TI 2015-033/VII
Tinbergen Institute Discussion Paper

Consumer Search and Prices in the Automobile Market

José Luis Moraga-González[1]
Zsolt Sándor[2]
Matthijs R. Wildenbeest[3]

[1] Faculty of Economics and Business Administration, VU University Amsterdam, and Tinbergen Institute, the Netherlands;
[2] Sapientia Hungarian University of Transylvania, Romania;
[3] Indiana University, United States.

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at http://www.tinbergen.nl

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: http://www.dsf.nl/

Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579

Consumer Search and Prices in the Automobile Market*

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

February 2015

Abstract

In many markets consumers have imperfect information about the utility they derive from the products that are on offer and need to visit stores to find the product that is the most preferred. This paper develops a discrete-choice model of demand with optimal consumer search. Consumers first choose which products to search; then, once they learn the utility they get from the searched products, they choose which product to buy, if any. The set of products searched is endogenous and consumer specific. Therefore imperfect substitutability across products does not only arise from variation in their characteristics but also from variation in the costs of searching them. We apply the model to the automobile industry. Our search cost estimate is highly significant and indicates that consumers conduct a limited amount of search. Estimates of own- and cross-price elasticities are lower and markups are higher than if we assume consumers have full information.

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

*We thank Jean-Pierre Dubé, Amit Gandhi, Elisabeth Honka, Jean-François Houde, Sergei Koulayev, Ariel Pakes, Steven Puller, Stephan Seiler, Marc Santugini, Philipp Schmidt-Dengler, Michelle Sovinsky, and Frank Verboven for their useful comments and suggestions. This paper has also benefited from presentations at the Düsseldorf Institute for Competition Economics, Illinois State University, Katholieke Universiteit Leuven, The Ohio State University, Texas A&M 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, and 2014 MaCCI Workshop on Consumer Search in Bad Homburg. 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.

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

‡Sapientia University Miercurea Ciuc, Affiliate Fellow at CERGE-EI, Prague, E-mail: `zsosan@gmail.com`.

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

1 Introduction

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

Earlier work on the estimation of demand models (Berry, Levinsohn, and Pakes, 1995, 2004; Nevo, 2001; Petrin, 2002) has proceeded by assuming that consumers have perfect information about all the products available in the market. Since in most markets product information is gathered by consumers and/or advertised by firms, there are two natural ways to interpret the perfect information assumption made in earlier work. The first is that search costs are negligible for all consumers. The second interpretation is that firms' advertisements reach all consumers and convey all relevant information. In the market settings referred to above, the full information assumption is, arguably, unrealistic.

In a recent study of the US computer industry, Sovinsky Goeree (2008) shows that departing from the perfect information assumption is important for obtaining realistic estimates of demand and supply parameters. In her model, firms distribute advertisements about the existence and characteristics of the computers they sell. Advertisements are perfectly informative but not all consumers are reached by them. Moreover, consumers differ in characteristics that affect their exposure to advertisements. As a result consumers end up having heterogeneous and limited information about the existing alternatives in the market. Yet, in the setting of Sovinsky Goeree (2008) consumers do not need to incur any search costs to evaluate the utility they derive from the alternatives they happen to be informed of via the advertisements.

This paper adds to the literature on the structural estimation of demand and supply by presenting a discrete choice model of demand with optimal consumer search. To the best of our knowledge our paper is the first to do this in a Berry, Levinsohn, and Pakes (1995) (BLP hereafter) framework.

[1] Several recent empirical papers have found that consumers search relatively little. For instance, Honka (2014) reports that consumers obtain an average of 2.96 quotes when shopping for car insurance. De los Santos, Hortaçsu, and Wildenbeest (2012) find that over 75 percent of consumers visited only one online bookstore before buying a book online, whereas De los Santos, Hortaçsu, and Wildenbeest (2015) find that the mean number of online retailers searched is less than 3 for MP3 players. Some other examples of markets in which consumers are found to search little are S&P 500 index funds (Hortaçsu and Syverson, 2004) and automobiles (Moorthy, Ratchford, and Talukdar, 1997; Scott Morton, Silva-Risso, and Zettelmeyer, 2011).

The distinctive feature of the BLP framework is that the utilities of the products also depend on a structural error term, which is known as an unobserved product characteristic in this literature. This structural error is crucial for modeling price endogeneity, and naturally leads to estimation based on aggregate data. An important difference with the demand model in BLP is that in our model consumers first decide which sellers to visit in order to inspect the products they sell. Doing so is costly; after having incurred search costs to visit sellers, consumers obtain all the relevant information about the products inspected, and decide which product to acquire, if any. Search costs vary across individuals and firms so consumers choose to visit distinct subsets of sellers even if they have similar preferences, as for instance happens in the conditional logit model. In our model consumer choice sets are thus endogenous and consumer specific. As a result, imperfect substitutability across products does not only arise from product differentiation but also from the variation in consumer choice sets generated by costly search. Similar to the effects of advertising in Sovinsky Goeree (2008), search frictions thus generate heterogeneous and limited consumer information.

We apply the model to the automobile market. The automobile market is precisely a market in which advertisements, reports in specialized magazines, television programs and the Internet convey much but not 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. We model consumers' search decisions by letting consumers make a tradeoff between the expected gains from searching, which are based on information about car characteristics that are observed without doing an in-store search (design, size, horsepower, fuel efficiency, prices, etc.), and the cost of searching, which we relate to dealership locations and certain consumer demographics.

The complexity of the real-world setting to which we apply our model—specifically, firms selling multiple products and search costs that are alternative specific—makes the computation of consumers' optimal search sets a challenge. Our problem is indeed related to the general class of portfolio problems discussed in Chade and Smith (2006). In these problems a decision maker must simultaneously choose among a set of ranked stochastic options; each choice is costly and only the best realized option is exercised. When there are many alternatives available in the market, finding the optimal choice set is an extremely complex task because of the large number of choice sets to be evaluated. Chade and Smith (2006) provide an algorithm, known as the Marginal Improvement Algorithm (MIA), that identifies the optimal solution for some classes of problems. Unfortunately we cannot apply the MIA algorithm to our setting, because of the aforementioned complexities that are specific to our application. Instead, to address the dimensionality problem that arises because of the large number of potential choice sets a consumer faces, we add an (unobservable) choice-set specific error term to the costs of searching a subset of alternatives. The role of this shock is analogous to the error term in logit models and allows us to compute, for every consumer, the probability with which she searches any given choice set.

We use data from the Dutch market for new cars to estimate the model. We provide background information on this market in Section 2. Survey data reveal two important facts. First, consumers visit a limited number of car dealers before buying a car and this number varies substantially across consumers. Second, a great deal of the dealer visits involve test-driving cars. We interpret these two facts as suggesting that models of perfect information and models of imperfect information in which consumers just shop for lower prices are likely to be inapt. Instead, we formulate a model in which consumers have imperfect information and consumers have to search to evaluate the alternatives on sale. In this section we also provide some reduced-form evidence that search behavior is related to demographics such as income, family size, age, and distances to dealerships.

We develop our search model in Section 3 and discuss estimation and identification in Section 4. The model can be estimated using car characteristics and aggregate-level data on prices and market shares, as well as data on dealership locations and consumer demographics. By exploiting variation in distances from consumer households to car dealerships the magnitude of search costs can be identified. To obtain more precise estimates and to help identification we supplement the aggregate-level data with the survey data mentioned above—this enables us to relate consumer demographics to the characteristics of products that are searched and purchased.

The data and estimation results are presented in Section 5. Our search cost estimates are highly significant. One advantage of our model is that it nests the demand model of BLP. Our paper shows that taking into account search costs leads to lower estimates of own- and cross-price elasticities and higher estimates of price-cost margins. 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.

In Section 5 we also use the estimates to perform several counterfactual analyses. Specifically, taking our estimates as a starting point, simulations indicate that for some car models prices may be higher in the full information model than in the costly search model. To understand this perhaps surprising result, we note two important aspects of our model. The first is that we assume consumers observe prices before searching. In such a case, costly search has two implications on firm pricing (Haan and Moraga-González, 2011). On the one hand we get the standard result from the search literature that costly search leads to smaller consumer choice sets, which gives firms incentives to raise prices. On the other hand, costly search makes it more important for a firm to compete for visits. Since price is one of the product features that consumers use to make search decisions, costly search pushes firms to cut their prices. Which of these two effects dominates depends on search costs as well as competition—both of which are determined at the car model level. The second aspect is that in our model the decision to participate is endogenous because consumers have the option to choose the outside alternative right away. In such a case, an increase in search costs results in an increase in the number of consumers who opt out of the new car market and, as shown in Moraga-González, Sándor, and Wildenbeest (2014), this may result in lower prices. We also simulate the competitive effects of changes in the way manufacturers use their dealer networks and find mergers of dealerships to have non-trivial effects on prices and profits.

