumer i derives from product j is given by:

$$u_{ij} = \bar{\delta}_j + \alpha_i p_j + x_j' \beta_i + \varepsilon_{ij}, \tag{1}$$

where $\bar{\delta}_j = \alpha p_j + x_j' \beta + \xi_j$ is the mean utility for product j, the variable p_j denotes the price of product j, the vector $(x_j, \xi_j, \varepsilon_{ij})$ describes different product attributes from which the consumer derives utility, α and β are scalar and vector coefficients, respectively, that capture the mean marginal utility of price and other product attributes, and α_i and β_i capture consumer heterogeneity in tastes for price and product attributes. We assume that the consumer observes the product attributes contained in 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 is 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 comfort, spaciousness, engine noise, and gearbox smoothness in the case of cars.[8] We assume that the econometrician observes price p_j and the product attributes contained in 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.

To capture that consumers differ in the way they value price and product characteristics, we follow Nevo (2001) in letting the distribution of consumer heterogeneity for product attributes be a function of demographic variables and standard normal draws, i.e., $\beta_i = \Pi D_i + \Sigma v_i$, where D_i is a $d \times 1$ vector of demographic variables, Π is a $(K+1) \times d$ matrix of parameters that measures how the marginal valuation for the K attributes varies with demographics, v_i is a $(K+1) \times 1$ vector of standard normal draws, and Σ is a scaling matrix so that covariances are equal to zero. Note that in our application, we set several of the parameters in Π and on the main diagonal of Σ equal to zero. Because income might affect price sensitivity, by for example facilitating access to credit, we let the price coefficient be inversely proportional to consumer i's yearly income y_i, that is, $\alpha_i = \tilde{\alpha}/y_i$.[9]

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 visit firms to find out the exact utility they derive from the products. We define a search as a trip from home to a single dealer and we assume that consumers then learn the utility they obtain from each of the products available at the firm visited.[10] Search is sequential, ordered, and consumers have costless recall. After each visit to a firm consumers decide whether to buy any of the inspected products so far, to opt for the outside option, or to continue searching. We assume that a consumer i before searching knows (i) the location of and the subset of products available at each firm, (ii) product characteristics x_j and ξ_j as well as the equilibrium price p_j for each product j, (iii) the distribution F of match values ε_{ij}, and (iv) the utility of her outside option ε_{i0}. This means we follow most of the consumer

[8] The term ε_{ij} might also include other information that can be retrieved from visiting a dealer, such as how much the dealer offers for a trade-in.

[9] As shown by Grigolon et al. (2018), for small capitalized car expenditures relative to capitalized income this specification approximates the Cobb-Douglas utility function used in BLP, i.e., $\tilde{\alpha}\log(y_i - p_j) \approx \tilde{\alpha}\log(y_i) - \tilde{\alpha}p_j/y_i$. Note that $\tilde{\alpha}\log(y_i)$ is common to all alternatives, so it drops out of the purchase probabilities.

[10] In general, one can distinguish between store search and brand search. In our model search is across different brands, which is different from searching across stores. Note that dealers of different brands often cluster—although our search model could be reformulated such that consumers search across clusters of brands instead of brands, a limitation of the survey data we use for estimation of the model is that it only contains information on which brands are visited and not which clusters (see Murry and Zhou, 2020, for a model in which consumers search across clusters in the context of a non-sequential search model). Also note that our assumptions imply that in our model a consumer would visit only one dealer for the same brand.

search literature (cf. Diamond, 1971; Wolinsky, 1986; Anderson and Renault, 1999) in assuming consumers have correct conjectures about equilibrium prices while they search, but they may find a deviation price upon visiting a dealer. This assumption has implications for the supply side of the model, which we discuss in detail in Section 2.6 as well as when discussing the counterfactuals in Section 5.

Let c_{if} denote the search cost of consumer i for visiting firm f. We assume that search costs vary across consumers and firms.[11] Let F^c_{if} be the cumulative distribution of consumer i's cost of searching firm f, which we allow to depend on certain search cost covariates through a location parameter μ_{if} (see Section 2.4), and has corresponding density f^c_{if}. 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 16).

2.2 Optimal sequential search

The utility function in equation (1) can be rewritten as $u_{ij} = \delta_{ij} + \varepsilon_{ij}$, where $\delta_{ij} = \bar{\delta}_j + \alpha_i p_j + x'_j \beta_i$ is the mean utility consumer i derives from product j. Because consumers have correct conjectures regarding equilibrium prices and observe x_j and ξ_j, they know δ_{ij} at equilibrium. However, consumers have to search to discover ε_{ij} and actual prices, which can deviate from equilibrium prices. The conditional distribution of consumer i's utility u_{ij} from a given product j is given by $F_{ij}(z) = F(z - \delta_{ij}) = \exp\left(-\exp\left(\delta_{ij} - z\right)\right)$, that is, conditional on demographics and product j's characteristics, consumers' utility is Gumbel with location parameter δ_{ij} and scale parameter 1.[12]

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, that is, $U_{if} = \max_{j \in G_f} \{u_{ij}\}$. Denoting the conditional distribution of U_{if} by F_{if}, we get

$$F_{if}(z) = \Pr[U_{if} \leq z] = \prod_{j \in G_f} F_{ij}(z) = \prod_{j \in G_f} F(z - \delta_{ij}) = F(z - \delta_{if}), \tag{2}$$

where, by the max-stability of the TIEV distribution, $\delta_{if} = \log\left(\sum_{h \in G_f} \exp(\delta_{ih})\right)$. This distribution is conditional on consumer demographics and all the products' attributes sold by firm f. From now on, we will refer to this conditioning as conditioning on δ_{if}.

We now turn to describing consumer i's optimal search strategy. Since dealers sell disjoint sets of products with match values that are independent across consumers and products, conditional on δ_{if}'s, the highest utility draws are independent across consumers and dealers as well.[13] This means we can apply the Weitzman rule to characterize optimal consumer search. To apply this rule, we first define the expected

[11] 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.

[12] Throughout the paper, when we refer to the Gumbel distribution, we mean the distribution with CDF $\exp(-\exp(-(z - m)/b))$, where m is a location parameter and b 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).

[13] Note that if the dealers would be selling overlapping sets of products, Weitzman's rule would not apply since the information obtained at one dealer would be informative about other dealers selling the same products as well. We also note that the observable characteristics are allowed to be correlated, but because consumers observe them before initiating a search, this does not affect the applicability of Weitzman's rule.

gains to consumer i from searching for a product at firm f when the best utility the consumer has found so far is r:

$$H_{if}(r) \equiv \int_r^\infty (z-r)dF_{if}(z). \tag{3}$$

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} of consumer i for firm f as the solution to equation

$$H_{if}(r) - c_{if} = 0 \tag{4}$$

in r. Notice that H_{if} is decreasing and strictly convex so equation (4) has 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, proven in Appendix A, decomposes the reservation value of a consumer i for a firm f into a utility component and a search cost component:

Lemma 1 *Under the assumption that the ε_{ij}'s are IID TIEV-distributed, $H_{if}^{-1}(c_{if}) = \delta_{if} + H_0^{-1}(c_{if})$, so consumer i's reservation value for firm f can be written as*

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

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

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.

2.3 Individual buying probabilities

Because we allow for both consumer and firm heterogeneity, Weitzman's (1979) solution is extremely hard to implement in our setting. To solve this problem, we next utilize a recent finding by Armstrong (2017) and Choi et al. (2018), which consists of a methodology for the computation of the purchase decisions without having to take into account the different search paths consumers may possibly follow.[14]

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

$$w_{if} = \min\{r_{if}, U_{if}\} = \min\left\{r_{if}, \max_{j\in G_f}\{u_{ij}\}\right\}. \tag{5}$$

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

Armstrong (2017) and Choi et al. (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, conditional on the δ_{if}'s and the search cost covariates, the probability that buyer i buys product j is given by

$$s_{ij} = P_{ij|f} P_{if}, \tag{6}$$

where

$$P_{ij|f} = \Pr\left(u_{ij} \geq \max_{h \in G_f} u_{ih}\right) \tag{7}$$

and

$$P_{if} = \Pr\left(w_{if} \geq \max_{g \in \{0\} \cup \mathcal{F}} w_{ig}\right). \tag{8}$$

Here $P_{ij|f}$ denotes the probability that product j is chosen out of the G_f products of firm f conditional on the δ_{ij}'s corresponding to the products of firm f, while P_{if} is the probability of buying from firm f conditional not only on all the δ_{if}'s of all the firms in the market but also on the search cost covariates.

Now we turn to computing the conditional probabilities in equations (7) and (8). Since ε_{ij} is an IID draw from a TIEV distribution, $P_{ij|f}$ in equation (7) has the familiar closed form:

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

To derive P_{if}, notice that r_{if} and U_{if} are independent conditional on δ_{if}. Hence, the conditional distribution of $w_{if} = \min\{r_{if}, U_{if}\}$, denoted F_{if}^{w}, can be obtained by computing the CDF of the minimum of two independent random variables:

$$F_{if}^{w}(z) = 1 - \left(1 - F_{if}^{r}(z)\right)\left(1 - F_{if}(z)\right) = F_{if}^{r}(z)(1 - F_{if}(z)) + F_{if}(z), \tag{9}$$

where F_{if} is given in equation (2) and F_{if}^{r} is the CDF of r_{if} conditional on the δ_{if}'s and the search cost covariates. To obtain F_{if}^{r} we can use equation (3):

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

Substituting the distribution of reservation values into equation (9) gives

$$F_{if}^{w}(z) = 1 - F_{if}^{c}\left(H_{if}(z)\right)\left(1 - F_{if}(z)\right). \tag{10}$$

Equation (10) 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.

Finally, because the w_{if}'s are independent conditional on the δ_{if}'s and the search cost covariates, the conditional probability P_{if} in (8) can be computed as:

$$P_{if} = \int \left(\prod_{g \neq f} F_{ig}^{w}(z) \right) f_{if}^{w}(z)\, dz. \tag{11}$$

2.4 Distributional assumptions and market shares computation

To compute the conditional individual buying probabilities in equation (6), we need to calculate the probabilities $P_{ij|f}$ and P_{if}, which are given in equations (7) and (11), respectively. The main difficulty here is that there is no closed-form solution for the probability P_{if} in equation (11) for arbitrary search cost distributions. Although we can compute P_{if} by first plugging equation (3) into the distribution of w given in equation (10), deriving the density f_{if}^{w}, and then performing the integration in equation (11) numerically, this approach makes the estimation of the demand model somewhat slow because in every iteration the integral in equation (11) needs to be computed. Although we can cope with these issues (see our estimates for normally distributed search costs in Section 4.2), 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 (11) 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 (11).

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

$$F_{if}^{c}(c) = \frac{1 - \exp\left(-\exp\left(-H_0^{-1}(c) - \mu_{if}\right)\right)}{1 - \exp(-\exp(-H_0^{-1}(c)))}, \tag{12}$$

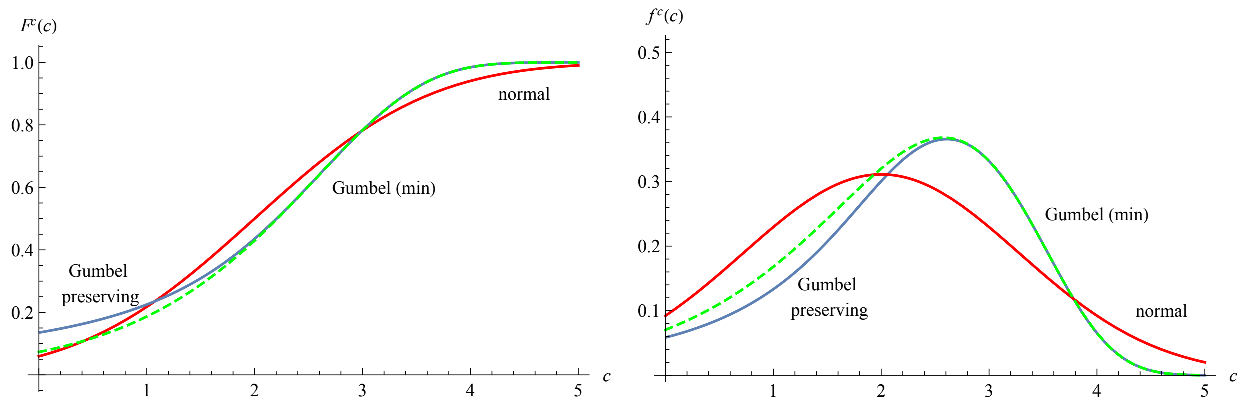
where μ_{if} is a consumer-firm specific location parameter of the search cost distribution, the conditional CDF of w_{if} is given by a Gumbel distribution with location parameter $\delta_{if} - \mu_{if}$, that is,

$$F_{if}^{w}(z) = \exp\bigl(-\exp\left(-(z - (\delta_{if} - \mu_{if}))\right)\bigr). \tag{13}$$

As a result, the conditional probability that individual i purchases product j is equal to

$$s_{ij} = \frac{\exp\left(\delta_{ij} - \mu_{if}\right)}{1 + \sum_{k=1}^{J} \exp\left(\delta_{ik} - \mu_{ig}\right)}, \tag{14}$$

The proof of this proposition, which can be found in Appendix B, builds on the idea that according to equation (10) there is a one-to-one relationship between the search cost distribution and the distribution of the random variable w that determines the buying probabilities. Proposition 1 then shows how an appropriate assumption on the search cost CDF produces closed-form expressions for the conditional purchase probabilities, which makes the estimation of the search model of similar difficulty as most standard discrete choice models of demand. Since the search cost distribution in equation (12) ensures that w is also Gumbel distributed (as is utility), we refer to this distribution as a *Gumbel preserving search cost distribution*.