Related literature

Our paper builds on the theoretical and empirical literature on consumer search. At least since the seminal article of Stigler (1961) on the economics of information a great deal of theoretical and empirical work has revolved around the idea that the existence of search costs has nontrivial effects on market equilibria. Part of the effort has gone into the study of the effects of costly search in homogeneous product markets (see for instance Burdett and Judd, 1983; Reinganum, 1979; Stahl, 1989). In this literature a fundamental issue has been the existence of price dispersion in market equilibrium. Another tradition has been the study of costly search in markets with product differentiation. The seminal paper is Wolinsky (1986), who notes that search costs generate market power even in settings with free entry of firms. More recent contributions investigate how product diversity (Anderson and Renault, 1999), product quality (Wolinsky, 2005), and product design (Bar-Isaac, Caruana, and Cuñat, 2012) are affected by costly search. As in our model, in this literature consumers search for a good product fit, and not for lower prices. Our search model is most closely related to the logit search model discussed in Anderson, De Palma, and Thisse (1992), but we allow for asymmetric multi-product firms, consumer heterogeneity in both preferences and search costs, and allow (deviation) prices to be observable before searching.[2]

Some recent empirical research on consumer search behavior has focused on developing techniques to estimate search costs using aggregate market data. Hong and Shum (2006) develop a structural method to retrieve information on search costs for homogeneous products using only price data. Moraga-González and Wildenbeest (2008) extend the approach of Hong and Shum (2006) to the case of oligopoly and present a maximum likelihood estimator. Hortaçsu and Syverson (2004)

[2]See Section 7.6 (pp. 246–248) of Anderson, De Palma, and Thisse (1992).

study a search model where search frictions coexist with vertical product differentiation. Our paper contributes to this line of work by incorporating consumer search into the BLP framework.

A number of recent papers present related models of search and employ micro- or aggregate-level data on search behavior to estimate preferences as well as the costs of searching (Kim, Albuquerque, and Bronnenberg, 2010; De los Santos, Hortaçsu, and Wildenbeest, 2012; Seiler, 2013; Dinerstein, Einav, Levin, and Sundaresan, 2014; Honka, 2014; Koulayev, 2014; Pires, 2014). 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. Moreover, while most of these papers require individual-specific choice data for estimation, our model, like Kim, Albuquerque, and Bronnenberg (2010), can be estimated using market level data. However, whereas Kim, Albuquerque, and Bronnenberg (2010) use aggregate product search data to estimate their search model, our model requires a search cost shifter but otherwise has similar data requirements as BLP-type discrete choice models of demand. Nevertheless, as shown later, our model can be easily extended to incorporate more detailed search data using micro moments.

Our paper also fits into a broader literature that estimates demand for automobiles, which includes BLP, Goldberg (1995), and Petrin (2002).[3] Recent papers in this literature have studied car dealership locations and how this affects consumer demand and competition. For instance, Albuquerque and Bronnenberg (2012) use transaction level data as well as detailed data on the location of consumers and car dealers to estimate a model of supply and demand and find that consumers have a strong disutility of demand for travel. In a related paper, Nurski and Verboven (2013) 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 distance from a consumer to a car dealer is interpreted as a transportation cost, i.e., distance is treated as a product characteristic that enters directly in the utility function. In contrast, in our paper distance enters as a search cost and as such generates variation in the subsets of cars sampled by consumers. We compare the two approaches in Section 5 and show that the elasticity estimates and markups from the search cost model are quite different from those obtained from the transportation cost model.

[3] See Murry and Schneider (2015) for an overview of studies on the economics of retail markets for new and used cars.

2 Dutch Car Market

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

In this section we use survey data to provide some background information on search behavior in the Dutch car market and to motivate the search model we will develop in Section 3. The data was obtained from TNS NIPO (www.tns-nipo.com), a Dutch survey agency. As part of their ongoing investigation (named "De Nederlandse Automobilist") on the characteristics and behavior of Dutch motorists, over 1,200 car drivers are surveyed every year. These drivers are part of TNS NIPObase, which is a panel of around 200,000 respondents. The dataset contains 2,530 observations—1,297 for the survey carried out in 2010 and 1,233 for the 2011 survey. Our data consists of a subset of the questions in the survey and focuses on two aspects of consumer decision making, namely the 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 which dealers they visited before buying the car, and at which dealers they did a test drive. Finally, the respondents answered questions about their household income, household size, age, whether there are children living in the household, and zip code.

Figure 1 gives a histogram of the number of dealers visited by respondents who bought a new car. We focus on purchases between 2003 and 2008 only, since that period overlaps with the aggregate data we will be using for the main analysis, giving us a total of 1,250 respondents who bought a car between those years, of which 540 were new cars. Note that the number of dealers visited only reflects the number of dealers of *different* brands visited, so if a respondent visited two dealers of the same brand, it is only counted as one visit. The average number of dealers visited for new car purchases is 2, which is slightly below the average of 3 dealers found by Moorthy, Ratchford, and Talukdar (1997) for the United States in the early nineties. Although close to 40 percent of

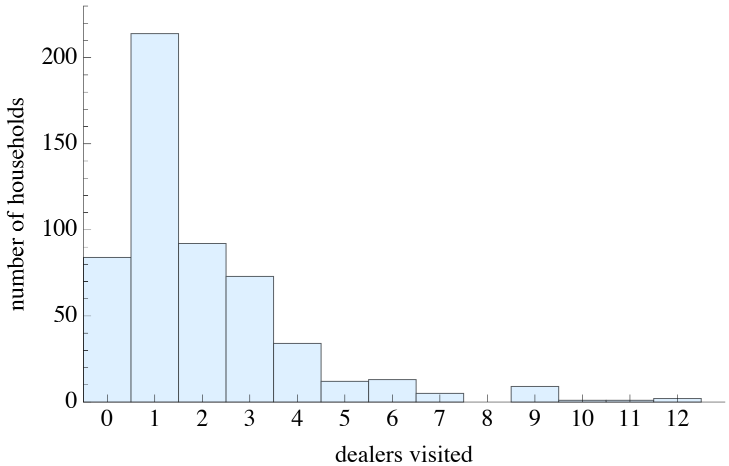


Figure 1: Histogram number of dealers visited for new car purchases

respondents visit at least one dealer of a single brand only, the distribution is positively skewed with some consumers visiting dealers of as many as 12 different brands. Approximately 16 percent of respondents claim they have not visited any dealers.[4]

The evidence obtained from our survey data serves to rule out a model in which consumers have perfect information as an accurate representation of demand. A first natural interpretation of the full information model is that consumers know all they need to know in order to conduct a purchase decision. Under this interpretation, consumers would visit one and only one dealer, and moreover every visit would result in a purchase. Our survey evidence clearly rejects this proposition. In fact, as shown in Figure 1, a substantial number of consumers visits more than one car dealership. Moreover, in 2008 around 44 percent of consumers did not buy a new car conditional on having visited at least one dealer that year.

An alternative interpretation of the full information model is that even though consumers do not observe all relevant car characteristics before visiting the dealerships, consumers can visit them and learn the utility they derive from the various cars at zero cost. Under this interpretation, consumers would then visit dealers of all brands in the market before conducting a purchase. The survey data also rejects this argument because, although some respondents visited as many as 12 dealers before having bought their car, not a single consumer visited a dealer of each of the 38 brands in the market.

[4]Possible explanations for purchases occurring without any dealer visit are online car purchases, parallel imports, and company car leases. A relatively large proportion of the non-visits are (company) car leases—although only 18 percent of new car purchases in the survey data are company car leases, they represent over 25 percent of the non-visits. Buying a new car online is only possible in the Netherlands since 2006, when the online car dealer nieuweautokopen.nl started operating.

The survey data also serves to rule out an imperfect information model in which the sole purpose of the visits is to compare prices. If the only reason for visiting dealers were price shopping, no test drive would have been observed in the market. The survey gives some information on what consumers do while visiting car dealers. In 45 percent of the dealer visits a test drive was involved. Moreover, among those who visited one or more dealers to shop for a new car, over 75 percent made at least one test drive at one of the visited dealerships. If price were the only relevant characteristic that consumers search for in this market, we would not observe this many test drives being made. Because, arguably, test drives are done to learn car characteristics (including whether the car is a good fit) that can hardly be learnt otherwise, the survey data is consistent with a search model in which the main reason for visiting car dealers is to learn more about the product and not just about its price.