(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 functional form, that is, $\mu_{if} = \log\left[1 + \exp\left(t'_{if}\lambda\right)\right]$, where the vector t_{if} may include a constant as well as search cost shifters that are consumer and/or firm specific, such as the distance from the household to the seller, the household's income and other demographics, and a standard normal distributed random constant, and λ is a vector of search cost parameters. Given this, it is straightforward to verify that the Gumbel preserving search cost distribution given in equation (12) 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.[15] Finally, we observe that, on the positive support, the distribution given by equation (12) has a shape relatively similar to the normal distribution. To see this, we plot in Figure 1 the Gumbel preserving CDF and the corresponding density for $\mu_{if} = 2$. Also shown is a normal distribution and density with mean $\mu_{if} = 2$ and variance set equal to 1.64 (in red), which is the variance of the Gumbel preserving distribution. Note that the two distributions are relatively similar on the positive real line.[16] The dashed green curves in Figure 1 are for a Gumbel distribution (for the minimum) with location parameter $\mu_{if} = 2 + \gamma = 2.577$ and scale parameter set to one, so the mean and variance correspond to the other two distributions.[17] This particular distribution is useful as a comparison because it is closed form and has a shape similar to the Gumbel preserving distribution.

Finally, to compute the market shares, denoted s_j, we aggregate the conditional individual buying probabilities that product j is purchased. Denoting by τ_i the vector of all consumer-specific variables in s_{ij}

[15]The estimated probability of the atom according to our estimates presented in Section 4.2 is rather small: it has mean 0.017 (with standard deviation 0.089) and median 0.000 across consumers and dealers.

[16]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.

[17]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 $1 - \exp(-\exp((z - m)/b))$.

(i.e., income y_i and other demographic characteristics that are contained in D_i and t_{if}, as well as the standard normal draws contained in v_i and t_{if}) with corresponding CDF F_τ, the probability that product j is purchased conditional on product characteristics is the integral

$$s_j = \int s_{ij} dF_\tau(\tau_i). \tag{15}$$

2.5 Search probabilities

As we explain in Section 3 when discussing estimation and identification, we supplement aggregate data with individual-specific search data. In particular, we use data on the shares of consumers not searching as well as searching once, where we aggregate these shares over various observed demographic variables. Below we present the expressions for consumer i's search probabilities that are needed to obtain predicted shares (see Appendix C for further details). Conditional on the δ_{if}'s and the search cost covariates, the probability that consumer i does not search at all is

$$\pi_{i0} = \int_{-\infty}^{\infty} \prod_{k \neq 0} F_{ik}^r(z) f(z) dz \tag{16}$$

and the conditional probability that consumer i searches once is

$$\pi_{i1} = \sum_f \int_{-\infty}^{\infty} \left(F_{if}^w(y) - F_{if}^r(y) \right) F_{i,-f}^r(y) f(y) dy + \sum_f \int_{-\infty}^{\infty} F_{i,-f}^r(x) F(x) f_{if}^w(x) dx. \tag{17}$$

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

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 calculating elasticities, profits, and markups implied by our estimates, as well as for several of the counterfactual simulations that are part of Section 5. 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 and the search behavior of consumers. Let $\mathbf{p}^*$ denote the vector of Nash equilibrium prices. Assuming a pure strategy equilibrium exists, any product j should have a price that satisfies the first order condition[18]

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

In computing these first order conditions we note that consumers' search behavior is not affected by deviation prices but purchasing decisions are. This means that the reservation values in Lemma 1 do not respond to price deviations. This is a distinctive feature of consumer search models and implies that firms have

[18]Note that what makes this search model different from a Varian-type of search model (Varian, 1980) is that products are differentiated. This distinction is important because Wolinsky-type search models (Wolinsky, 1986) usually have pure-strategy equilibria, while Varian-type search models usually have mixed-strategy equilibria.

an incentive to hold up consumers that visit them, which, reminiscent of the Diamond paradox, leads to equilibrium prices that are higher than when price deviations are observed, as in full information models (see Armstrong, 2017; Choi et al., 2018; Haan et al., 2018). Details on the computation of the market share derivatives for our search model are given in Appendix D. To obtain the price-cost markups for each product we can rewrite the first order conditions as $\mathbf{p}^* - \mathbf{mc} = \boldsymbol{\Delta}(\mathbf{p}^*)^{-1}\mathbf{s}(\mathbf{p}^*)$, where the element of $\boldsymbol{\Delta}(\mathbf{p}^*)$ in row j column r is denoted by Δ_{jr} and $\Delta_{jr} = -\partial s_r/\partial p_j$ if r and j are produced by the same firm and zero otherwise.

3 Estimation

In this section we present our methodology to estimate the theoretical search model discussed in Section 2. Proposition 1 shows that for the Gumbel preserving distribution given by equation (12), the expression for the buying probabilities given by equation (14) is similar to that of a BLP-type full information model, except that search costs enter the expression through μ_{if}. However, this expression also shows that the part that reflects the observed part of utility (δ_{ij}) is not separately identified from the part that reflects search costs (μ_{if}) since both δ_{ij} and μ_{if} enter the buying probabilities additively. One way to achieve identification of search costs is to assume that variables that enter the search cost specification do not also enter the utility specification. For instance, one could assume that distance from a consumer to the nearest dealer affects search costs, but does not directly affect utility. As we discuss in more detail in Section 3.3, an advantage of excluding search cost covariates from utility is that the model can be estimated using aggregate data, as in BLP. However, while it might be possible to argue that this restriction holds for search cost shifters such as distance, it also applies to the search cost constant, i.e., a search cost constant term is not separately identified from the utility constant. This means that the level of search costs is not identified, but only the variation in search costs that is due to the excluded search cost shifters. We show that a solution to these identification issues is to use individual-specific search data. As we will explain in more detail in Section 3.3, even though we can only identify the combined effect of common shifters using purchase data, variation in observed search decisions allows us to separate the combined effect into an effect due to search costs and an effect due to preferences.

We will incorporate information on search through moments based on survey data. Our estimation procedure closely resembles Berry et al. (2004), who use a combination of micro and aggregate data to estimate a differentiated products demand system. However, whereas Berry et al. (2004) 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 use for estimation in more detail. We finish with an informal discussion of the identification of the key parameters of the model.

3.1 Outline of the estimation procedure

Our goal is to estimate the parameters of the model, which include the mean utility parameters $\theta_1 = (\alpha, \beta)$ and the utility parameters that capture consumer heterogeneity as well as search cost parameters $\theta_2 = (\tilde{\alpha}, \Pi, \Sigma, \lambda)$. Following standard practice in the literature (see, e.g., Nevo, 2001), we will refer to the components of θ_2 as the nonlinear parameters. As shown by Berry et al. (2004) in the context of a full information model, micro data allows for estimation of choice-specific constants that capture product j's mean utility $\bar{\delta}_j$, so one approach is to first estimate the vector $(\bar{\delta}, \theta_2)$ followed by using the estimated $\bar{\delta}_j$'s to estimate the mean utility parameters θ_1 in a second step (see also Goolsbee and Petrin, 2004; Train and Winston, 2007). Alternatively, if we have a set of instruments $\mathbf{z}$ that satisfy the conditional moment restriction $E[\xi|\mathbf{z}] = 0$, we could estimate (θ_1, θ_2) directly instead of $(\bar{\delta}, \theta_2)$. In what follows, we use the two-step procedure. Even though we lose efficiency by not imposing such moment restrictions on ξ when estimating θ_2, the main advantage of using the two-step procedure is that we do not lose consistency if the restrictions are not correct. This is especially important in our setting because the unobserved quality variable ξ is not only expected to be correlated with prices, but also potentially correlated with distances from consumers to sellers. This means that the joint estimation of θ_1 and θ_2 requires an instrument for distance, which may be difficult to obtain in practice.[19]

We use aggregate (product-level) data on sales, prices, and other product characteristics, combined with micro (individual-specific) data on purchases and search behavior. In the first step of the procedure, we estimate $(\bar{\delta}_j, \theta_2)$ by generalized method of moments (GMM) using three sets of moments. The first set of moments relates demographic information to buying decisions and is useful for estimating the parameters that capture consumer heterogeneity in the utility function ($\tilde{\alpha}$, Π, and Σ). The second set of moments relates demographic information to search decisions, which allows us to estimate the search parameters of the model (λ). We estimate $\bar{\delta}_j$ by using a third set of moments that match the observed market shares to the model's predicted market shares. Finally, in a second step we use the $\bar{\delta}_j$'s estimated in the first step to estimate α and β by two-stage least squares, using cost shifters as instruments for price.

3.2 Moments and GMM estimation

In this section, we describe how we compute the moments we use for estimation. We use aggregate and micro moments. As shown by Berry et al. (2004), the aggregate moments ($G^3(\cdot)$ below) can be used to concentrate out $\bar{\delta}$ from the GMM objective function, which decreases the dimensionality of the GMM minimization problem, since we then only have to search over the nonlinear parameters θ_2. Specifically, for given θ_2, we can obtain $\bar{\delta}$ as the vector that sets the difference between the observed market shares and the predicted market shares equal to zero. For this we use the result in Berry (1994) and BLP that $\bar{\delta}$ can be computed as the unique fixed point of a contraction mapping. The fact that this mapping is also a contraction in our

[19] Since dealers may prefer to locate in areas where there is most demand for their cars, distances from consumers to car dealers are potentially correlated with unobserved quality differences captured by ξ. In an earlier version of this paper we did jointly estimate θ_1 and θ_2 using variables that relate to the cost of operating a dealership (variation in property values and local taxes) as instruments for distance from consumers to dealerships.

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).[20] Since the predicted market shares cannot be computed directly, we follow BLP and most of the subsequent literature in estimating these using Monte Carlo methods. Specifically, we jointly draw income y_i, the demographic characteristics used in D_i and t_{if}, and the standard normal variables used in v_i and t_{if} of, say, ns consumers from regional demographic data provided by Statistics Netherlands (see Section 4.1 for details), and then compute s_{ij} for each simulated consumer $i = 1, \ldots, ns$. This means in practice we compute $\bar{\delta}$ as the vector that sets the difference between the observed market shares and the Monte Carlo estimator of the predicted market shares, i.e.,

$$G^3(\theta_2, \bar{\delta}) = s_j^{\text{obs}} - \frac{1}{ns} \sum_{i=1}^{ns} s_{ij}, \tag{18}$$

equal to zero, where s_{ij} in this moment is given by equation (14).

In order to describe the computation of the moments that relate demographic information to buying decisions in the micro data, it is useful to introduce some notation. Suppose in the micro data we observe the demographic characteristics and purchase decisions of N consumers. We follow Petrin (2002) in using moments based on the following type of conditional expectations:

$$E\left[\mathbb{1}\{a_i \in T\} | D_i^g \in R_k\right], \quad k = 1, 2 \ \text{ where } \ \mathcal{R} = \{R_1, R_2\}. \tag{19}$$

In this expression, $\mathbb{1}\{a_i \in T\}$ is an indicator for the event that consumer $i \in \{1, \ldots, N\}$ makes choice $a_i \in \{0, 1, \ldots, J\}$ from a certain group of products T (e.g., new cars, large cars, cars from luxury brands, etc.), and D_i^g is demographic characteristic g (i.e., income, age, etc.), which is partitioned in two subsets according to $\mathcal{R}$. For the specific moments we use in our application we refer to Table A2 in the Appendix.

Let $R \in \mathcal{R}$. Since we observe the choice of each consumer $i \in \{1, \ldots, N\}$ in the micro data, the moments boil down to aggregation of the choice a_{iT} of i regarding T over those consumers i whose demographic characteristic satisfy that $D_i^g \in R$, where a_{iT} equals 1 if $a_i \in T$ and 0 otherwise. Note that a_{iT} is a Bernoulli random variable with success probability

$$\begin{aligned} s_{iT}\left(\theta_2, \bar{\delta}\right) &= \Pr\left(a_{iT} = 1 | p, \mathbf{x}, \theta_2, \bar{\delta}, D_i, v_i, y_i, t_{if}\right) = \Pr\left(a_i \in T | p, \mathbf{x}, \theta_2, \bar{\delta}, D_i, v_i, y_i, t_{if}\right) \\ &= \sum_{j \in T} \Pr\left(a_i = j | p, \mathbf{x}, \theta_2, \bar{\delta}, D_i, v_i, y_i, t_{if}\right) = \sum_{j \in T} s_{ij} \end{aligned}$$

and independent across i conditional on $\left(p, \mathbf{x}, \theta_2, \bar{\delta}, D_i, v_i, y_i, t_{if}\right)$, so that $E\left[a_{iT} | p, \mathbf{x}, \theta_2, \bar{\delta}, D_i, v_i, y_i, t_{if}\right] = s_{iT}\left(\theta_2, \bar{\delta}\right)$. Aggregation over i such that $D_i^g \in R$ yields the moment

$$G_{gR}^1\left(\theta_2, \bar{\delta}\right) = \frac{1}{N_R} \sum_{\substack{i=1 \\ D_i^g \in R}}^{N} \left(a_{iT} - s_{iT}\left(\theta_2, \bar{\delta}\right)\right), \tag{20}$$

[20] 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 et al., 2012).

where N_R is the number of consumers i for which $D_i^g \in R$. This is just the sample counterpart of the moment condition $E\left[a_{iT} - s_{iT}\left(\theta_2, \bar{\delta}\right)|p, \mathbf{x}, \theta_2, \bar{\delta}, D_i, v_i, y_i, t_{if}\right] = 0$ over the sample of consumers i with $D_i^g \in R$. Note that while we aggregate over i we use standard random normal draws for the v_i's and the random search cost constant in t_{if}.