Besides information on dealer visits, the survey data contains demographic information such as zip code, household income, family size, and household composition. To obtain a better insight into what explains the differences in search behavior across the respondents, we run several regressions. We first use the information on dealer visits from the survey to investigate what determines the number of dealer visits. Column (A) of Table 1 gives the results of an ordered probit regression in which we explain the number of dealer visits by the log of household income, a dummy for whether there are kids living in the household, a dummy for whether the partner of the head of household is 65 years or older, and a dummy for whether the respondent purchased a new car. In addition we include year fixed effects as well as fixed effects for the make that was ultimately bought. As shown in the table, the log income coefficient is positive and highly significant. Even though this suggests that higher income leads to more search, this does not necessarily mean that higher income respondents have lower search costs because more wealthy consumers also tend to buy more expensive cars and the benefits from search may be higher for this type of cars. Although only significant at the ten percent level, having children in the household reduces the number of searches, while being older increases the number of searches. The new car dummy indicates that people visit more dealers when buying a new car than when buying a used car, which may reflect the fact that it is less common to buy a used car at a car dealer than it is when buying a new car. In specification (B) we focus on new car purchases only. Although this does not change the income coefficient much, the effect of children in the household is now twice as large, whereas the senior dummy is no longer significantly different from zero.

To see how the physical distance from a respondent to a dealer location affects decisions on whether or not to visit a dealer, in specifications (C) and (D) we regress an indicator for whether

Table 1: Dealer visits

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

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Data is for 2003-2008. All specifications include year and make dummies. Specifications (B) and (D) use data for new car purchases only.

a dealer is visited by a respondent on the same set of covariates as before, as well as the Euclidian distance between the centroid of the zip code where the respondent resides and the nearest dealer of each of the car brands in our data. Specification (C) takes all car purchases into account, whereas in specification (D) we only focus on new car purchases. The effects of income, kids, and senior are similar to the results for the ordered probit regressions: income is positively related to a dealer visit, children negatively, and senior positively, although the effect for the latter disappears when we condition on new car purchases. In both specifications distance has a negative impact on the probability of visiting a dealer and is highly significant.

We interpret these results as suggesting that distance from the consumer to a car dealer is related to search frictions. This interpretation is consistent with the fact that, according to the survey, 41 percent of new car buyers responded that distance was a factor they took into account when determining which dealers to visit. That distance matters is also reported in related work. Albuquerque and Bronnenberg (2012), using individual car transaction data in the San Diego metropolitan area between 2004 and 2006, find that consumers have a strong disutility for travel when buying a car. Similarly, Nurski and Verboven (2013) find that dealer proximity is an important determinant of demand for automobiles in Belgium. While these papers interpret distance as a transportation cost parameter that directly lowers utility, we treat it also as a variable that increases the cost of searching cars and creates limited and heterogeneous information.

The above findings suggest search frictions play a role in the car market. Not only is there substantial heterogeneity across respondents in how many dealers they visit, but also consumers tend to visit car dealers to learn more about the characteristics of the cars they sell instead of just for making a purchase or obtaining price information. Moreover, demographics such as income and location seem to play a role in the decision which dealers to visit. These observations lead us to formulate our search model in the next section. Specifically, we develop a discrete choice model with optimal consumer search, in which consumers search for a good fit and in which we relate search costs to consumer demographics including the distance to the nearest car dealer of each brand. Although the model can be estimated using aggregate data only, we estimate it supplemented with some micro moments from the survey data in order to aid identification.

3 Economic Model

3.1 Utility and demand

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

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

where α_i is a consumer-specific price coefficient, the variable p_j denotes the price of car j and the vector $(x_j, \xi_j, \varepsilon_{ij})$ describes different product attributes from which the consumer derives utility. We assume x_j and ξ_j are product attributes the consumer observes without searching, like horse-power, weight, transmission type, ABS, air-conditioning, number of gears, etc. Information on car characteristics and dealership locations can easily be retrieved from for instance advertisements, the Internet, specialized magazines, and consumer reports. The variable ε_{ij}, which is assumed to be independently and identically type I extreme value 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 visiting the dealership, inspecting, and possibly test driving the car, like comfortability, spaciousness, engine noisiness, and gearbox smoothness. It is assumed that the econometrician also observes the product attributes contained in x_j but cannot observe those in ξ_j and ε_{ij}. Consumers differ in the way they value price and product characteristics. The parameter α_i and the expression $(\beta + V_i\sigma)$ capture consumer heterogeneity in tastes for price and product attributes. Here V_i is a diagonal matrix that contains either demographic characteristics or standard normal draws on its main diagonal such that the first component corresponds to the first component of x_j, the second to the second component of x_j, and so on. We let α_i relate to income in the following way:

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

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

The utility from not buying any of the cars is

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

Therefore, we regard product $j = 0$ as the "outside" option; this includes the utility derived from a non-purchase, or the purchase of a used car. 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 cars, so $\mathcal{G}_f \cap \mathcal{G}_g = \varnothing$ for any $f \neq g,\ f, g \in \mathcal{F}$.

We assume consumers must search to find out the exact utility they derive from each of the cars available as well as the utility of the outside option. To be more specific, we assume that before searching consumers know (*i*) the location of each car dealership and the subset of makes and models available at each dealership, (*ii*) car characteristics p_j, x_j and ξ_j for each car j, and (*iii*) the distribution of match values ε_{ij}. Therefore, we regard the process of search of a consumer i as a process by which she discovers the exact values of the matching parameter ε_{ij} upon visiting dealership j.[5]

Consumers are assumed to use a non-sequential search strategy, i.e., they choose which subset of dealers to visit in order to maximize their expected utility; once they have visited the chosen dealers and have learned all the attributes of the cars they are interested in, they decide whether to buy any of the inspected cars or else opt for the outside option. An advantage of a non-sequential search rule is that it allows the searcher to collect information quickly. Although the optimal search strategy often has non-sequential and sequential elements (Morgan and Manning, 1985), a few institutional details of the market for new cars in the Netherlands help justify the non-sequential search assumption. For instance, in this market consumers often need to make appointments with car dealers in order to test-drive their most preferred cars. One motivation for this is simply that the seller, wishing to offer the best possible service, likes to plan test drives such that unnecessary and

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

unpleasant waiting time for his customers is minimized. Moreover, because of space limitations and the associated costs of having cars ready to test-drive, the typical dealer does not have showroom vehicles available for all models at all times. In those cases, test drives need be planned with sufficiently advanced notice. Consumers thus save costs by planning various test drives on the same day, often the single day of the week in which opening hours extend to 9pm as well as Saturdays.

Solving for a consumer's optimal search strategy is difficult. As mentioned in the Introduction, our search problem is related to the general class of search problems discussed in Chade and Smith (2006). In these problems a decision maker must simultaneously choose among a set of ranked stochastic options; each choice is costly and only the best realized option is exercised. When there are many alternatives available in the market, as is the case in our empirical analysis, finding the optimal choice set is extremely complex because there are many choice sets to be evaluated. Chade and Smith (2006) provide a procedure, known as the Marginal Improvement Algorithm (MIA), that finds the solution under some assumptions. In our model, unfortunately, we cannot apply the MIA algorithm. One reason is that the firms in our application sell different numbers of cars and this implies that the various dealer utility distributions cannot a priori be ranked according to the first- or second-order stochastic dominance criterion.[6] Another problem is that, as we have shown in the previous section, an important part of the cost of visiting dealers is the distance from the consumer's home to the different car dealerships. Because not all dealers locate in the same place, the costs of searching cars for a given consumer are brand-specific, which according to Chade and Smith (2006), invalidates the MIA procedure. In order to solve this problem, we add a choice-set specific random term to the costs of searching subsets of dealers that is unobservable to the econometrician. The idea is analogous to the idea of adding an error term to the utility in discrete-choice models. By doing this, we are able to readily compute the probability with which any given subset of dealers is searched by a consumer.

Specifically, we model search costs as follows. Let $\mathcal{S}$ be the set of all subsets of dealers in $\mathcal{F}$ and let S be an element of $\mathcal{S}$. We shall denote consumer i's search cost for visiting all the dealerships in S by c_{iS}. Besides distances we include other variables that potentially capture variation in search costs. One such variable is $|S|$ (i.e., the number of dealers in S) multiplied by log income, which may potentially serve as a proxy for the opportunity cost of time. Other demographic information

[6]The problem is how to compare dealer utility distributions with different means and variances. See also the discussion in Honka (2014).

can be included as well. An example of a search cost specification for consumer i is

$$c_{iS} = \gamma_1 \sum_{f \in S} d_{if} + \gamma_2 \, |S| \log y_i + \gamma_3 \, |S| \, kids_i + \lambda_{iS},$$

where d_{if} denotes the distance between consumer i and dealer f, y_i is consumer i's income, $kids_i$ is an indicator for whether there are children living in the household of consumer i, and λ_{iS} is a consumer-specific search cost shock for visiting a set of dealers S.[7]

We interpret this search cost shock as capturing choice-set specific variation in search costs which we are unable to control for, such as time of the day or day of the week the search occurred, or variation in travel patterns. For instance, we are unable to control for traffic congestion, even though this may affect search costs. Similarly, during the weekend consumers may value time differently than during a weekday, which may lead to variation in choice-set specific search costs that is not captured by our controls.