The type of moments we use to relate demographic information to search decisions is similar to the type given in equation (19), but with the number of searches as the choice variable instead of the type of car bought. This gives the moment

$$G_{gR}^2\left(\theta_2, \bar{\delta}\right) = \frac{1}{N_R} \sum_{\substack{i=1 \\ D_i^g \in R}}^{N} \left(a_{iS} - \pi_{iS}\left(\theta_2, \bar{\delta}\right)\right), \tag{21}$$

where a_{iS} equals 1 if consumer i is observed to search S times in the micro data and 0 otherwise, and π_{iS} is the predicted probability that consumer i searches S times (see Section 2.5 for expressions for the probability of not searching π_{i0} and the probability of searching once π_{i1}).

We stack $G_{gR}^1(\cdot)$ and $G_{gR}^2(\cdot)$ into $G(\cdot)$, so the GMM estimator of θ_2 is then

$$\widehat{\theta}_2 = \arg\min_{\theta_2} G(\cdot)'\Xi G(\cdot),$$

where $G(\cdot)$ is the column vector of moments and Ξ is a weighting matrix.[21] In our application we follow Petrin (2002) and use weights that are determined using the inverse of the variance of the survey moments (see also Appendix B.1 of Petrin, 2002). In a second step, assuming θ_2 known, we can obtain θ_1 as the linear IV estimator

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

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

3.3 Identification

In this section we informally discuss identification of the model parameters. Identification is based on the assumption that we observe aggregate market share data and search and purchase data for a large number of individuals. It is important to reiterate, 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 parameters θ_1 (that is, the components of $\bar{\delta}$) and the nonlinear parameters in θ_2 (that is, the price/income coefficient $\tilde{\alpha}$, the observed and unobserved heterogeneity parameters (Π, Σ), and the search cost parameters λ) based on parametric random coefficient discrete choice identification arguments (see also Berry et al., 2004).

Since in our application the number of products is relatively large, it is not practical to estimate the

[21] To obtain estimates for θ_2, we use the `fminunc` function in Matlab, which finds the minimum of an unconstrained multivariable function using a Quasi-Newton algorithm that uses the BFGS Quasi-Newton method with a cubic line search procedure.

product-specific mean utilities $\bar{\delta}$ directly. Therefore, following Berry et al. (2004), we compute them as a function of θ_2 by using the observed market shares, as described in Section 3.2. Given the estimates of $\bar{\delta}$, the identification of the mean utility parameters in θ_1 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 $\mathbf{z}$, that is, the conditional moment restrictions $E[\xi_j|\mathbf{z}] = 0$, $j = 1, \ldots, J$ 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. We use cost shifters as instruments for price (see Section 4.2 for details); according to Armstrong (2016) such instruments result in more precise estimates of the price coefficient than BLP instruments, especially when the number of products per market is large and the number of markets is small (as in our application).

The moment conditions we use for estimation of the model are tightly linked to the nonlinear parameters θ_2. The first set of moment conditions, given by G^1_{gR} in equation (20), relates purchase probabilities to demographics. We use several moment conditions for the interaction of household demographics and product attributes. For instance, in our application we include moment conditions for the interaction of price and income to identify the price/income coefficient and use moment conditions for the interaction of larger vehicles and whether children are living in the household to identify random coefficients on size and size interacted with a children indicator. The second set of moment conditions, given by G^2_{gR} in equation (21), relates search probabilities to demographics. Here we use a separate moment condition for each search cost covariate. Because the observed share of consumers that searches more than twice is very small in our data, we only use expressions for the probability of not searching and searching once, as given by equations (16) and (17). However, expressions can be obtained for a higher number of searches, and the use of such moments may help identifying the search cost distribution non-parametrically.[22] Relatedly, it is also possible to derive expressions for the probability of searching conditional on the brand that was purchased. Although such moments will help identification of substitution patterns between brands, the number of purchases in the survey for an individual brand is too low to implement this with our data.

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. These effects cannot be separately identified without search data with Gumbel preserving distributed search costs. Here we provide somewhat more formal arguments to show 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

[22] Hong and Shum (2006) show that non-parametric estimation of the search cost distribution can be achieved in the context of a homogenous non-sequential search model (see also Moraga-González et al., 2013). Note that non-parametric identification of the search cost distribution when preferences are unknown (as in our model) is still an open issue in the literature.

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 probability of not searching π_{i0}, which is given by equation (16).[23] Rearranging equation (9) gives $F_{if}^r = (F_{if}^w - F_{if})/(1 - F_{if})$, so the distribution of reservation utilities that appears in equation (16) can be written as

$$F_{if}^r(z) = \frac{\exp(-\exp(\delta_{if} - \mu_{if} - z)) - \exp(-\exp(\delta_{if} - z))}{1 - \exp(-\exp(\delta_{if} - z))},$$

which means the expression $\prod_{k \neq 0} F_{ik}^r(z)$ in the integral of equation (16) contains both $\delta_{if} - \mu_{if}$ and δ_{if} for $f = 1, \ldots, F$. As explained above, the expressions $\delta_{ij} - \mu_{if}$ are identified from individual-level purchase data, which, since $\delta_{if} - \mu_{if} = \log\left(\sum_{h \in G_f} \exp(\delta_{ih} - \mu_{if})\right)$, also means $\delta_{if} - \mu_{if}$ are identified. Therefore, in the probability π_{i0} the only unknowns left are the utility parameters involved in the δ_{if}'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 $1 - \pi_{i0} - \pi_{i1}$, which allows us to separately identify the distance coefficient in the search cost specification and the distance coefficient 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 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.

Note that throughout the identification section we have assumed that consumers are searching sequentially for a good product fit. A caveat on identification is that the individual-level search data may be consistent with alternative models of search. For instance, if consumers are searching non-sequentially, that is, if they

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

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

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

determine before searching how many (and which) dealers to visit, then the search probabilities in Section 2.5 will be different, which may have implications for the parameter estimates.[24]

4 Application

In this section we apply our model to the Dutch market for new cars. We first briefly discuss the aggregate and individual level data we use for estimation of the model and provide some background information on how consumers search in this market.[25] In the remainder of Section 4, we report the estimation results of our search model as well as estimates of demand elasticities and markups.

4.1 Data

Aggregate data

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. The data on car 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. Sales data are obtained from BOVAG, which is the Dutch trade organization for car dealers. All prices are list (post-tax) prices, deflated to 2006 euros using the Consumer Price Index. We regard list prices as equilibrium prices in the model—list prices are relatively easy to figure out in the Netherlands without visiting a dealership (e.g., by visiting a website, making a phone call, etc.), but to capture that consumers have to visit a dealer to ascertain the price of a specific car, we follow most of the theoretical search literature (e.g., Diamond, 1971; Burdett and Judd, 1983; Wolinsky, 1986) in assuming that deviations from equilibrium prices are only observed when visiting a dealership. Even though this does not affect estimation of the demand side of the model (since consumers have rational expectations and therefore base their search decisions on correct beliefs about equilibrium prices), it does affect calculation of elasticities and markups, and, as shown later, has important implications for counterfactuals in which the dealership network changes. Because we do not have information on transaction prices, we cannot model dealers' decision-making. This implies that we are not modeling other potential sources of price variation that may induce consumers to search in the car market, such as price dispersion across dealers of the same brand. Such dispersion may arise because of dealer cost heterogeneity and differences in bargaining skills. However, as pointed out by Nurski and Verboven (2016), in Europe manufacturers control retail prices either directly or indirectly via price restraints such as non-linear pricing schemes and bonus systems. Moreover, most dealers in the Netherlands have very few cars in stock due to space limitations, which means there is typically less incentive for dealers to offer prices different from the manufacturer price in comparison to for

[24] In an earlier version of this paper we assumed non-sequential search instead of sequential search (Moraga-González et al., 2015)—although the parametric search specifications we used led to very similar purchase probabilities, the search probabilities were different. Both De los Santos et al. (2012) and Honka and Chintagunta (2017) provide methods to test for the search method consumers are using (sequential vs. non-sequential), but because these tests assume consumers are searching for prices only, they do not directly apply here.

[25] We refer to Appendix F for a more detailed overview of the data.

instance the United States, where it is more common to have a large number of cars in stock. By ignoring these additional sources of price variation, we are implicitly assuming that the main reason for visiting dealerships is to learn more about the products and (brand-level) price deviations.

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: 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	19,951	0.749	0.788	7.184	0.308	14.696	11.737	0.080
2005	233	457,897	20,541	0.727	0.794	7.270	0.301	14.861	10.987	0.096
2006	231	475,636	20,367	0.715	0.804	7.272	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	19,917	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. HP/Weight is $\times 10$.

Table 1 gives the sales weighted means for the main variables we use in our analysis. The table shows that the number of models has increased, prices have mostly gone up, and that the share of European cars has decreased over the sampling period. Moreover, cars have become more powerful and more fuel efficient. Finally, the share of luxury brand cars, where luxury 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.

For estimation of the mean utilities we need to match predicted market shares to observed market shares, using the moment in equation (18). The predicted market shares are estimated using Monte Carlo methods by sampling from demographic data at the neighborhood level provided by Statistics Netherlands. In addition to demographic data we have information on the exact location of each neighborhood as well as the addresses of all dealerships in the Netherlands, which we use to compute the Euclidean distance from the center of the neighborhood to the closest dealer of each brand. 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: more than 89% of all households live within 7 kilometers from an Opel dealer. Rover has the lowest percentage of households within 7 kilometers: only 5% of households is within easy reach.

Table 2: Descriptive statistics for distances

Brand	Number of dealerships	Weighted average distance	Percentage of households within 7 km	Brand	Number of dealerships	Weighted average distance	Percentage of households within 7 km
Alfa Romeo	75	7.96	57.90	Mazda	121	5.57	73.77
Audi	161	4.68	82.32	Mercedes-Benz	83	6.58	66.32
BMW	57	8.35	55.13	Mini	37	11.41	41.94
Cadillac	15	18.96	22.92	Mitsubishi	108	5.57	72.55
Chevrolet/Daewoo	137	5.00	76.98	Nissan	114	6.02	68.19
Chrysler/Dodge	32	12.51	37.82	Opel	233	3.55	89.40
Citroën	162	4.40	82.35	Peugeot	187	4.14	84.17
Dacia	98	7.25	67.69	Porsche	8	25.70	12.88
Daihatsu	99	6.23	68.49	Renault	196	4.18	82.73
Fiat	142	4.86	79.72	Rover	7	34.16	5.12
Ford	233	3.66	89.13	Saab	37	11.46	35.78
Honda	65	7.99	57.26	Seat	127	6.08	71.18
Hyundai	138	5.08	77.63	Skoda	97	6.18	68.35
Jaguar	16	18.32	24.99	Smart	21	13.65	30.40
Jeep	43	10.13	49.40	Subaru	45	10.96	38.60
Kia	115	5.74	71.39	Suzuki	124	5.00	78.02
Lancia	50	11.12	49.24	Toyota	141	4.69	80.92
Land Rover	20	14.43	27.63	Volkswagen	188	4.05	86.07
Lexus	13	19.55	24.63	Volvo	114	5.33	75.94

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

Micro-level data

The micro-level data on individual searches and purchases used for the moments in equations (20) and (21) is obtained from Dutch survey agency TNS NIPO. For 2,530 respondents, we have information about the make and model of their current car and the year in which the car was bought, as well as information on how consumers search in this market, including the brands of the dealerships visited before buying the car, and for which brands a test drive was made at the dealer. The respondents also answered questions about their household income, household size, age, whether there are children living in the household, and zip code.

Figure 2 gives a histogram of the number of dealers visited of the brands in our sample (conditional on searching). The average number of dealers of different brands visited for new car purchases is 2, which is slightly below the average of 3 dealers found by Moorthy et al. (1997) for the United States in the early nineties. Although 47 percent of respondents who actively searched visited one brand only, the distribution is positively skewed with some consumers visiting dealers of as many as 12 different brands.

Consumers visit car dealerships for various reasons, such as learning more about the characteristics of cars or to bargain for a lower price (or higher price for the car they trade in). 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% of the dealer visits for those that bought a new car, a test drive was involved. Moreover, among those who visited one or more dealers to shop for a new car, approximately 69% 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

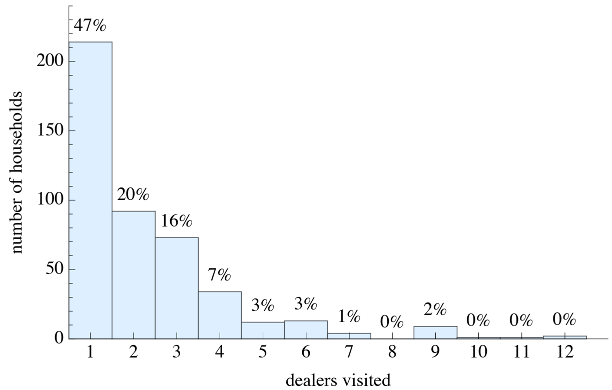


Figure 2: Histogram number of dealers visited for new car purchases

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.[26]

4.2 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 preference heterogeneity and search cost parameters and allows us to separately identify the impact of 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 determinants of mean utility, 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, non-parametric 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. This also

[26] According to data from the survey, of the respondents who visited dealerships to buy a new car, about a third made multiple trips to a dealer of the brand that was ultimately purchased. While price shopping could explain this behavior, there are other potential explanations for this observation. For example, consumers may visit several dealers of the same brand because they are initially also interested in used cars, because they may be open to buy a showroom model, or 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 which cars are on inventory.

means a search cost constant is not separately identified from the utility constant; without search data we can only identify how search costs move with distance (assuming distance is not part of utility), but not the level of search costs.