This choice-set specific error is observed by consumers before searching, but not by the researcher. We shall assume that shocks λ_{iS} are such that $(-\lambda_{iS})$ are i.i.d. type I extreme value distributed across consumers and subsets of dealers. This distributional assumption along with the search cost errors being choice-set specific provides us with a very convenient way to compute the outcome of the optimal search behavior of a consumer.

Regarding this distributional assumption, we note that because the matching term ε_{ij} in the utility also follows a double exponential distribution we are making a double normalization. This is not necessary. We make this assumption because, as we will show later, it allows us to integrate out the choice-set probabilities and obtain a closed-form solution for the probability a car is bought by a given consumer. This closed-form solution allows for a direct comparison to the full information model of BLP, and it is extremely helpful in the estimation phase. Nonetheless, we have experimented with the case in which the scale parameter of the distribution of the choice-set specific errors is not normalized to 1. In such a case, we can no longer compute the choice probabilities in closed form and need to use importance sampling methods to estimate them.[8]

[7] The assumption that the error term is choice-set specific is very convenient because it allows us to compute the choice-set probabilities in closed form no matter the number of available options. When the number of options is small, this is not needed. We compared numerically the case where the λ's are choice-set specific to the case in which they are dealer specific instead. The implied choice probabilities were not very different and the latter case was computationally far more demanding. Details on this comparison are available from the authors upon request.

[8] In Appendix E we discuss how to estimate the model using importance sampling. We have used this method for robustness purposes and found that our qualitative results do not change for other values of the scale parameter. In particular, for the conditional logit case most utility and search cost parameters are not affected by the scale parameter being different from 1 and, correspondingly, the estimated elasticities are very similar. We note however that the model is computationally much more difficult to estimate. We have also tried to estimate the conditional

To simplify the formulas of the choice probabilities, it is convenient to assume consumers always include the outside good in their choice set. Of course, consumers are allowed to pick a choice set that only includes the outside good, i.e., $S = \varnothing$, for a cost $c_{i\varnothing} = \lambda_{i\varnothing}$.[9]

For simplicity of notation let

$$c_{iS} = t'_{iS}\gamma + \lambda_{iS},$$

where

$$t'_{iS} = \left(\sum_{f \in S} d_{if}, \ \ |S| \log y_i, |S| \, kids_i \right)$$

and $\gamma = (\gamma_1, \gamma_2, \gamma_3)'$.

3.2 Optimal non-sequential search

A consumer i first decides which subset of dealers to visit; then, upon visiting the dealers and inspecting and test-driving the cars that are sold at those dealers, she makes a purchase decision. In order to decide which (subset of) dealers to visit, consumer i must compare the expected gains from searching all the possible subsets of dealers. The expected gains to consumer i from searching the dealerships in a subset S are

$$E\left[\max_{j \in \mathcal{G}_f \cup \{0\}, \ f \in S} \{u_{ij}\} \right] - c_{iS},$$

where E denotes the expectation operator, taken in this case over the search characteristics ε_{ij}'s. We now define

$$m_{iS} = E\left[\max_{j \in \mathcal{G}_f \cup \{0\}, \ f \in S} \{u_{ij}\} \right] - t'_{iS}\gamma.$$

Letting F denote the CDF of ε_{ij}, the random variable

$$\max_{j \in \mathcal{G}_f \cup \{0\}, \ f \in S} \{u_{ij}\}$$

logit model with a free scale parameter. The problem we have encountered is that for certain values of the scale parameter the contraction mapping property is violated, so the nested fixed point algorithm breaks down. Extending the model to accommodate micro data appears to make the contraction mapping problem more serious. Currently we are studying algorithms that are able to invert the demand system without the contraction mapping property. Although our results are still very preliminary, we intend to report them as part of future work.

[9] An interpretation of this assumption is that if a consumer i does not search then she does not know ε_{i0}; if this consumer searches some firms then she gets to know ε_{i0} at no additional cost.

has a CDF given by $\prod_{j\in\mathcal{G}_f\cup\{0\},\ f\in S} F(u-\delta_{ij})$, where δ_{ij} is the mean utility consumer i derives from alternative j, i.e., $\delta_{ij} = \alpha_i p_j + x_j'\beta_i + \xi_j$. Using this, we obtain

$$m_{iS} = \log\left(1 + \sum_{f\in S}\exp[\delta_{if}]\right) - t_{iS}'\gamma, \tag{3}$$

where $\delta_{if} = \log\left(\sum_{j\in\mathcal{G}_f}\exp[\delta_{ij}]\right)$.[10]

Since we normalize the mean utility of the outside option to zero, i.e., $\delta_{i0} = 0$, the expected maximum utility of not searching is 0. In our model a consumer can opt for the outside option in two different ways: immediately by deciding not to search at all (i.e., by choosing the empty choice set $S = \varnothing$) or by choosing to visit a number of sellers and deciding to not buy any of the products after inspecting them. In our application in any given year around 93 percent of households do not buy a new car and therefore choose the outside option. Although this group includes consumers who search but do not find a satisfactory option, the vast majority will choose the outside option right away because they are not in the market for a new car that year. To make the model flexible enough that it can accommodate the tradeoffs that both groups of non-buyers face, we allow the value of the outside option when not searching to be different from the outside option after having searched. We do so by letting the expected gains to consumer i from not searching depend on an additional parameter ρ.[11] The expected gains from not searching are then

$$\begin{aligned} m_{i\varnothing} &= E\left[\max_{j\in\{0\}}\{u_{ij}\}\right] + \rho - t_{i\varnothing}'\gamma \\ &= \rho, \end{aligned}$$

[10] Note that

$$\begin{aligned} E\left[\max_{j\in\mathcal{G}_f\cup\{0\},\ f\in S}\{u_{ij}\}\right] &= \int_{-\infty}^{\infty} u\frac{d}{du}\left(\prod_{j\in\mathcal{G}_f\cup\{0\},\ f\in S} F(u-\delta_{ij})\right)du; \\ &= \int u\frac{d}{du}\left(\prod_{j\in\mathcal{G}_f\cup\{0\},\ f\in S}\exp\left[-\exp\left[-(u-\delta_{ij})\right]\right]\right)du; \\ &= c + \log\left(1 + \sum_{j\in\mathcal{G}_f,\ f\in S}\exp[\delta_{ij}]\right), \end{aligned}$$

where c is the Euler constant. So

$$m_{iS} = c + \log\left(1 + \sum_{j\in\mathcal{G}_f,\ f\in S}\exp[\delta_{ij}]\right) - t_{iS}'\gamma.$$

In the expression in equation (3) we omit c because it does not affect choices.

[11] Only the difference in expected utility between searching and not searching is identified, which means that the parameter ρ can also be interpreted as a fixed search cost.

where, as in equation (3), we have omitted the Euler constant because it does not affect choices.

Consumer i will pick the subset S_i that maximizes the expected gain $m_{iS} - \lambda_{iS}$, i.e.,

$$\begin{aligned} S_i &= \arg\max_{S \in \mathcal{S}} [m_{iS} - \lambda_{iS}]; \\ &= \arg\max_{S \in \mathcal{S}} \left[\log \left(1 + \sum_{f \in S} \exp[\delta_{if}] \right) - t'_{iS}\gamma - \lambda_{iS} \right]. \end{aligned}$$

Since we assume $(-\lambda_{iS})$ is i.i.d. type I extreme value distributed, the probability that consumer i finds it optimal to sample the set of dealers S_i is P_{iS_i}, where

$$P_{iS} = \frac{\exp[m_{iS}]}{\sum_{S' \in \mathcal{S}} \exp[m_{iS'}]}. \tag{4}$$

Given that consumer i searches the set S_i, the probability that she buys product j is equal to the probability that product j provides the highest utility out of the products of the firms in S_i. Denoting this probability by $P_{ij|S_i}$, we have

$$P_{ij|S} = \frac{\exp[\delta_{ij}]}{1 + \sum_{r \in S} \exp[\delta_{ir}]}.$$

In order to obtain the unconditional probability s_{ij} that consumer i purchases product j, we need to "integrate out" S_i from this probability, i.e.,

$$\begin{aligned} s_{ij} &= \sum_{S \in \mathcal{S}_f} P_{iS} P_{ij|S} \\ &= \sum_{S \in \mathcal{S}_f} \frac{\exp[m_{iS}]}{\sum_{S' \in \mathcal{S}} \exp[m_{iS'}]} \frac{\exp[\delta_{ij}]}{1 + \sum_{r \in S} \exp[\delta_{ir}]}, \end{aligned}$$

where f denotes the firm producing j and $\mathcal{S}_f$ is the set of all choice sets containing firm f. In Appendix A we show this probability can be written as

$$s_{ij} = \frac{\exp\left[\delta_{ij} - \ln\left(1 + \exp\left[t'_{i\{f\}}\gamma\right]\right)\right]}{\dfrac{1-\phi}{\phi \Pi_{iF}} + 1 + \sum_{k=1}^{J} \exp\left[\delta_{ik} - \ln\left(1 + \exp\left[t'_{i\{g\}}\gamma\right]\right)\right]}, \tag{5}$$

where $t_{i\{f\}}$ contains the search cost of firm f only, $\phi = \exp[-\rho]$, and $\Pi_{iF} = \prod_{g \in F} \left(1 + \exp\left[-t'_{i\{g\}}\gamma\right]\right)$. This equation suggests that our search model nests the full information model of BLP. In fact, we get the full information choice probability when $\gamma \to -\infty$ (because the support of the search cost distribution is $(-\infty, \infty)$); in such a case, $\exp\left[t'_{if}\gamma\right] \to 0$, $\Pi_{iF} \to \infty$ and we obtain

$$s^f_{ij} = \frac{\exp\left[\delta_{ij}\right]}{1 + \sum\limits_{k=1}^{J} \exp\left[\delta_{ik}\right]}.$$