Table 3 gives the parameter estimates for the simplest version of the model, the conditional logit model. We use a simplified version of equation (1), i.e., we assume all consumers share the same price coefficient and do not allow for preference heterogeneity for the other product attributes either, so the indirect utility function is given by $u_{ij} = \alpha p_j + x_j' \beta + \xi_j + \varepsilon_{ij}$. 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. 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 Armstrong (2016) we use proper cost shifters as excluded instruments. In particular, as in Reynaert and Verboven (2014) we include unit labor costs and steel prices interacted with weight (both normalized by country-specific producer price indices). Furthermore, to allow for economies of scale in production, we include model-specific annual production as a cost shifter as well.[27] Since production is distributed worldwide and the Netherlands is a relatively small country, we expect this variable to be exogenous in the demand and supply model for the Netherlands.

All parameter estimates in column (A) of Table 3 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 next three columns of Table 3 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.[28] Next we use the distances to the nearest

[27] The estimated coefficients in the first-stage price regression (heteroskedasticity robust standard errors within parentheses) are 9.568 (2.614) for ULC (normalized by PPI), 0.390 (0.090) for (CRU) × weight (divided by 10,000 and normalized by PPI), and −5.514 (1.870) for production (measured in millions of units), with a Cragg-Donald Wald F-stat for the excluded instruments of 15.68.

[28] When generating these random draws, instead of resorting to the commonly used computer-generated *pseudo*-random numbers, we employ *quasi*-random numbers, which are designed to mimic a given distribution better than pseudo-random draws. As a result, we expect to obtain more precise estimates of the underlying probabilities. Specifically, we use quasi-random draws constructed from a $(0, 2, 47)$-net in base 47, which contains $47^2 = 2{,}209$ draws (for more details see e.g., Sándor and András, 2004).

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 $(\mathbf{z}'\mathbf{z})^{-1}$ as the weighting matrix, where $\mathbf{z}$ contains the instruments. The results shown in column (B) of Table 3 are obtained using the Gumbel preserving search cost distribution specified in equation (12), and show that search costs are positively related to distance and precisely estimated. Most 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 and higher average markups.

Table 3: Estimation results conditional/nested logit model

	Full information		Search						Full information	
	Logit Demand (A)		Logit Demand (B)		Logit Demand (C)		Logit Demand (D)		Nested Logit Demand (E)	
PREFERENCE PARAMETERS										
constant	-20.728	(1.978)	-17.441	(1.792)	-17.002	(1.785)	-17.230	(1.788)	-17.657	(1.638)
HP/weight	4.706	(1.294)	3.905	(1.145)	3.899	(1.144)	3.895	(1.143)	3.535	(1.017)
non-European	-1.237	(0.187)	-1.064	(0.166)	-1.062	(0.165)	-1.060	(0.166)	-1.132	(0.148)
cruise control	0.508	(0.142)	0.260	(0.131)	0.259	(0.130)	0.259	(0.130)	0.513	(0.113)
fuel efficiency	2.627	(0.354)	2.500	(0.305)	2.494	(0.304)	2.496	(0.304)	2.182	(0.307)
size	13.453	(2.362)	10.591	(2.109)	10.573	(2.107)	10.564	(2.106)	11.012	(1.920)
luxury brand	1.923	(0.403)	1.763	(0.346)	1.758	(0.345)	1.761	(0.345)	1.193	(0.341)
price	-0.168	(0.035)	-0.130	(0.031)	-0.130	(0.031)	-0.130	(0.031)	-0.132	(0.028)
$\log(s_{j\|g})$	—		—		—		—		0.316	(0.047)
SEARCH COST PARAMETERS										
distance	—		0.183	(0.040)	0.143	(0.028)	0.195	(0.041)	—	
Avg. price elast.	-4.906		-3.806		-3.801		-3.796		-5.408	
Avg. markup	3,713		4,813		4,839		4,833		4,713	
Search cost distr.	–		Gumbel preserving		Normal		Gumbel (min)		–	

Notes: The number of observations is 1,382. Heteroskedasticity robust standard errors are in parenthesis. The number of simulated consumers used for the estimation of specifications (B)-(D) is 2,209. The linear part of utility includes car segment fixed effects. Instruments for price include unit labor costs (normalized by PPI), steel prices (normalized by PPI) interacted with weight, and annual production. We include average distance to dealerships as an instrument for specifications (B)-(D). The instrument for $\log(s_{j|g})$ in specification (E) is the number of products within nest. The Cragg-Donald Wald F test of excluded instruments for specification (A) is 15.68.

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 (11). 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.[29] 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

[29]The variance we use for the normal distribution is normalized to 1.64 (the variance of the Gumbel preserving search cost distribution). If we instead normalize the variance of the normal distribution to 1, the estimate of distance goes down to 0.110, whereas all the preference parameters are virtually unchanged in comparison to those reported in column C of Table 3.

similar in shape to the Gumbel preserving search cost distribution (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 Gumbel preserving distribution, a comparison of the estimation results with search to those without search shows that the price coefficient goes down in absolute value, which suggests that at least in the conditional logit model, ignoring search frictions may result in an overestimation of consumer price sensitivity. A possible explanation for this result is misspecification bias—higher priced car models tend to be sold by dealerships that are on average located further away from consumers, so by not controlling for distance-related search costs it seems consumers are more price sensitive then they really are.

In the last column of Table 3 we go back to the full information model, but now estimate it using a nested logit formulation, where the nests are the different brands. The purchase probabilities in the nested logit model can be written as the product of the probability of choosing a nest and the probability of choosing a product conditional on having chosen the nest (see, e.g., Anderson and De Palma, 1992). As shown in equation (6), the probability of buying product j in our search model is $s_{ij} = P_{if} P_{ij|f}$, which has a similar structure if the nests are the products of the different firms. However, despite the similarity in structure, the nested logit model does not allow for search costs. Moreover, in the nested logit model a firm can only be part of one nest, while in our search model we allow for a more flexible structure in which nests can be overlapping. The results in column (E) are obtained by regressing $\log(s_j) - \log(s_0)$ on product characteristics, prices, and $\log(s_{j|f})$ (the log of within-nest market shares), using the same cost shifters as before to instrument for price as well as the number of car models per brand as an instrument for $\log(s_{j|f})$. Notice that despite the price coefficient and average markup being similar to those in the search specifications, the estimated average own-price elasticity is larger in absolute value.[30]

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. We use the survey data 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 and those who buy without visiting a dealership, which leaves us with 2,014 observations.[31] 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,

[30] In the nested logit model the own-price elasticity is given by $\alpha/(1-\sigma) \cdot p_j \left(1 - \sigma \cdot s_{j|g} - (1-\sigma) \cdot s_j\right)$, which shows that the own-price elasticity now also depends on the nest parameter σ. Firms set prices such that the effect of price changes on other own products are internalized. In this particular case, the nests correspond to the brands, which means that the nesting structure is subsumed by the product ownership matrix that is used to determine markups. The estimated nest parameter in column (E) is 0.316 which suggests that within-brand substitution is important. Within-brand substitution reinforces the positive effect on equilibrium prices of internalizing the effect of price changes on own products (since consumers are likely to substitute towards other car models of the same brand), which gives firms additional market power and in this case leads to higher markups than in the full information model, despite more elastic demand.

[31] Our structural model does not allow consumers to make a purchase without having visited any dealers.

which equals the percentage of households in the Netherlands that bought a new car in 2008.[32]

The demand side estimates for the complete model are based on the utility function in equation (1) and the Gumbel preserving search cost distribution that is derived from Gumbel distributed w's. We use the same car attributes as those shown in Table 3 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 the constant, size, and luxury brand dummy to differ across consumers by estimating a variance term for these attributes.[33] We also interact consumer demographics with some of the attributes to account for observed consumer heterogeneity. Specifically, to allow substitution to the outside good to depend on demographics, we interact the constant with an urban dummy, a dummy for whether there are kids present in the household, and a senior dummy, while we interact size with the kids dummy to capture that families with children might prefer larger cars.[34] As explained in Section 2.1, we allow for heterogeneity in the price parameter by making it inversely proportional to consumer i's household income y_i, that is, $\alpha_i = \tilde{\alpha}/y_i$. 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 dummy for whether the head of household is senior and whether there are kids living in the household. To allow for the effect of distance to differ according to income, we also interact household income with distance. In addition, we estimate the mean and standard deviation of a normal distributed random search cost constant.

For the survey respondents we observe income, whether there are children living in the household, age, as well as zip code, so we can include this information directly when matching the search and purchase probabilities from the survey data to the their predicted counterparts from the model. The predicted probabilities depend on choice-specific constants (the linear part of utility $\bar{\delta}_j$)—following the estimation procedure outlined in Section 3, to solve for these we use the aggregate data to invert the market shares for $\bar{\delta}_j$. Since we allow for random coefficients, predicted market shares are estimated by Monte Carlo by drawing from distributions of demographic characteristics and the standard normal distribution.

Table 4 builds up the model starting from a full information model in which distance to dealerships only enters the utility specification. Next we present the estimates from an intermediate specification which adds a simple search costs specification in which search costs only depend on distance and a constant, and finally the full specification in which we add additional search cost shifters and interactions. The nonlinear parameters, which include the price/income coefficient, the random coefficients, and the search cost parameters in the search specifications, are obtained using the GMM approach outlined in Section 3. The mean utility parameters are obtained in a second step by regressing mean utility $\bar{\delta}_j$ on car attributes and

[32]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.

[33]Following BLP and papers that employ the BLP framework (such as Nevo, 2001; Petrin, 2002; Sovinsky Goeree, 2008) we allow for a random constant in the utility function, which is identified because the mean utility of the outside option is normalized to zero and captures unobserved heterogeneity in substitution towards the outside good.

[34]The urban dummy is one if the number of addresses per square kilometer in a neighborhood or zip code exceeds 1,000 (moderately urban to highly urban) and zero otherwise.

Table 4: Estimation results complete model

	Full Information		Search					
	(A)		(B)		(C)		(D)	
PREFERENCE PARAMETERS								
price	-0.044	(0.036)	-0.047	(0.034)	-0.093	(0.035)	-0.093	(0.035)
$\times$1/income	-6.297	(0.449)	-5.137	(0.035)	-3.103	(0.064)	-3.132	(0.064)
HP/weight	4.847	(1.213)	4.791	(1.244)	4.802	(1.265)	4.796	(1.264)
non-European	-1.278	(0.182)	-1.258	(0.183)	-1.257	(0.185)	-1.255	(0.184)
cruise control	0.589	(0.141)	0.506	(0.140)	0.504	(0.141)	0.503	(0.140)
fuel efficiency	2.399	(0.363)	2.550	(0.355)	2.545	(0.354)	2.547	(0.353)
size	12.768	(2.255)	12.405	(2.274)	11.645	(2.309)	11.626	(2.307)
$\times$ normal	2.183	(0.399)	1.018	(0.131)	0.999	(0.159)	1.059	(0.154)
$\times$ kids	6.359	(0.148)	5.373	(0.083)	5.774	(0.107)	5.711	(0.110)
luxury brand	-4.842	(3.513)	-2.053	(0.716)	0.338	(0.796)	0.298	(0.831)
$\times$ normal	5.463	(1.605)	3.698	(0.279)	2.148	(0.407)	2.180	(0.429)
constant	-24.431	(1.910)	-13.608	(1.896)	-13.068	(1.931)	-13.098	(1.929)
$\times$ normal	5.218	(0.320)	0.648	(0.087)	0.984	(0.092)	1.072	(0.092)
$\times$ urban	-0.539	(0.040)	-0.403	(0.011)	-0.440	(0.016)	-0.456	(0.016)
$\times$ kids	-4.784	(0.119)	-3.733	(0.067)	-5.068	(0.083)	-5.011	(0.085)
$\times$ senior	1.296	(0.075)	1.078	(0.013)	0.107	(0.021)	0.104	(0.022)
distance	-0.012	(0.002)	0.015	(0.002)	0.004	(0.005)	0.004	(0.005)
exclusive dealer	—		—		—		0.056	(0.005)
SEARCH COST PARAMETERS								
distance	—		0.032	(0.002)	0.195	(0.065)	0.201	(0.065)
$\times$ log(income)	—		—		-0.048	(0.019)	-0.050	(0.019)
log(income)	—		—		-1.150	(0.141)	-1.161	(0.142)
kids	—		—		-0.956	(0.013)	-0.959	(0.014)
senior	—		—		-1.048	(0.016)	-1.062	(0.016)
constant	—		9.936	(0.028)	13.993	(0.483)	14.094	(0.490)
$\times$ normal	—		4.414	(0.017)	4.269	(0.024)	4.314	(0.026)
Objective function	1,320.8		16,372.5		4,903.3		5,062.3	
Average price elasticity	-5.730		-4.988		-4.983		-4.971	

Notes: The number of observations is 1,382. Heteroskedasticity robust standard errors are in parentheses. The number of (quasi-)random draws used for the aggregate data is 2,209. The number of (quasi-)random draws used for the micro data is 25 per respondent. The linear part of utility includes car segment fixed effects. Instruments for price include unit labor costs (normalized by PPI), steel prices (normalized by PPI) interacted with weight, and annual production. The reported objective function value is $\times$1000.

segment dummies. We allow the price to be correlated with ξ_j, and use two-stage least squares to estimate the mean utility parameters of the model. As before, we use unit labor cost, steel prices interacted with weight, and production as instruments for price. Although not reported, we allow for segment fixed effects in all specifications of Table 4.