Equation (4) can also be used to calculate the probability consumer i searches k times for low values of $k = 1, 2, 3, \ldots, F$, which is useful for constructing search-related micro moments. Let $M_i = \sum_{S' \in \mathcal{S}} \exp[m_{iS'}]$, the denominator of P_{iS} in equation (4). As shown in Appendix A, this can be simplified to

$$M_i = \Pi_{iF}\left(\frac{1-\phi}{\phi\Pi_{iF}} + 1 + \sum_{k=1}^{J} \frac{\exp[\delta_{ik}]}{1 + \exp[t'_{i\{g\}}\gamma]}\right).$$

Consumer i's probability of not searching is then $p_{i0} = \exp\left[m_{i\varnothing}\right]/M_i = \exp\left[\rho\right]/M_i$, whereas the probability of searching once is $p_{i1} = \sum_{f \in \mathcal{F}} \exp\left[m_{i\{f\}}\right]/M_i$, where $m_{i\{f\}}$ is given by equation (3) for the singleton subset that only contains firm f. Similarly, the probability of searching two firms is $p_{i2} = \sum_{f=1}^{F-1} \sum_{g=f+1}^{F} \exp\left[m_{i\{f,g\}}\right]/M_i$.

Let τ_i be the vector of all consumer-specific random variables in s_{ij}, that is, parameters and demographic characteristics. Then the probability that product j is purchased is the integral

$$s_j = \int s_{ij}\, dF_\tau(\tau_i), \tag{6}$$

where F_τ is the CDF of τ_i. Such an integral is difficult to compute analytically but it can be estimated by Monte Carlo simulations by drawing the demographic characteristics and random coefficients of, say, N consumers, and then computing s_{ij} for each consumer $i = 1, \ldots, N$. The Monte Carlo estimator of s_j is then taken as the sample mean $\widehat{s}_j = \frac{1}{N}\sum_{i=1}^{N} s_{ij}$.

3.3 Supply side

We include the supply side in order to obtain estimates of price-cost markups. We assume each firm $f \in \{1, \ldots, F\}$ supplies a subset $\mathcal{G}_f$ of the J products. Let $\mathcal{M}$ denote the number of consumers and let mc_j denote the marginal cost of producing product j. Then the profit of firm f is given by

$$\Pi_f(p) = \sum_{j \in \mathcal{G}_f} (p_j - mc_j)\mathcal{M}s_j(p).$$

Following BLP we assume mc_j depends log-linearly on observed product characteristics affecting cost, w_j, and an unobserved cost characteristic ω_j:

$$\ln(mc_j) = w_j'\eta + \omega_j. \tag{7}$$

We expect the unobserved cost characteristics ω_j to be correlated with the unobserved demand characteristics ξ_j. For instance, if the researcher does not observe whether a car has a navigation system as standard equipment, then cars having this characteristic will have a higher unobserved demand characteristics and, because it is more costly for the firm to include a navigation system, a higher unobserved cost characteristics as well. We will account for this correlation in the estimation procedure.

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 dealers.[12] Let p denote the vector of Nash equilibrium prices. Assuming a pure strategy equilibrium exists for this game, any product j should have a price that satisfies the first order condition

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

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

$$p - mc = \Delta(p)^{-1} s(p), \tag{8}$$

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

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

For the derivation of the partial derivatives of the market shares with respect to price it matters whether consumers observe deviation prices before or after search. In our context, because consumers can easily observe list prices while being at home, we adopt the assumption that consumers observe (deviation) prices before they start searching. Notice that this assumption differs from most of the literature on consumer search for differentiated products (Wolinsky, 1986; Anderson and Renault, 1999) and, as demonstrated in recent work (see, e.g., Armstrong and Zhou, 2011; Haan

[12] Although we do not model optimal dealer locations, we do not regard it as exogenous because we do not use distances as instruments.

and Moraga-González, 2011), this assumption has implications for the behavior of prices and search costs.[13]

4 Estimation and identification

Our estimation procedure closely resembles BLP, except that we allow for an endogenous choice-set selection stage which is the outcome of an optimal consumer search problem. As shown by BLP, the parameters of the demand and supply model without search frictions can be estimated by generalized method of moments (GMM). Their GMM procedure accounts for price endogeneity by solving for the unobservables ξ_j and ω_j in terms of the observed variables and taking these as the econometric error term of the model. As in BLP, we can compute the vector $\xi = (\xi_1, \ldots, \xi_J)$ of unobserved characteristics as the unique fixed point of a contraction mapping. The fact that this mapping is indeed a contraction 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 B for more details). In this section we provide a method to estimate the search model by GMM as well.

4.1 Moments

We consider macro- and micro moments. For the macro moments, following BLP, the predicted market share $s_j(\theta)$ of product j should match observed market shares $\hat{s}_j$, or

$$s_j(\delta(\theta), \theta) - \hat{s}_j = 0.$$

We use the contraction mapping mentioned in the previous paragraph to solve for $\delta(\theta)$. The first moment unobservable follows from $\delta(\theta)$ and is

$$\xi_j = \delta_j(\theta) - x_j\beta. \tag{9}$$

[13] In a standard search model a firm chooses its price to maximize the payoff from the consumers who visit. By changing the price a firm thus affects the selling probability, but not the visiting probability. In contrast, when prices are observed from home, changing the price affects both the visiting and the buying probability in our model. While in most standard models prices increase in search costs, in models where prices are observable before search, prices can be decreasing in search costs (Haan and Moraga-González, 2011). We will return to this point later in the paper when we study the equilibrium effects of changing search costs.

The second moment unobservable follows from the parametric marginal cost specification and the first order conditions—combining equations (7) and (8) and solving for ω_j gives

$$\omega_j = \ln\left(p - \Delta^{-1} s(\theta)\right) - w_j' \eta. \tag{10}$$

Since we have individual-level data, we are able to construct micro moments as well. The survey data provides for each respondent information on their latest car purchase as well as their search behavior related to that car purchase. We supplement the BLP moments with micro moments that are based on these survey responses as well as corresponding model predictions—following Petrin (2002) we let the GMM estimation routine select the parameters of the model such that the predicted probabilities match the observed probabilities in the survey data. We discuss some specific examples of micro moments in Section 5. For computational details we refer to Appendix C.

4.2 GMM estimation

We use GMM to estimate the model. The original estimation relies on the assumption that the true values of the demand and cost unobservables are mean independent of observed product characteristics, that is,

$$E[(\xi_j, \omega_j)|(X, W)] = 0.$$

Let Z be a matrix of instruments with $2J$ rows and let $\psi(\theta) = (\xi_1(\theta), \ldots, \xi_J(\theta), \omega_1(\theta), \ldots, \omega_J(\theta))'$. Let the sample moments be

$$g_J(\theta) = \frac{1}{2J} Z' \psi(\theta). \tag{11}$$

Denote the column vector of micro moments by $g_N(\theta)$; let

$$g(\theta) = \begin{pmatrix} g_J(\theta) \\ g_N(\theta) \end{pmatrix}$$

be the column-vector of both macro- and micro moments.