In the full information model shown in column (A), distance from the dealership to the consumer only enters the utility function and as such can be considered a transportation cost. The price/income coefficient is very precisely estimated. All else equal, the positive HP/weight parameters suggests consumers prefer more powerful cars. The estimated coefficient of the non-European dummy is negative, 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 of the normal distribution interacted with 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. Furthermore, families with children put positive value on larger cars as well. The mean parameter estimate for the luxury brand dummy is negative and imprecisely 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 interactions of the constant with various demographics suggest there is substantial heterogeneity in how different groups of consumers value the outside option. We allow distance to enter 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.

The remaining columns of Table 4 show estimates for various specifications of the search model. Column (B) gives the estimation results from an intermediate specification in which search costs only depend on distance and a constant. In comparison to the results of the full information model in column (A), the parameter estimate for price/income decreases in magnitude. Moreover, several of the preference parameter interactions decrease in magnitude as well. This is most pronounced for the constant×normal preference parameter—the search-related moments allow us to separate the effect of the constant on preferences and search costs, and according to the estimates substantial variation in the search cost constant is needed to explain observed search behavior. The distance parameter in the search cost specification is positive and precisely estimated. Somewhat surprisingly, the distance parameter in the utility function is now positive—if this parameter would capture distance as a transportation cost, its effect is expected to be negative. However, once we allow for additional search cost shifters (see specification (C) below), the estimate for distance in the utility specification becomes indistinguishable from zero.

Specification (C) adds additional search cost shifters to the model. Most search cost parameters are very precisely estimated. The estimates indicate that search costs are positively related to distance, but that the effect is smaller for households with higher income. Household income, the kids dummy, and the senior dummy are all negatively related to search costs. 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 estimated distance-related search cost parameter corresponds to a mean distance-related search cost of €148 per kilometer, which is similar to the effect of distance found in other studies.[35] For instance, Nurski and Verboven (2016) get an average distance-related travel cost of €112 per kilometer, while Albuquerque and Bronnenberg (2012) obtain very similar travel costs for their logit specification.

Table A2 in the appendix gives the estimated probabilities used for the moments as well as those from the survey data. Most estimated probabilities used as moments are close to the corresponding probabilities

[35] Euro values of the search cost estimates are obtained by normalizing the distance-related search cost parameter by the estimated price coefficients—since we allow for heterogeneity in the price coefficient, we obtain a distribution of search cost in euros.

from the survey data, so all specifications are able to match the probabilities from the survey data relatively well.

4.3 Demand elasticities and markups

Table 5: Demand elasticity estimates

	Toyota Auris	Renault Megane	Opel Zafira	VW Touran	VW Passat	Ford Mondeo	Audi A4	BMW 3-series	Mercedes C-class
Search									
Toyota Auris	-3.0941	0.0117	0.0180	0.0114	0.0275	0.0339	0.0194	0.0176	0.0143
Renault Megane	0.0096	-3.3638	0.0184	0.0116	0.0286	0.0356	0.0202	0.0182	0.0146
Opel Zafira	0.0094	0.0118	-3.9612	0.0117	0.0292	0.0367	0.0210	0.0189	0.0150
Volkswagen Touran	0.0094	0.0117	0.0184	-4.1859	0.4723	0.0362	0.0206	0.0186	0.0148
Volkswagen Passat	0.0095	0.0122	0.0194	0.1979	-3.9948	0.0399	0.0227	0.0202	0.0156
Ford Mondeo	0.0096	0.0124	0.0200	0.0123	0.0327	-3.9974	0.0241	0.0213	0.0162
Audi A4	0.0065	0.0083	0.0137	0.0086	0.0229	0.0299	-4.1993	0.1079	0.0824
BMW 3-series	0.0065	0.0082	0.0135	0.0085	0.0222	0.0288	0.1193	-4.3075	0.0814
Mercedes C-class	0.0062	0.0078	0.0127	0.0081	0.0204	0.0262	0.1068	0.0958	-4.5868
Full information									
Toyota Auris	-4.1243	0.0306	0.0471	0.0311	0.0748	0.0909	0.0183	0.0161	0.0116
Renault Megane	0.0252	-4.4398	0.0496	0.0326	0.0816	0.1017	0.0196	0.0167	0.0116
Opel Zafira	0.0243	0.0310	-5.1161	0.0342	0.0883	0.1128	0.0206	0.0172	0.0117
Volkswagen Touran	0.0240	0.0305	0.0512	-5.2247	0.0861	0.1091	0.0204	0.0172	0.0119
Volkswagen Passat	0.0242	0.0320	0.0555	0.0361	-5.2920	0.1323	0.0235	0.0187	0.0119
Ford Mondeo	0.0240	0.0326	0.0579	0.0374	0.1081	-5.4287	0.0250	0.0194	0.0119
Audi A4	0.0056	0.0072	0.0121	0.0080	0.0221	0.0288	-6.2191	0.6123	0.4350
BMW 3-series	0.0055	0.0069	0.0114	0.0076	0.0197	0.0250	0.6858	-6.3484	0.4544
Mercedes C-class	0.0052	0.0063	0.0102	0.0069	0.0165	0.0202	0.6406	0.5974	-6.6693

Notes: Demand elasticities are calculated for 2008. The table 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. Elasticities for the search model are calculated using estimates from specification (C) in Table 4; those for the full information model are based on specification (A) in Table 4.

Table 5 gives demand elasticity estimates for a selection of car models sold in 2008 for both the search model (using the estimates in column (C) of Table 4) and the full information model (using the estimates in column (A) of the table). For all car models, demand is estimated to be more inelastic in the search model than in the full information model. This means that in this application assuming consumers have full information, while in reality they do not, leads to an overestimation of price sensitivity for most car models. This finding is similar to what we found for the logit estimates reported in Table 3 and suggests that the bias due to misspecification of the full information model persists even when allowing for a richer specification that includes distance as a utility shifter, as in specification (A) of Table 4. The cross-price elasticities show a similar pattern for most models: 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 full information model than in the search model. One exception among the car models in the table relates to the substitution patterns between the Volkswagen Touran and Volkswagen Passat—in the search model consumers are much more likely to switch from a Passat to a Touran than to any of the other models listed when the price of a Passat increases and vice versa. This is not specific to these particular models: for all car models consumers are far more likely to switch within brand than across brands. Since search happens at the brand level, a consumer does not have to pay an additional search cost when considering one of the other models sold by the same brand, making it more likely that a consumer buys within brand following a price increase.

Table 6: Distance elasticity estimates

	Smart	Lexus	Honda	BMW	VW	Ford
Smart	2.7435	-0.0011	-0.0116	-0.0113	-0.0315	-0.0267
Lexus	-0.0007	1.8144	-0.0042	-0.0698	-0.0111	-0.0099
Honda	-0.0023	-0.0013	1.2397	-0.0115	-0.0291	-0.0254
BMW	-0.0012	-0.0117	-0.0061	1.0487	-0.0163	-0.0146
Volkswagen	-0.0024	-0.0013	-0.0111	-0.0116	0.6308	-0.0265
Ford	-0.0024	-0.0014	-0.0113	-0.0122	-0.0310	0.5842
Outside good	-0.0002	-0.0001	-0.0008	-0.0012	-0.0023	-0.0020

Notes: Distance elasticities are calculated for 2008. The table gives the percentage change in market share of brand i when the dealers of brand j are located 10% closer to consumers, where i indexes rows and j columns. Elasticities are calculated using estimates from specification (C) in Table 4.

Table 6 gives the distance elasticity estimates for a selection of brands, which are obtained as the predicted change in market share of brand i following a 10% decrease in distances to the dealerships of brand j. On average, a brand's market share increases by 0.94% if its dealers are located 10% closer to consumers. However, there is substantial variation—Smart and Lexus have relatively few dealerships and therefore see some of the biggest increases in market shares when moving closer to consumers. On the other hand, both Volkswagen and Ford have extensive dealership networks and their market shares only go up by around 0.6% when moving 10% closer to consumers. The substitution patterns shown in the table indicate that there is very little substitution with the outside good, so by moving closer to consumers, most of the additional consumers are substituting away from other brands. Moreover, consumers are more likely to substitute towards similar brands—for instance, if BMW dealers are 10% closer to consumers, Lexus' market share goes down by 0.07%, which is about six times larger than for the other brands in the table.

Table 7: Markups and variable profits

				Search			Full information		
	price	pre-tax price	sales	markup over MC	percentage markup	variable profit	markup over MC	percentage markup	variable profit
Toyota Auris	16,857	11,345	3,932	6,278	55.33	24.68	2,706	23.85	10.64
Renault Megane	18,647	12,048	4,310	6,213	51.57	26.78	2,797	23.22	12.06
Opel Zafira	22,812	15,037	5,641	6,489	43.15	36.60	2,947	19.60	16.63
Volkswagen Touran	23,416	15,180	3,668	7,172	47.25	26.31	3,096	20.40	11.36
Volkswagen Passat	24,129	15,447	8,496	7,189	46.54	61.08	3,125	20.23	26.55
Ford Mondeo	25,230	16,536	9,941	6,948	42.02	69.07	3,213	19.43	31.94
Audi A4	30,383	19,324	7,175	8,811	45.60	63.22	3,616	18.71	25.95
BMW 3-series	30,875	20,481	6,304	8,017	39.15	50.54	3,405	16.62	21.46
Mercedes C-class	32,609	20,878	4,540	7,150	34.25	32.46	3,319	15.90	15.07

Notes: Prices and markup over MC are in euro. Variable profit is in €1 mln. Markups and profits are calculated using estimates from specification (C) in Table 4 for the search model and specification (A) for the full information model. 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.

Table 7 compares the estimated markups between the search model and the full information model.

Consistent with the elasticity patterns reported in Table 5, estimated markups in the search model are higher for all the cars. The estimated average percentage markup across all models in 2008 is 41% for the search model versus 21% for the full information model. The full information markups are in line with several other studies of the automobile market: BLP report an average ratio of markup to retail price of 24% 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% for the model with moments. Goldberg (1995), on the other hand, obtains higher markup estimates: wholesale price markups of on average 38%, 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 7—the most profitable car model among all cars in the sample is the Volkswagen Golf.

5 Counterfactuals

In this section we study the effects of three changes in the primitives of the model. First, we look at what happens to equilibrium prices when the costs of visiting dealers decrease. For instance, during the COVID-19 pandemic it has become more common to offer at-home test driving, which has accelerated an already ongoing trend in the industry towards facilitating test driving by bringing cars to the buyers' homes or work places.[36] We model this situation as if the cost of transportation to the dealership went down all the way to zero. Second, we make a comparison between equilibrium prices in the search model and equilibrium prices in a full information model by simulating a market in which consumers do not need to search to find which cars suit their needs best. Finally, we study the stability of the exclusive dealing arrangements that currently prevail in the Dutch market and whether these arrangements can work as an anticompetitive mechanism by limiting search and raising profits at the expense of consumers. To do this, we explore the extent to which car manufacturers' variable profits change when they move away from an exclusive dealing arrangement and start selling other brands in their dealerships. We do this in two settings. In the first setting, we let the leading brands add smaller brands that are part of the same business group and show that car manufacturers have no incentive to do this. In the second setting, we study what happens if we pair the leading brands with brands of competing manufacturers only.

5.1 Changes 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 (C) of Table 4 and use the supply side model discussed in Section 2.6 to 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. Table 8 summarizes the effects on prices for the same car models as in Table 7. As expected, we observe that reducing distance-related search costs by fifty percent results in lower prices for all the models, while the impact on prices is more

[36]See https://www.caranddriver.com/news/a32713343/infiniti-at-home-shopping.

Table 8: Simulated prices for different levels distance-related search costs

Price	percentage of distance-related search costs 100%	50%	0%	prices observed	full info ($\lambda = -\infty$)
Toyota Auris	16,857	16,793	16,726	12,173	12,073
Renault Megane	18,647	18,590	18,531	14,127	14,017
Opel Zafira	22,812	22,751	22,688	17,963	17,875
Volkswagen Touran	23,416	23,349	23,280	17,818	17,846
Volkswagen Passat	24,129	24,063	23,994	18,544	18,582
Ford Mondeo	25,230	25,168	25,104	19,940	19,917
Audi A4	30,383	30,220	30,044	22,794	22,775
BMW 3-series	30,875	30,814	30,729	23,746	23,577
Mercedes C-class	32,609	32,577	32,534	26,784	26,611
Average all models	29,214	29,159	29,100	25,941	25,840
Share not searching	0.881	0.876	0.870	0.867	—
Average percentage markup	40.52	40.41	40.28	32.59	32.19

Notes: Prices are in euros. The prices shown are simulated prices when search costs are 100, 50, or 0 percent of the distance-related part of the estimated search costs from specification (C) in Table 4. The second-to-last column gives simulated prices when price deviations are observed in the search model. The last column gives simulated prices for the full information model.

pronounced if distance-related search costs are eliminated altogether. In the last column of this table we also report the equilibrium prices that manufacturers would charge if consumers had full information and did not need to engage in costly search at all.[37] A comparison of the average price under full information with actual prices suggests that prices are on average €3,374 higher because of search frictions. As we will discuss in more detail in the next subsection, this substantial effect on equilibrium prices is primarily caused by the fact that under full information price deviations are observed before search, which completely removes the incentives firms have to hold up consumers upon visiting the dealership (cf. Diamond, 1971).[38] Indeed, as shown in the second-to-last column of the table, if we take the estimates from our search model and simulate equilibrium prices in our search model assuming price deviations are observed as well, prices go down by almost as much as when assuming full information.[39]

[37] 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 $\lambda \to -\infty$.

[38] In fact, in a related study where consumers are assumed to observe price deviations before venturing the dealerships, Murry and Zhou (2020) find that search frictions lead to an average price increase for new cars of only \$333.

[39] Haan et al. (2018) discuss the effect of making price deviations observable in a simplified version of our model. When price deviations are observable, an individual firm has an incentive to cut its price to attract consumers to its store. Once consumers are at the store, they can be held up like in our model. They show that the first effect has a dominating influence so that equilibrium prices when price deviations are observable before search are lower than in the full information setting. The extent to which prices are cut (in the case of observable price deviations) or increased (in the case of unobservable price deviations) relative to the full information price depends on the gains from search. When the gains from search are low, consumers do not easily move from firm to firm. Hence, firms whose price deviations are observed do not need to cut prices much to attract consumers to their stores, and firms whose price deviations are not observed have large incentives to hold up consumers. On the other hand, when the gains from search are large, it becomes more likely to lose a customer to the rival. This gives firms whose price deviations are observed incentives to price aggressively to attract consumers to the store, and firms whose price deviations are not observed incentives not to hold up consumers much. In terms of the results shown in Table 8, this suggests that in our setting the variance of the observable part of utility, which is captured by the magnitude of the utility parameters, is much larger than the variance of the unobservable part of utility, which is given by the variance of the match values ε.

Table 9: Exclusive dealing versus multi-branding

	Exclusive dealing							Multi-branding			
	current variable profit	change in variable profit when deviating to multi-branding			change in variable profit when collectively moving to multi-branding			new variable profit	change in variable profit when deviating to exclusive dealing		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	joint	main	other	joint	main	other	joint	joint	main	other	joint
PANEL A: ADD OWN BRANDS											
Ford	432.75	-20.53	-32.89	-53.43	2.64	-23.17	-20.53	412.22	21.68	36.62	58.31
Opel	298.65	-8.10	-9.20	-17.29	18.69	-5.79	12.90	311.55	8.62	10.78	19.40
Renault	298.53	-10.04	-14.03	-24.07	13.00	-8.57	4.44	302.96	10.70	16.22	26.92
Volkswagen	581.74	-30.78	-48.18	-78.95	-8.16	-34.03	-42.19	539.55	32.09	52.69	84.78
Peugeot	396.98	-19.78	-25.72	-45.50	2.61	-15.29	-12.69	384.30	21.27	28.70	49.97
Fiat	119.73	-3.24	-1.22	-4.46	9.53	0.55	10.08	129.81	3.65	1.51	5.16
Toyota	292.02	-8.64	-9.59	-18.23	17.03	-5.81	11.22	303.24	9.20	11.10	20.30
Mercedes-Benz	91.28	-1.77	-0.20	-1.97	6.57	1.97	8.55	99.82	1.96	0.31	2.27
Multi-brand firms	2,511.68				61.91	-90.14	-28.23	2,483.46			
Independent brands	415.15						53.38	468.53			
Total	2,926.83						25.16	2,951.99			
PANEL B: ADD COMPETING BRANDS											
Ford	686.53	43.95	54.20	98.15	4.43	-0.80	3.63	690.16	-39.13	-50.46	-89.59
Opel	474.88	34.49	11.05	45.54	-7.69	-23.08	-30.76	444.12	-29.13	-11.01	-40.14
Renault	463.27	28.41	13.19	41.60	-10.07	-23.34	-33.40	429.86	-23.87	-12.62	-36.50
Volkswagen	394.95	33.47	3.60	37.07	-22.94	-10.43	-33.37	361.59	-26.11	-3.11	-29.22
Peugeot	427.57	21.79	31.51	53.30	-23.40	-2.06	-25.46	402.11	-17.42	-26.79	-44.22
Fiat	134.59	1.33	2.32	3.65	-18.43	-3.73	-22.16	112.44	-0.86	-1.90	-2.76
Toyota	260.08	0.39	1.69	2.08	-42.33	-1.49	-43.82	216.26	0.07	-1.33	-1.26
Mercedes-Benz	84.95	4.61	1.29	5.90	-12.12	-1.10	-13.22	71.73	-3.43	-1.01	-4.44
Total	2,926.83				-132.54	-66.02	-198.57	2,728.26			

Notes: Variable profit is in €1 mln and excludes fixed costs. Results are obtained using the estimates from specification (D) in Table 4. Column (1) gives the estimated joint variable profits for each of the incumbents listed and their paired entrants, where the pairs are according to the agreements listed in Table A1 of Appendix G. Columns (2)-(4) give the counterfactual change in variable profits when a pair unilaterally moves to multi-brand dealerships, where 'main' represents the incumbent, 'other' the entrants, and 'joint' the pair. Columns (5)-(7) give the counterfactual change in variable profits when the industry collectively moves to multi-brand dealerships. Column (8) gives the counterfactual joint variable profits for each of the incumbents and their paired entrants when collectively moving to multi-brand dealerships. Columns (9)-(11) give the counterfactual change in variable profits when unilaterally using exclusive dealing arrangements when the rest of the industry is using multi-brand dealerships. Independent firms in Panel A are brands that are not part of one of the eight incumbent brands, and therefore do not participate in the counterfactual multi-brand agreements.

5.2 Exclusive-dealing versus multi-branding

Exclusive dealing is the prevalent form of automobile distribution in the European Union, which historically has allowed manufacturers to use several types of vertical restraints that may prevent dealers from selling multiple brands. Specifically, since 1985 various versions of the European Motor Vehicle Block Exemption Regulation have been in effect, allowing car manufacturers to adopt either selective distribution or exclusive distribution and permitting exclusive dealing for up to 30-80% of a dealer's sales.[40] As a result, 69% of dealerships in our data are exclusive to one brand, which is also the average in Europe, but higher than the 57% reported for the United States (see Nurski and Verboven, 2016, and references therein).

In this section we study the stability of the exclusive dealing arrangements that currently prevail in the Dutch market and their effect on consumers. Exclusive and non-exclusive dealing each has its advantages and disadvantages for both consumers and car manufacturers. One important advantage of exclusive dealing is that it may generate value for consumers by encouraging retailers to provide extra services such as attractive showrooms, specialized sales people, test-driving vehicles, or fast after-sales service (Nurski and Verboven, 2016). Moreover, exclusive dealing may make the entry of new brands more difficult, thereby reducing potential competition. Multi-branding, in contrast, has the advantage that it lowers search costs for consumers, which makes multi-brand dealerships more attractive to visit. This alters the competition patterns and affects prices, which makes it difficult to establish a priori which of these effects will dominate. Further, it is likely that the relative strength of these effects depends on the distribution arrangement adopted by the rest of the industry, which potentially leads to multiple equilibria in terms of dealership structure (Bernheim and Whinston, 1998). In what follows, we use our model to study whether exclusive dealing is preferred by car manufacturers over an arrangement in which car dealerships sell multiple brands and how this affects welfare. Moreover, we also look at the stability of exclusive-dealing and multi-branding arrangements.

To capture the value of exclusivity to consumers we follow Nurski and Verboven (2016) and add an exclusivity dummy to our utility specification, which takes on value one if the dealer selling the car model is exclusive and zero otherwise. The estimates of this specification are reported in column (D) of Table 4. Note that to estimate the additional exclusivity variable, we add additional micro-moments to the model, as reported in column (D) of Table A2. According to the estimates, consumers put positive value on purchasing from an exclusive dealer—the parameter estimate of 0.056 corresponds to €275 on average. We use the estimates in column (D) of Table 4 in our counterfactual simulations and simulate market shares, profits, and prices when manufacturers allow dealerships to start selling additional brands in their locations.

Following Nurski and Verboven (2016) we make a distinction between more established brands (incumbent brands) and smaller brands (entrants), where the incumbent brands are the flagship brands of the eight largest car manufacturers in our data. For simplicity, we assume the counterfactual multi-branding agreements give entrants access to the dealership network of an incumbent while simultaneously closing the existing dealership

[40] Before 1985 vertical restraints were allowed through individual exceptions, which was a costly process, and led the European Commission to move to block exemptions for the entire car industry. See Nurski and Verboven (2016) for a historic overview of vertical restraints in the European car market.

network of the entrants.[41] We first let dealerships of incumbent brands only add entrant brands that are part of their own business group and show that car manufacturers have no incentive to do this. For instance, using the ownership structure of 2007, this means a Toyota dealer would pair up with Daihatsu and Lexus, whereas Fiat would give both Alfa Romeo and Lancia access to its dealers. Next we study what happens if we pair the incumbent brands with brands of competing manufacturers only. Here we pair up incumbents with competing entrants according to the number of dealerships, i.e., we pair up the incumbent brand with the largest number of dealerships to the four entrant brands with the most dealerships, the second largest incumbent with the four entrants than come next in terms of number of dealerships, etc. Table A1 in Appendix G shows the list of multi-brand agreements under both scenarios.[42]

Table 9 reports the effects of changes in dealership structure on the variable profits of the brands involved. Column (1) of the table gives the combined current variable profits for each of the incumbent brands and their paired entrants. Our first goal is to see whether there is an incentive for each individual combination of incumbent and entrants to move to multi-branding. To analyze this, columns (2)-(4) give changes in the variable profits of the main (incumbent) brand, the other (entrant) brands, as well as the change in their joint variable profits when each of the eight incumbents unilaterally shifts to multi-branding, while assuming the other incumbents stick to exclusive dealing arrangements.

First consider adding own brands. The results in Panel A indicate that starting from the current exclusive dealing arrangements, variable profits decrease when unilaterally moving to multi-branding for each of the incumbents and their affiliated brands. This result may seem surprising at first, because adding additional brands to an incumbent's dealership network should make these dealerships more attractive to consumers in comparison to those that have not added any brands. However, there are several effects at work. First, a move to multi-branding reduces consumers' utility for the brands involved because consumers value exclusivity. By this effect, prices and profits of the cars sold through multi-brand dealerships tend to decrease. Second, the fact that multi-brand dealers become more prominent in the market place tends to reduce their prices, rather than increase them. The reason for this is that market prominence expands the pool of visitors who visit the dealer early on in their search spell—because these "early" visitors tend to be more price sensitive than later ones, this puts downward pressure on prices, although this generally has a positive effect on profits (see also Armstrong, 2017). Third, following the logic of the Diamond paradox (Diamond, 1971), firms have a greater incentive to hold up visiting consumers because of search frictions. A move to multi-brand dealerships makes the cars of the brands involved more substitutable because they can be inspected and compared at no additional costs. When the brands that are added are part of the same business group, the incentive to hold up consumers increases because visiting consumers are more likely to substitute towards other brands owned by the same firm, which creates upward pressure on prices. The effects on profits are negative nevertheless,

[41] Note that the incumbent brands tend to have more dealership locations than the entrant brands, which means entrants would still get better coverage by getting access to the dealership network of an incumbent despite losing their own dealership network. However, alternative arrangements such as keeping existing locations of the entrants and allowing for two-way access (so the incumbents get additional dealerships as well) lead to similar results.

[42] This type of assortative matching guarantees that all entrants are paired with an incumbent that has more dealerships. However, alternative ways to pair entrants to incumbents (such as pairing the incumbent with the largest number of dealerships to the entrant brands with the lowest number of dealerships, as in Nurski and Verboven, 2016), lead to very similar results.

because in anticipation of this price increase consumers adjust their search behavior, which results in lower demand and lower profits. Note that a crucial feature of the Diamond paradox is that the temptation to hold up consumers by raising prices cannot be avoided—in Diamond (1971) prices increase enough to lead to a break down of the market, while in our setting, because of product differentiation, this manifests itself as a decrease in profits. Taking all three effects together, the loss of exclusivity and prominence effects are dominated by the hold-up effect, with higher prices as a result. The drop in demand that follows is significant enough for variable profits to go down, which means none of the manufacturers benefit from unilaterally moving to multi-brand dealerships that sell own-brands only. However, when all incumbents collectively shift to multi-branding, because all firms raise prices the drop in individual demand is less severe, and as shown in columns (5)-(7), some of the parties might actually see higher variables profits under the new multi-branding arrangement. Still, the combined change in variable profits of the firms that move to multi-brand dealerships is negative (-€28.23 mln). Moreover, as can be seen from columns (9)-(11) of the table, not sharing dealerships when everyone else is doing so results in higher variable profits for the business group that sticks to exclusive dealing, which shows there is an incentive to deviate from a multi-brand arrangement, suggesting this is not a stable industry structure.

The results from Panel A show that there is little incentive to form multi-brand dealerships that consist of brands that are part of the same firm, so we next focus on multi-brand dealerships that pair up incumbents with competing brands only. Panel B shows the results for this case. Starting from the exclusive dealing status-quo, columns (2)-(4) show that when an incumbent unilaterally grants competing brands access to its network of dealerships, unlike the case of adding own brands, the variable profits of both the main brand as well as the entrants increase. The main difference between the two cases is that now there is less incentive to hold up consumers relative to the status-quo because within-dealership markets become more competitive —cross-price elasticities between the incumbent and the added brands again increase, but now consumers are more likely to switch to the added, competing, brands. The overall effect is for prices to go down significantly, and joint profits to go up, which suggests this type of arrangement could destabilize the exclusive-dealing status-quo. The table also shows that when the incumbent dealerships collectively move to multi-branding, the overall effect on variable profits is dramatically different: variable profits decrease for everyone except Ford. The total decrease in joint variable profits when moving to this new multi-branding setting is close to €200 mln. Moreover, as shown in columns (9)-(11), the change in variable profits when deviating to exclusive dealing when everyone else is using multi-branding suggests that multi-branding could be a stable situation—for all multi-branding arrangements in the table joint profits go down when deviating to exclusive dealing. However, the comparison of variable profits under multi-branding to current profits shows that exclusive dealing is Pareto dominant, so manufacturers collectively prefer to stick to exclusive dealing agreements. This could explain why from an industry perspective, exclusive dealing regulation might be necessary to prevent the industry from moving to this type of multi-branding arrangement.