The GMM estimator of θ is

$$\hat{\theta} = \arg\min_{\theta} g(\theta)' \Xi g(\theta),$$

where Ξ is a weighting matrix; see Appendix D for various alternatives for this. Some parameters enter linearly in the model, so we can concentrate them out of the above GMM minimization. By using equations (9) and (10) and denoting

$$X_1 = \begin{pmatrix} x & 0 \\ 0 & w \end{pmatrix} \quad \text{and} \quad \delta\left(\theta_2\right) = \begin{pmatrix} \delta_D\left(\theta_2\right) \\ \delta_C\left(\theta_2\right) \end{pmatrix}$$

we get

$$\psi\left(\theta\right) = \delta\left(\theta_2\right) - X_1\theta_1.$$

By assuming θ_2 known we can obtain θ_1 as the linear IV estimator

$$\widehat{\theta}_1 = \left(X_1' Z \Xi Z' X_1\right)^{-1} X_1' Z \Xi Z' \delta\left(\theta_2\right),$$

and substituting this in $g_J\left(\theta\right)$ we obtain a new sample moment, which is a function of θ_2 only

$$\overline{g}_J\left(\theta_2\right) = \frac{1}{J} Z' \overline{\psi}\left(\theta_2\right), \quad \text{where} \quad \overline{\psi}\left(\theta_2\right) = \delta\left(\theta_2\right) - X_1\left(X_1' Z \Xi Z' X_1\right)^{-1} X_1' Z \Xi Z' \delta\left(\theta_2\right).$$

Note that by using $\delta_D\left(\theta_2\right)$ in $s_{ij}\left(\theta\right)$, the micro moment (see equation (A16) in the Appendix) does not depend on θ_1; denote by $\overline{g}_N\left(\theta_2\right)$ the vector of micro moments that are the same as $g_N\left(\theta\right)$ but formally depend on $\delta_D\left(\theta_2\right)$ instead of ξ and θ_1. Also, let

$$\overline{g}\left(\theta\right) = \begin{pmatrix} \overline{g}_J\left(\theta\right) \\ \overline{g}_N\left(\theta\right) \end{pmatrix}.$$

The GMM estimator of θ_2 based on this is

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

4.3 Identification

In this section we provide an informal discussion on identification. In our model, variation in the sales across brands is due to (i) variation in the characteristics of cars, (ii) variation in the unobserved consumer characteristics, which also include demographic characteristics like income, number of children, and distances from the households to the closest dealers of the brands, and (iii) variation in the observed demographic characteristics and choices by the consumers in the survey. We are interested in the identification of the parameters in the utility function, which are α, β and σ, the marginal cost parameter vector η, the search cost parameter vector γ and the gains from not searching ρ.

Usually the α and β parameters of the utility function can be identified because the econometrician observes different market shares corresponding to different product characteristics. Further, the σ parameter vector that is responsible for unobserved consumer heterogeneity can be identified due to the fact that certain types of consumers prefer cars with certain characteristics. In our model, since we can control for the distances from the households to the dealerships, the parameters β and σ can be identified in the same way. Although these parameters are identified from aggregate data, the observed consumer-level characteristics, choices, and choice sets aid their identification, as shown by Petrin (2002). The marginal cost parameter vector η is identified from variation in prices and market shares with respect to the observed marginal cost variables.

The set of instruments we use to control for possible correlation between unobserved characteristics and price as well as unobserved characteristics and distances are similar to those used by BLP. That is, in addition to product characteristics, which are exogenous by assumption, we add the number of cars and the sum of characteristics of the cars produced by the same firm, as well as the number of competing cars and the sum of characteristics of all competing cars. It is known from Armstrong (2014) that in large markets like the car market, the instruments constructed in this way are weak and may not identify the price coefficients; BLP deal with this problem by using sales data from 20 years. Since in addition to aggregate sales data from 6 years we use consumer-level choice and choice-set data, the weak instruments problem is solved because at the level of consumers prices can be regarded as exogenous, so they need not be instrumented in the micro moments. This is because in the micro moments we can treat the unobserved characteristics as observed by the econometrician since they can be computed from the market shares.

The parameter vector γ, which appears in our formulation of search costs, can be identified from aggregate data due to nonlinearities (see equation (5)). The variables involved in the search cost specification are distance, income, and “kids,” that is, a dummy variable indicating whether there are children in the household. In our most complete model (see specification (B) in Table 5) these variables also appear in the utility either directly (distance) or in combination with other variables (income and “kids”). Again, consumer-level data reinforces the identification of γ. For example, as it can be seen from the survey probabilities in Table 6, the probability of buying a new car for high income consumers (i.e., $y_i \geq \overline{y}$) is roughly the double of the same probability for low income consumers (0.105 versus 0.053). At the same time, among the low income consumers there are more who search at least twice than those who buy (0.061 versus 0.053), while among the high income consumers there are fewer who search at least twice than those who buy (0.096 versus 0.105). This means that income affects search costs negatively.

Finally, the gains from not searching ρ can be identified from aggregate data due to its nonlinear connection to the constant in the utility (see equation (5)). Consumer-level data provides additional information on ρ because the probability of not searching p_{i0} (see end of Section 3.2) is strictly increasing in ρ and in the survey we observe the proportion of consumers who do not search.[14]

5 Data and Results

5.1 Data

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

The data on product characteristics are obtained from *Autoweek Carbase*, which is an online database of prices and specifications of all cars sold in the Netherlands from the early eighties until now.[15] Characteristics include horsepower, number of cylinders, maximum speed, fuel efficiency, weight, size, and dummy variables for whether the car's standard equipment includes air-conditioning, power steering, cruise control, and a board computer. Unfortunately transaction prices are not available, so all prices are listed (post-tax) prices.[16] We have used the Consumer Price Index to normalize all prices to 2006 euros.

We have supplemented the data set 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 market shares (calculated as sales divided by the number of

[14]Note that

$$p_{i0} = \frac{1}{\phi\Pi_{iF}\left(\frac{1-\phi}{\phi\Pi_{iF}} + 1 + \sum_{k=1}^{J}\frac{\exp[\delta_{ik}]}{1+\exp[t'_{i\{g\}}\gamma]}\right)} = \frac{1}{1+\phi\left[\Pi_{iF}\left(1+\sum_{k=1}^{J}\frac{\exp[\delta_{ik}]}{1+\exp[t'_{i\{g\}}\gamma]}\right)-1\right]},$$

which is decreasing in ϕ because the expression in the square brackets is positive. Therefore, p_{i0} is strictly increasing in ρ.

[15]See http://www.autoweek.nl/carbase.

[16]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.

households), while average gasoline prices are used to construct our kilometers per euro (KM€) 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 40 different brands in our sample are owned by 17 different companies. For instance, in 2007 Ford Motor Company owned Ford, Jaguar, Land Rover, Mazda, and Volvo.

Table 2: Summary statistics

Year	No. of Models	Sales	Price	European	HP/ Weight	Size	Cruise Control	KPL	KP€	Family Car	MPV or SUV
2003	213	481,913	19,562	0.762	0.787	7.153	0.229	14.480	12.497	0.426	0.216
2004	228	476,581	19,950	0.749	0.788	7.184	0.308	14.696	11.737	0.408	0.235
2005	233	457,897	20,540	0.727	0.794	7.270	0.301	14.861	10.987	0.403	0.242
2006	231	475,636	20,367	0.715	0.804	7.271	0.308	15.120	10.707	0.361	0.234
2007	236	495,091	20,509	0.712	0.810	7.330	0.281	15.112	10.356	0.363	0.240
2008	241	489,584	18,613	0.714	0.813	7.271	0.293	15.813	10.290	0.381	0.188
All	1,382	479,450	19,916	0.730	0.799	7.247	0.286	15.018	11.091	0.390	0.226

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

Table 2 gives the sales weighted means for the main variables we use in our analysis. The number of models has increased from 213 in 2003 to 241 in 2008. Sales were lowest in 2005 and peaked in 2007. Prices have been going up mostly in real terms, although 2008 saw a sharp decrease, possibly as a result of the onset of a recession. The share of European cars sold shows a downward trend, mainly to the benefit of cars that originate from East Asia. The ratio of horsepower to weight has been increasing steadily. The share of cars with cruise control as standard equipment increased in the first half of the sampling period, but then decreased somewhat. Cars have become more fuel efficient during the sampling period. Nevertheless, as shown in the KP€ column of Table 2, fuel efficiency has not increased enough to offset rising gasoline prices—the number of kilometers that can be traveled for one euro has decreased over the sample period. The share of family cars has steadily declined over time, while cars that are classified as either multi-purpose vehicle (MPV) or sport utility vehicle (SUV) saw their market shares increase until 2007, followed by a sharp drop in 2008. During this period the share of cars classified as supermini went up from 32 percent to over 40 percent of the market.[17]

[17]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, as well as the small and large off-road 4x4 classes into a single MPV/SUV class (combined market share of 0.226). The combined market share of cars in other