Table 10 provides an in-depth look at the welfare changes following a move from exclusive-dealing to multi-branding for the case in which only competing brands are added to the network of the incumbents.

Table 10: Welfare effects

	current	total change	change due to distance	change due to multi-branding	change due to prices
market shares (percentages)					
outside good	93.24	-1.14	-0.05	0.06	-1.15
inside goods	6.76	1.14	0.05	-0.06	1.15
main brand	56.41	-1.45	-0.87	0.01	-0.59
other brands	43.59	1.45	0.87	-0.01	0.59
variable profits (million euros)					
main brands	1,779.94	-132.54	-13.20	-16.48	-102.87
other brands	1,146.89	-66.02	29.63	-11.18	-84.47
total	2,926.83	-198.57	16.44	-27.66	-187.34
consumer surplus (million euros)		1,042.38	37.18	-60.42	1,065.63
welfare (million euros)		843.81	53.62	-88.09	878.28

Notes: Results are obtained using the estimates from specification (D) in Table 4.

The table splits out the effect by changes due to lower distance-related search costs, changes due to dealerships losing their exclusivity status, and changes due to the new equilibrium prices. The table shows that when the incumbents add competing brands, the market share of the outside good decreases by 1.14%, which corresponds to 82,634 additional cars sold. All brands are harmed from their dealers losing exclusivity. Overall consumer surplus increases by over one billion euros. This increase occurs despite the impact on exclusivity, which decreases consumer surplus, and is due to consumers benefitting from brands being closer to them, making it easier to shop around, and the substantial decrease in equilibrium prices, which go down by an average of €1,800. Since the change in consumer surplus more than offsets the decrease in variable profits, welfare is expected to go up as well, by approximately €844 mln. We conclude that exclusive dealing regulation substantially benefits car manufacturers at the expense of consumers, who not only face higher search costs due to exclusive dealing, but also pay significantly higher prices due to the absence of competition within dealerships.

6 Conclusions

In this paper we have presented a differentiated product model in which consumers are initially unaware of whether a given product is a good match or not. Consumers search sequentially, taking into account their preferences for the various alternatives as well as the costs of searching them. We have solved the model using recent findings from consumer search theory that re-characterize the search problem as a standard discrete choice problem (Armstrong, 2017; Choi et al., 2018) and have shown that for a Gumbel preserving search cost distribution we obtain a closed-form expression for the buying probabilities.

We have proposed an estimation method that supplements aggregate data on market shares and product characteristics with individual-specific data on searches and purchases, which makes it possible to estimate the effect of covariates that enter both utility and search cost specifications. We have estimated the model using data from the Dutch automobile market using a combination of aggregate and survey data. Our estimation results have shown that search costs are precisely estimated and economically meaningful. Assuming, instead, that search frictions are negligible and consumers have full information results in higher (absolute) own-price elasticity estimates, lower within-brand price elasticities, and lower estimated markups. According to our counterfactual estimates, the price of the average car is approximately 13% higher than in the absence of search frictions. We have also investigated the competitive and welfare effects of exclusive dealing regulation, which historically has been allowed in the European Union and has been used to limit the number of brands that are sold by dealerships. Our counterfactuals indicate that the effect on variable profits of adding brands to dealerships depends on whether the added brands are owned by the same manufacturer: adding own brands generally results in lower variable profits while unilaterally adding competing brands increases profits. This difference is due to the different ways in which a Diamond-type hold-up problem manifests itself and how this affects equilibrium prices—when own brands are added to a dealership, consumers are more likely to substitute towards own brands, which creates upward pressure on prices despite lowering profits, while the opposite happens when competing brands are added. Even though this means there is only an incentive to add competing brands, we have shown that when the entire industry moves to multi-brand arrangements of this type, profits go down substantially. Firms therefore benefit from exclusive dealing regulation at the expense of consumers who face higher search costs and higher prices.

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.[43] 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.

[43] In recent work, D'Haultfœuille et al. (2019) 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

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 André de Palma: "Multiproduct Firms: A Nested Logit Approach," *The Journal of Industrial Economics* 40, 261–276, 1992.

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.

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

Bernheim, B. Douglas and Michael D. Whinston: "Exclusive Dealing," *Journal of Political Economy* 106, 64–103, 1998.

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.

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

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," *Review of Economic Studies* 86, 1973–1998, 2019.

Diamond, Peter A.: "A Model of Price Adjustment," *Journal of Economic Theory* 3, 156–168, 1971.

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.

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

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.

Grigolon, Laura, Mathias Reynaert, and Frank Verboven: "Consumer Valuation of Fuel Costs and Tax Policy," *American Economic Journal: Economic Policy* 10, 193–225, 2018.

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, and Pradeep Chintagunta: "Simultaneous or Sequential? Search Strategies in the U.S. Auto Insurance Industry," *Marketing Science* 36, 21–42, 2017.

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.

Jiang, Zhenling, Chan, Tat, Che, Hai and Youwei Wang: "Consumer Search and Purchase: An Empirical Investigation of Retargeting Based on Consumer Online Behaviors," *Marketing Science*, forthcoming, 2021.

Jolivet, Grégory and Hélène Turon: "Consumer Search Costs and Preferences on the Internet," *Review of Economic Studies* 86, 1258–1300, 2019.

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.

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, Zsolt Sándor, and Matthijs R. Wildenbeest: "Semi-Nonparametric Estimation of Consumer Search Costs," *Journal of Applied Econometrics* 28, 1205–23, 2013.

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

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 Colocation," *Mangement Science* 66, 1909–1934, 2020.

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.

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.

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.

Seiler, Stephan and Fabio Pinna: "Estimating Search Benefits from Path-Tracking Data: Measurement and Determinants," *Marketing Science* 36, 565–589, 2017.

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.

Varian, Hal R.: "A Model of Sales," *American Economic Review* 70, 651–659, 1980.

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.

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 (10) 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, that is,

$$F_{if}^c(H_{if}(z)) = \frac{1-F_{if}^w(z)}{1-F_{if}(z)}. \tag{A22}$$

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 (A22). Specifically, use the change of variables $c = H_{if}(z)$ in equation (A22) to obtain

$$F_{if}^c(c) = \frac{1-F_{if}^w(H_{if}^{-1}(c))}{1-F_{if}(H_{if}^{-1}(c))} = \frac{1-F_{if}^w(\delta_{if}+H_0^{-1}(c))}{1-F_{if}(\delta_{if}+H_0^{-1}(c))} = \frac{1-F_{if}^w(\delta_{if}+H_0^{-1}(c))}{1-F(H_0^{-1}(c))}, \tag{A23}$$

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 (A23) 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 $\delta_{if} - \mu_{if}$, that is,

$$F_{if}^w(z) = \exp\left(-\exp\left(-\left(z-(\delta_{if}-\mu_{if})\right)\right)\right). \tag{A24}$$

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 (A23) gives

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

Because we have taken the utility shock distribution F to be the TIEV distribution we get:

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

Given equation (A24), calculation of the probability of buying from firm f in equation (11) is straightforward, that is,

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

The probability that consumer i buys product j is $s_{ij} = P_{if} P_{ij|f}$ and simplifies to

$$\begin{aligned} s_{ij} &= \frac{\exp\left(\delta_{if} - \mu_{if}\right)}{1 + \sum_{g=1}^{F} \exp\left(\delta_{ig} - \mu_{ig}\right)} \times \frac{\exp(\delta_{ij})}{\sum_{h \in G_f} \exp(\delta_{ih})} \\ &= \frac{\exp\left(\delta_{ij} - \mu_{if}\right)}{1 + \sum_{g=1}^{F} \exp\left(\delta_{ig} - \mu_{ig}\right)} = \frac{\exp\left(\delta_{ij} - \mu_{if}\right)}{1 + \sum_{g=1}^{F} \sum_{h \in G_g} \exp\left(\delta_{ih} - \mu_{ih}\right)} \\ &= \frac{\exp\left(\delta_{ij} - \mu_{if}\right)}{1 + \sum_{k=1}^{J} \exp\left(\delta_{ik} - \mu_{ig}\right)}, \end{aligned} \tag{A25}$$

where in the last line g denotes the firm that produces product k. ■

C Derivation of Search Probabilities

Conditional on the δ_{if}'s and the search cost covariates, 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

$$\pi_{i0} = \Pr\left[u_{i0} \geq \max_{k \neq 0} \{r_{ik}\}\right] = \int_{-\infty}^{\infty} \int_{-\infty}^{u_{i0}} dF_i^{r_{max}}(z) dF(u_{i0}) = \int_{-\infty}^{\infty} F_i^{r_{max}}(z) f(z) dz, \tag{A26}$$

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 (9) for F_{if}^r, that is,

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

Using $F_{if}^{w}(z)$ as specified in equation (13), we get

$$F_{if}^{r}(z) = \frac{\exp\left(-\exp\left(\delta_{if} - \mu_{if} - z\right)\right) - \exp(-\exp(\delta_{if} - z))}{1 - \exp(-\exp(\delta_{if} - z))}. \tag{A27}$$

We now compute the conditional 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, that is,

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

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

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

This probability can be written as

$$\pi_{i1} = \sum_f \int_{-\infty}^{\infty} \left(1 - F_{if}^{r}(y)\right) F_{i,-f}^{r}(y) f(y) dy + \sum_f \int_{-\infty}^{\infty} F(x)\left(1 - F_{if}(x)\right)\left(1 - F_{if}^{r}(x)\right) f_{i,-f}^{r}(x) dx,$$

where $F_{i,-f}^{r}$ and $f_{i,-f}^{r}$ are the CDF and PDF of $\max_{k\neq 0,f} r_{ik}$. Note that by equation (10),

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

and by integration by parts

$$
\begin{aligned}
\int_{-\infty}^{\infty} F(x)\left(1-F_{if}^{w}(x)\right) f_{i,-f}^{r}(x)dx &= \int_{-\infty}^{\infty} F(x)\left(1-F_{if}^{w}(x)\right) dF_{i,-f}^{r}(x) \\
&= -\int_{-\infty}^{\infty} F_{i,-f}^{r}(x) d\left[F(x)\left(1-F_{if}^{w}(x)\right)\right] \\
&= \int_{-\infty}^{\infty} F_{i,-f}^{r}(x)F(x)f_{if}^{w}(x)dx - \int_{-\infty}^{\infty}\left(1-F_{if}^{w}(x)\right) F_{i,-f}^{r}(x)f(x)dx,
\end{aligned}
$$

where we use

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

We finally get that

$$
\pi_{i1} = \sum_{f}\int_{-\infty}^{\infty}\left(F_{if}^{w}(y) - F_{if}^{r}(y)\right) F_{i,-f}^{r}(y)f(y)dy + \sum_{f}\int_{-\infty}^{\infty} F_{i,-f}^{r}(x)F(x)f_{if}^{w}(x)dx. \tag{A28}
$$

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

D Market Share Derivatives

Own-price derivatives

Denote $\alpha_i^* = \alpha + \alpha_i$. The market share derivative with respect to price p_j is given by

$$
\frac{\partial s_j}{\partial p_j} = \int \frac{\partial s_{ij}}{\partial p_j} dF_\tau(\tau_i) = \int \alpha_i^* \frac{\partial s_{ij}}{\partial \delta_{ij}^d} dF_\tau(\tau_i),
$$

where $\delta_{ij}^d = \delta_{ij} - \alpha_i^* p_j^* + \alpha_i^* p_j$. The derivative of the buying probability s_{ij} with respect to δ_{ij}^d is

$$
\frac{\partial s_{ij}}{\partial \delta_{ij}^d} = \frac{\partial P_{if}}{\partial \delta_{if}^d}\frac{\partial \delta_{if}^d}{\partial \delta_{ij}^d} P_{ij|f} + P_{if}\frac{\partial P_{ij|f}}{\partial \delta_{ij}^d} = \frac{\partial P_{if}}{\partial \delta_{if}^d} P_{ij|f}^2 + s_{ij}(1-P_{ij|f}),
$$

where $\delta_{if}^d = \log\left(\sum_{h\in G_{fJ}} \exp(\delta_{ih}) + \exp\left(\delta_{ij}^d\right)\right)$. The derivative of P_{if} with respect to δ_{if}^d is given by

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

In computing $\partial f_{if}^{w}/\partial \delta_{if}^d$, we note that consumers' reservation values are not affected by deviation prices but actual utility levels are. To capture this explicitly, let us rewrite the distribution of w in equation (10) as $F_{if}^{w}(z) = 1 - F_{if}^{c}(H_{if}(z))(1-F_{if}^{d}(z))$, where the super-index d in $F_{if}^{d}(z)$ is meant to indicate the dependence of realized utilities on deviation prices. The density of w is given by

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

where only $F^d_{if}(z)$ and $f^d_{if}(z)$ depend on δ^d_{if}, and where we have used $\partial H_{if}(z)/\partial\delta_{if} = 1 - F_{if}(z)$.

The derivative of f^w_{if} with respect to δ^d_{if} is given by

$$\begin{aligned}\frac{\partial f^w_{if}}{\partial \delta^d_{if}} &= f^c_{if}(H_{if}(z))(1-F_{if}(z))f^d_{if}(z) + F^c_{if}(H_{if}(z))f^d_{if}(z)(1-\exp(\delta^d_{if}-z));\\ &= \left(f^c_{if}(H_{if}(z))(1-F_{if}(z)) + F^c_{if}(H_{if}(z))(1-\exp(\delta^d_{if}-z))\right) f^d_{if}(z).\end{aligned}$$

Cross-price derivatives

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

$$\frac{\partial s_k}{\partial p_j} = \int \frac{\partial s_{ik}}{\partial p_j} dF_\tau(\tau_i) = \int \alpha^*_i \frac{\partial s_{ik}}{\partial \delta^d_{ij}} dF_\tau(\tau_i);$$

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

$$\frac{\partial s_{ik}}{\partial \delta^d_{ij}} = \frac{\partial P_{if}}{\partial \delta^d_{if}} \frac{\partial \delta^d_{if}}{\partial \delta^d_{ij}} P_{ik|f} + P_{if}\frac{\partial P_{ik|f}}{\partial \delta^d_{ij}} = \frac{\partial P_{if}}{\partial \delta^d_{if}} P_{ij|f}P_{ik|f} - s_{ik}P_{ij|f}.$$

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

$$\frac{\partial s_{ik}}{\partial \delta^d_{ij}} = \frac{\partial P_{ih}}{\partial \delta^d_{if}} \frac{\partial \delta^d_{if}}{\partial \delta^d_{ij}} P_{ik|h} = \frac{\partial P_{ih}}{\partial \delta^d_{if}} P_{ij|f}P_{ik|h}.$$

The derivative of P_{ih} with respect to δ^d_{if} is given by

$$\frac{\partial P_{ih}}{\partial \delta^d_{if}} = \int \left(\prod_{g\neq f,h} F^w_{ig}(z)\right) \frac{\partial F^w_{if}(z)}{\partial \delta^d_{if}} f^w_{ih}(z)dz. \tag{A29}$$

The derivative of F^w_{if} with respect to δ^d_{if} is given by

$$\frac{\partial F^w_{if}}{\partial \delta^d_{if}} = -F^c_{if}(H_{if}(z))f^d_{if}(z), \tag{A30}$$

where once again we have taken into account that the gains from search $H_{if}(z)$ do not depend on deviation prices.

E Contraction Mapping

Here we present a version of the Contraction Theorem from BLP in which we specify conditions on the market share function instead of the contraction mapping itself. The conditions of our theorem imply those from BLP in the following way. Condition 1 from our theorem implies condition (1) from BLP's (p.887), condition 1 and 2 from our theorem imply condition (2) from BLP's, and condition 3 from our theorem implies condition (3) from BLP's.

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

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

(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 function $\overline{\sigma}_j$ defined in Condition 1 satisfies

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

3. 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

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

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

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

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

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 $\overline{f}_j(\xi) = \min\left[\overline{\xi}, f_j(\xi)\right]$ 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 (A25) 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. 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 (14) we have

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

where $\alpha_i^* = \alpha + \alpha_i$ and $\beta_i^* = \beta + \beta_i$, 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.

Condition 3. 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 $\delta_{ij}(x)$ denotes the expression δ_{ij} where ξ_j is replaced by x. From this it is straightforward to obtain that $\lim_{x\to-\infty}\widetilde{s}_0^j(x)=1$ and $\lim_{x\to\infty}\widetilde{s}_0^j(x)=0$. In conclusion, the contraction property holds.

F Data

Aggregate data

Data on car characteristic and prices are obtained from Autoweek Carbase (see autoweek.nl/carbase), where for each car model we use the characteristics of the least expensive version. Sales data are obtained from BOVAG. We include a car model in a given year if more than fifty cars are 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 sold during this period, which corresponds to about 230 different models in any given year. We treat each model-year combination as one observation, which results in a total of 1,382 observations. We use post-tax prices. The tax when buying a new car in the Netherlands consists of a sales tax as well as an additional automobile tax. The 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.

In all our specifications we use segment dummies. 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.

We obtain predicted market shares by sampling from neighborhood level data provided by Statistics Netherlands. We only include neighborhoods with a strictly positive number of inhabitants, which leaves us with a total of 11,122 neighborhoods for 2007. Most neighborhoods are relatively small; the mean number of inhabitants is 1,471. 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). 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 households with children, as well as mean (after-tax) household income level. We obtain income draws by randomly drawing 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. 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.

Micro data

The micro-level data on individual searches and purchases used for the moments in equations (20) and (21) is obtained from two separate surveys that were administered in 2010 and 2011 by the Dutch survey agency TNS NIPO (see tns-nipo.com). 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 on current car owners—1,297 for the survey carried out in 2010 and 1,233 for the 2011 survey.[44] 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.[45] 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.[46] Finally, the respondents answered questions about their household income, household size, age, whether there are children living in the household, and zip code. The income bound $\overline{y}$ we use in the moments corresponds to a household income (after tax) of €25,000. Note that 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.

[44] We have excluded 34 observations for households that do not own a car.

[45] 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?". Approximately 14 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 28 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.

[46] We only have information on same-day visits for the survey administered in 2011. The vast majority of respondents who bought a new car and visited dealerships of distinct brands (and for which we have this information) did not make all of those visits on the same day, which is consistent with a sequential search strategy.

Instruments

We follow Reynaert and Verboven (2014) in using unit labor costs and steel prices interacted with car weight (both normalized by country-specific producer price indices). To construct these two instruments, we have collected information about the main country of production for each car model, using publicly available information from annual reports and industry reports. Unit labor costs are country-specific and are calculated using data from the Conference Board International Labor Comparisons program. For steel prices, we use the CRU Steel Price Indicator for Europe, which we obtained through Datastream. Producer Price Indices are obtained from the IMF International Financial Statistics.

To construct model-specific annual production, we use the total global production and sales for each of the 16 manufacturers in our data (using data from the Global Market Data Book, OICA, and manufacturers' annual reports), and use European sales data at the model level (using data from carsalesbase.com) to create a proxy for production at the model level. Specifically, we created this variable as follows:

$$\text{production}_j = \text{global production}_m \times \frac{\text{European sales}_j}{\text{global sales}_m},$$

where subscript j is used to indicate a car model and m a manufacturer.

G Additional Tables

Table A1: List of agreements used for counterfactuals

	Incumbent	Entrant 1	Entrant 2	Entrant 3	Entrant 4
PANEL A: ADD OWN BRANDS					
Agreement 1	Ford	Jaguar	Land Rover	Mazda	Volvo
Agreement 2	Opel	Cadillac	Chevrolet-Daewoo	Saab	
Agreement 3	Renault	Dacia	Nissan		
Agreement 4	Volkswagen	Audi	Seat	Skoda	
Agreement 5	Peugeot	Citroën			
Agreement 6	Fiat	Alfa Romeo	Lancia		
Agreement 7	Toyota	Daihatsu	Lexus		
Agreement 8	Mercedes-Benz	Jeep	Chrysler	Dodge	Smart
PANEL B: ADD COMPETING BRANDS					
Agreement 1	Ford	Citroën	Audi	Hyundai	Chevrolet-Daewoo
Agreement 2	Opel	Seat	Suzuki	Mazda	Nissan
Agreement 3	Renault	Kia	Volvo	Mitsubishi	Skoda
Agreement 4	Volkswagen	Dacia	Daihatsu	Alfa Romeo	Jeep
Agreement 5	Peugeot	Honda	BMW	Lancia	Chrysler
Agreement 6	Fiat	Subaru	Mini	Saab	Dodge
Agreement 7	Toyota	Smart	Land Rover	Porsche	
Agreement 8	Mercedes-Benz	Jaguar	Cadillac	Lexus	

Table A2: Fit moments

	Survey	Full info (A)	Search (B)	(C)	(D)
Purchase related probabilities					
$E[\mathbb{1}\{i \text{ purchases a new vehicle}\} \mid \{\text{median distance } i \text{ to dealers} \geq 7\text{km}\}]$	0.065	0.068	0.067	0.067	0.067
$E[\mathbb{1}\{i \text{ purchases a new vehicle}\} \mid \{\text{median distance } i \text{ to dealers} < 7\text{km}\}]$	0.064	0.065	0.065	0.065	0.065
$E[\mathbb{1}\{i \text{ purchases a new vehicle}\} \mid \{y_i < \overline{y}\}]$	0.038	0.039	0.038	0.037	0.037
$E[\mathbb{1}\{i \text{ purchases a new vehicle}\} \mid \{y_i \geq \overline{y}\}]$	0.085	0.087	0.087	0.089	0.089
$E[\mathbb{1}\{i \text{ purchases a new vehicle}\} \mid \{i \text{ is senior}\}]$	0.079	0.079	0.077	0.077	0.077
$E[\mathbb{1}\{i \text{ purchases a new vehicle}\} \mid \{i \text{ is not a senior}\}]$	0.061	0.063	0.063	0.063	0.063
$E[\mathbb{1}\{i \text{ purchases a new vehicle}\} \mid \{i \text{ has children in the household}\}]$	0.062	0.064	0.070	0.067	0.068
$E[\mathbb{1}\{i \text{ purchases a new vehicle}\} \mid \{i \text{ has no children in the household}\}]$	0.066	0.067	0.064	0.065	0.065
$E[\mathbb{1}\{i \text{ purchases a new vehicle}\} \mid \{i \text{ lives in an urban area}\}]$	0.062	0.062	0.062	0.062	0.062
$E[\mathbb{1}\{i \text{ purchases a new vehicle}\} \mid \{i \text{ does not live in an urban area}\}]$	0.070	0.072	0.072	0.073	0.073
$E[\mathbb{1}\{i \text{ purchases vehicle} \geq €20{,}000\} \mid \{y_i < \overline{y}\}]$	0.006	0.005	0.005	0.007	0.007
$E[\mathbb{1}\{i \text{ purchases vehicle} \geq €20{,}000\} \mid \{y_i \geq \overline{y}\}]$	0.041	0.037	0.037	0.036	0.036
$E[\mathbb{1}\{i \text{ purchases a large vehicle}\} \mid \{i \text{ has children in the household}\}]$	0.022	0.025	0.026	0.026	0.026
$E[\mathbb{1}\{i \text{ purchases a large vehicle}\} \mid \{i \text{ has no children in the household}\}]$	0.011	0.013	0.012	0.012	0.012
$E[\mathbb{1}\{i \text{ purchases from a dealer} \geq 7\text{km}\} \mid \{i \text{ lives in an urban area}\}]$	0.007	0.007	0.007	0.007	0.006
$E[\mathbb{1}\{i \text{ purchases from a dealer} < 7\text{km}\} \mid \{i \text{ lives in an urban area}\}]$	0.055	0.056	0.056	0.056	0.056
$E[\mathbb{1}\{i \text{ purchases a luxury brand}\} \mid \{i \text{ lives in an urban area}\}]$	0.005	0.005	0.006	0.006	0.006
$E[\mathbb{1}\{i \text{ purchases a non-luxury brand}\} \mid \{i \text{ lives in an urban area}\}]$	0.057	0.057	0.057	0.056	0.056
$E[\mathbb{1}\{i \text{ purchases from an exclusive dealer}\} \mid \{i \text{ lives in an urban area}\}]$	0.033				0.033
$E[\mathbb{1}\{i \text{ purchases from a non-exclusive dealer}\} \mid \{i \text{ lives in an urban area}\}]$	0.029				0.030
Search related probabilities					
$E[\mathbb{1}\{i \text{ searches once}\} \mid \{y_i < \overline{y}\}]$	0.041		0.040	0.037	0.037
$E[\mathbb{1}\{i \text{ searches once}\} \mid \{y_i \geq \overline{y}\}]$	0.057		0.057	0.059	0.059
$E[\mathbb{1}\{i \text{ searches once}\} \mid \{\text{median distance } i \text{ to dealers} \geq 7\text{km}\}]$	0.051		0.050	0.049	0.049
$E[\mathbb{1}\{i \text{ searches once}\} \mid \{\text{median distance } i \text{ to dealers} < 7\text{km}\}]$	0.049		0.049	0.050	0.050
$E[\mathbb{1}\{i \text{ searches once}\} \mid \{i \text{ is senior}\}]$	0.048		0.052	0.053	0.053
$E[\mathbb{1}\{i \text{ searches once}\} \mid \{i \text{ is not a senior}\}]$	0.050		0.049	0.049	0.049
$E[\mathbb{1}\{i \text{ searches once}\} \mid \{i \text{ has children in the household}\}]$	0.062		0.051	0.053	0.053
$E[\mathbb{1}\{i \text{ searches once}\} \mid \{i \text{ has no children in the household}\}]$	0.045		0.049	0.048	0.048
$E[\mathbb{1}\{i \text{ searches at least twice}\} \mid \{y_i < \overline{y}\}]$	0.045		0.062	0.048	0.048
$E[\mathbb{1}\{i \text{ searches at least twice}\} \mid \{y_i \geq \overline{y}\}]$	0.083		0.068	0.081	0.081
$E[\mathbb{1}\{i \text{ searches at least twice}\} \mid \{\text{median distance } i \text{ to dealers} \geq 7\text{km}\}]$	0.062		0.061	0.063	0.063
$E[\mathbb{1}\{i \text{ searches at least twice}\} \mid \{\text{median distance } i \text{ to dealers} < 7\text{km}\}]$	0.069		0.067	0.069	0.069
$E[\mathbb{1}\{i \text{ searches at least twice}\} \mid \{i \text{ is senior}\}]$	0.079		0.068	0.077	0.077
$E[\mathbb{1}\{i \text{ searches at least twice}\} \mid \{i \text{ is not a senior}\}]$	0.064		0.064	0.064	0.064
$E[\mathbb{1}\{i \text{ searches at least twice}\} \mid \{i \text{ has children in the household}\}]$	0.081		0.066	0.083	0.083
$E[\mathbb{1}\{i \text{ searches at least twice}\} \mid \{i \text{ has no children in the household}\}]$	0.061		0.065	0.060	0.060

Notes: The income bound $\overline{y}$ corresponds to a household income (after tax) of €25,000.