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

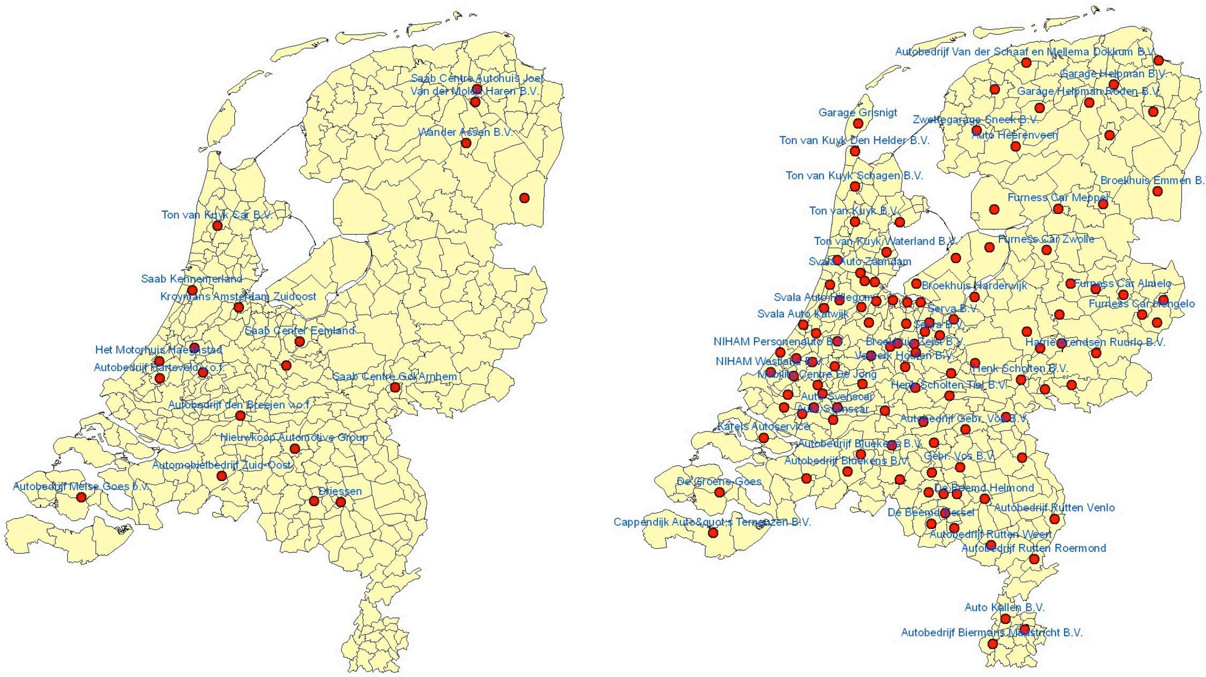
Our demographic and socioeconomic data on households are obtained from Statistics Netherlands. These data are available at various levels of regional disaggregation (neighborhoods, districts, city councils, counties, provinces, etc.). Since the purpose of our study is to estimate the importance of search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

For every neighborhood, the demographic data include the number of inhabitants and their distribution by age groups, the number of households, the average household size, the proportion of single-person households, and the proportion of households with children. The socioeconomic data include the average home value, the average income per inhabitant and income earner, as well as the total number of cars and their ownership status (company leased versus privately owned). We only include neighborhoods with a strictly positive number of inhabitants, which leaves us with a total of 11,122 neighborhoods for 2007.[18] Most neighborhoods are relatively small; the mean number of inhabitants is 1,471.

In addition to demographic data we have information on the exact location of each neighborhood on the map of the Netherlands. Using a geographical software package we use this information to construct a proxy for the distance that needs to be travelled when visiting a car dealership. To be able to do this, for every brand we have first obtained the addresses of all its dealerships in the Netherlands. For instance, Saab has a total of twenty dealers in the Netherlands, which are spread over the country as shown in Figure 2(a). Since we have the exact addresses of the twenty dealerships of Saab, for every neighborhood, we can compute the Euclidean distance from the center of the neighborhood to the closest Saab dealer. We do this for all car manufacturers and obtain a matrix of 11,122 by 38 containing the minimum distances from the center of a neighborhood to a car dealer.

classes (executive, luxury, sports cars, and vans) is 0.037.

[18]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).



(a) Saab dealerships (b) Volvo dealerships

Figure 2: Locations selected dealers

There is a lot of variation in the distances to the closest dealer of each brand across neighborhoods. Figure 2(b) gives the spread of Volvo dealerships across the country—clearly on average the minimum distance to a Volvo dealer is much smaller than the minimum distance to a Saab dealer. A similar picture arises for other brands. For instance, Audi has 161 dealers, whereas BMW has only 20, even though both brands are active in the luxury segment of the market and both have similar sales figures. Table 3 gives some descriptive statistics for the distances to the nearest dealer for all the car brands in our data. Opel is the most accessible: almost 79% of all households live within 5 kilometer from an Opel dealer. Porsche has the lowest percentage of households within 5 kilometer: only 6.3% of households is within easy reach.

Since car dealers may be located in areas where there is most demand for the cars they are selling, the location of car dealers may be endogenous to the model; this, if not properly dealt with, may result in biased estimates of the distance coefficient. For instance, if dealers of luxury brands are located near wealthy neighborhoods, while mainstream brands are not, then the effect of distance may seem more important than it is. In order to explore whether this is an issue in our data, in the last two columns of Table 3 we compute the weighted average distances to the dealers for the top ten percent of neighborhoods in terms of average income. The table indicates that the weighted average distances to dealers of luxury brands such as Audi, BMW, Mercedes,

Table 3: Descriptive statistics for distances

Brand	Number of dealerships	Number of cars sold in 2008	All neighborhoods: Weighted average distance	All neighborhoods: Percentage of households within 5 km	Top income decile: Weighted average distance	Top income decile: Percentage of households within 5 km
Alfa Romeo	75	3,050	7.98	42.1	6.00	46.5
Audi	161	16,738	4.70	67.6	3.87	76.7
BMW	20	15,170	17.11	16.8	11.47	31.1
Cadillac	15	198	19.01	14.1	12.82	21.7
Chevrolet/Daewoo	137	7,421	5.00	64.9	3.98	76.6
Chrysler/Dodge	20	2,589	17.94	16.8	12.23	24.1
Citroën	162	24,139	4.40	69.9	3.42	83.2
Dacia	20	4,549	31.21	19.2	18.14	36.7
Daihatsu	99	9,186	6.26	52.9	5.24	59.2
Fiat	142	21,010	4.86	66.6	4.05	74.1
Ford	233	42,504	3.67	78.0	3.10	87.6
Honda	20	8,479	19.11	15.6	12.77	22.3
Hyundai	138	17,433	5.10	63.1	4.16	69.0
Jaguar	16	752	18.40	13.9	11.14	21.9
Jeep	20	784	17.94	16.8	12.23	24.1
Kia	115	12,236	5.70	54.5	4.95	56.2
Lancia	20	761	17.28	17.0	11.01	27.6
Land Rover	20	1,421	14.45	17.0	9.38	24.1
Lexus	13	1,044	19.57	16.4	13.08	25.0
Mazda	121	7,582	5.59	57.8	4.40	69.8
Mercedes-Benz	83	10,446	6.59	47.9	5.05	54.6
Mini	20	3,417	17.34	18.5	11.08	34.8
Mitsubishi	108	7,805	5.59	53.8	5.11	51.3
Nissan	114	10,259	6.03	51.6	5.19	59.1
Opel	233	40,405	3.55	78.7	3.09	86.8
Peugeot	187	40,250	4.14	72.9	3.36	83.7
Porsche	8	531	25.78	6.3	16.77	11.1
Renault	197	37,526	4.19	70.8	3.40	81.7
Saab	20	1,938	20.10	14.4	11.78	23.7
Seat	127	13,061	6.06	57.7	4.63	67.2
Skoda	97	9,461	6.14	51.3	4.84	59.3
Smart	20	952	14.37	20.2	10.33	23.7
Subaru	20	1,422	19.10	16.4	12.82	29.4
Suzuki	124	14,547	5.04	59.0	4.65	59.1
Toyota	141	38,997	4.70	66.7	4.10	70.8
Volkswagen	188	45,034	4.04	74.1	3.43	84.4
Volvo	114	16,487	5.34	61.5	4.14	73.3

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

and Lexus is indeed lower for the ten percent richest neighborhoods than for all neighborhoods. In addition, the percentage of households within 5 kilometer is higher in the top decile in comparison to all neighborhoods taken together. While this is suggestive of luxury brands locating closer to neighborhoods where there is most demand for their cars, Table 3 also shows that these two facts are also true for the rest of the brands, not just for the upscale brands. In fact, for both luxury and mainstream brands, the weighted average distance for the top decile neighborhoods is on average about two-thirds of that of all neighborhoods together. If endogeneity of location were a serious issue then one would not expect to find a similar relation for the non-luxury brands. Nonetheless, the possible endogeneity of location can in principle be addressed by including a rich set of interactions between car characteristics and demographics (see also Nurski and Verboven, 2013). Moreover, we do not use distance as an instrument, and therefore do not treat the location of dealers as exogenous during the estimation of the model.

Our last dataset, discussed in Section 2, is obtained from two separate surveys that were administered by TNS NIPO, a Dutch survey agency, in 2010 and 2011. The focus of the survey is on characteristics and the behavior of Dutch car owners. Each of the in total 2,530 respondents that participated in the survey has answered specific questions on which dealers were visited in relation to the most recent purchased car, which provides useful information on how consumers search in this market. In addition, we have information about the respondent's household income, household size, age, kids, and zip code. We use these data for the estimation of the micro moments. We exclude respondents for which we do not observe income or a zip code, which leaves us with 2,024 observations. For the micro moments we focus on new car purchases in 2008 only—we assume that all respondents that did not buy a new car in 2008 went for the outside option, which includes not buying a car and buying a used car. According to data from the survey, slightly over 7 percent of the respondents bought a new car in 2008, which equals the share of households in the Netherlands that bought a new car in 2008.

5.2 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. We first show results for the conditional logit model. The advantage of the logit model is that it allows us to explore the effects of search frictions in a very simple setting. We next estimate a more complex model in which we estimate the supply side alongside the demand model and allow for random coefficients. In this specification we also use moments that are based on individual-level data from the survey, which will help identification (Armstrong, 2014) and can improve the precision of the estimates (Petrin, 2002).

Conditional logit model

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

$$u_{ij} = \alpha p_j + x_j' \beta + \xi_j + \varepsilon_{ij}. \tag{12}$$

As a benchmark case, in the first two columns we present the demand estimates for a model without search frictions. The results in column (A) are obtained by regressing $\ln(s_j) - \ln(s_0)$ on product characteristics and price using ordinary least squares. The results in column (B) are obtained by using an instrumental variables (IV) approach to control for possible correlation between unobserved characteristics and price. As in most of the previous literature, we assume the car characteristics to be exogenous. Since markups relate to how far away other car models are in the characteristics space, this means prices will be correlated with the characteristics of other car models and these characteristics can be used as instrumental variables.[19]

Except for horsepower per weight and the family car dummy, in both specifications all parameter estimates have the expected sign and are significant—horsepower per weight has an unexpected negative impact in the OLS specification, while the family car dummy is not significantly different from zero in the IV specification in column (B). The price coefficient increases in magnitude in the IV specification, which is to be expected given that higher unobserved quality components should lead to higher prices. As a result the number of products with inelastic demands decreases substantially, from 97 percent of all cars for the estimates in column (A) to less than 15 percent in column (B). 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.). Size, a higher mileage per euro, and cruise control as standard equipment all affect the consumers' mean utility in a positive way. Finally, consumers consider MPVs and SUVs to have a positive marginal utility.

In the last column of Table 4 we present the demand estimates using our consumer search model. We only use distance as a search cost shifter—we relate the search cost parameter γ 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 closed form solution for the market share equations, so we proceed by simulating buying probabilities. Specifically, we randomly draw 2,209 neighborhoods from the demographic data, where each neighborhood is weighted by number of

[19]Specifically, as instruments we use own product characteristics, the number of other cars produced by the firm, the number of cars produced by rival firms, the sum of product characteristics (cruise control, fuel efficiency, and whether the car is a family car or MPV/SUV) of other cars produced by the firm, as well as the sum of the same product characteristics of cars produced by rival firms.

Table 4: Estimation results conditional logit

	Full information				Search	
	OLS Logit Demand (A)		IV Logit Demand (B)		GMM/IV Logit Demand (C)	
Variable	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
Preference parameters						
constant	-12.111	(0.705)***	-15.577	(1.208)***	-14.109	(1.163)***
HP/weight	-1.049	(0.237)***	1.908	(0.822)**	1.745	(0.829)**
non-European	-0.512	(0.074)***	-0.951	(0.142)***	-0.879	(0.144)***
cruise control	0.101	(0.087)	0.289	(0.110)***	0.176	(0.114)
fuel efficiency	2.671	(0.272)***	2.348	(0.317)***	2.307	(0.279)***
size	2.386	(0.634)***	6.736	(1.348)***	5.910	(1.319)***
family car	0.266	(0.097)***	-0.248	(0.174)	-0.159	(0.170)
MPV/SUV	0.276	(0.100)***	0.259	(0.112)**	0.238	(0.106)**
price	-0.012	(0.004)***	-0.082	(0.019)***	-0.070	(0.019)***
Search cost parameters						
distance	—		—		0.070	(0.032)**
R^2	0.344		n.a.		n.a.	
Objective function	n.a.		n.a.		0.021	

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. The number of observations is 1,382. Standard errors are in parenthesis. The number of simulated consumers used for the estimation of specification (C) is 2,209.

inhabitants.[20] Next we use the distances to the nearest dealer for each of the brands in our sample to simulate search behavior for the 2,209 selected "consumers." We estimate the model by GMM, using the same instruments as when estimating specification (B), and use $(Z'Z)^{-1}$ as the weighting matrix. The results shown in column (C) of Table 4 show that search costs are positively related to distance and significantly different from zero at the five percent level. 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 ignoring search frictions may result in an overestimation of consumer price sensitivity. We will come back to this when discussing the results for the complete model below.

Estimation results complete model

The demand side estimates for the complete model are based on the utility function in equation (1). We use the same attributes as those shown in Table 4 for the estimation of the simplified model. As before, we estimate the mean marginal utility of each of these attributes, but now we allow the marginal utility for some of the attributes to differ across consumers by estimating

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

a variance term for these attributes. Specifically, if a specific car characteristic has a random coefficient, for all characteristics except for "family car" we use a standard normal draw for the corresponding component of the diagonal matrix V_i. For "family car" the corresponding component of the diagonal matrix is (kids) $\times\, \upsilon_{3i}$, where kids is a dummy for whether there are children present in the household and υ_{3i} is a $\chi^2(3)$-distributed draw truncated at 95%. An advantage of using this distribution is that it is bounded and skewed toward positive taste (see also Petrin, 2002). We allow for heterogeneity in the price parameter in accordance with equation (2). This means that in addition to normalizing prices by household income, we allow α to differ according to income groups (see also Petrin, 2002). A simulated consumer's household income is randomly drawn from a log-normal distribution with scale parameter 0.28 (which is estimated outside the model) and neighborhood-specific location parameter such that the mean (after-tax) household income level in the neighborhood where the simulated consumer resides matches the neighborhood data from Statistics Netherlands. The income bound $\overline{y}$ we use corresponds to a household income (after tax) of €31,000.[21] The kids dummy is obtained from the neighborhood-specific percentage of households with children, i.e., the kids dummy equals 1 if that percentage is larger than a uniform draw on $(0, 1)$ and zero otherwise.

The estimation of the supply side is based on equation (7), where we use the fact that model j's marginal cost mc_j equals the difference between its price and markup $\Delta(p)^{-1}s(p)$. Our cost-side variables are based on the attributes in the utility function and include a constant, indicators for non-European and cruise control, and the natural logarithm of HP/weight, kilometers per liter, and size.

We use two sets of micro moments. The first relates demographic information to buying decisions. Specifically, we use two micro moments that match the model's predicted average probability of buying a new car conditional on income level to the survey data:

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

where $\mathbb{1}\{i \text{ purchases new vehicle}\}$ is an indicator for the event that consumer i purchases a new vehicle and $\{y_i < \overline{y}\}$ and $\{y_i \geq \overline{y}\}$ correspond to the events that consumer i is in the low or high income group, respectively. We also use two micro moments that relate the model's predicted

[21]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 €47,600 before taxes, which corresponds to one of the cutoffs used to create bins in the survey data.

probability of buying a larger car (MPV/SUV or family car) conditional on either family size or having children in the household to those observed in the survey data:

$$E[\mathbb{1}\{i \text{ purchases an MPV or SUV}\} \mid \{i \text{ has a family size } \leq 3\}],$$
$$E[\mathbb{1}\{i \text{ purchases a family car}\} \mid \{i \text{ has children present in the household}\}].$$

The second set of micro moments matches the model's predicted average probability of searching at least twice conditional on income level and the number of dealers nearby to the survey data:

$$E[\mathbb{1}\{i \text{ searches at least twice}\} \mid \{y_i < \overline{y}\}],$$
$$E[\mathbb{1}\{i \text{ searches at least twice}\} \mid \{y_i \geq \overline{y}\}],$$
$$E[\mathbb{1}\{i \text{ searches at least twice}\} \mid \{\text{number of dealers within 10km of } i < 15\}],$$
$$E[\mathbb{1}\{i \text{ searches at least twice}\} \mid \{\text{number of dealers within 10km of } i \geq 15\}],$$

We use the approximation to the optimal weighting matrix as specified in Appendix D. To obtain a first-consistent estimator of θ^0, denoted by $\hat{\theta}$, we use

$$\Xi = \begin{pmatrix} (Z'Z)^{-1} & 0 \\ 0 & I_M \end{pmatrix}$$

where I_M is the identity matrix of dimension M, which is the number of micro moments.

By including micro moments we can estimate a richer search cost specification than the one used in the estimation of the conditional logit model. Specifically, in addition to distance we let search costs depend on the logarithm of household income as well as a dummy for whether there are children present in the household, where both are multiplied by the number of searches. In addition, we estimate the parameter ρ, which can be interpreted as the expected value of not searching.

The estimation results for the search model are presented in the first two columns of Table 5. In specification (A) we allow for a random coefficient on the interaction of family car and the kids dummy as well as a random coefficient on the constant.[22] The price coefficients are statistically significant at high significance levels. All else equal, consumers prefer larger and more

[22] Although not reported, we have estimated versions of the model with more random coefficients. Even though most of the estimated base coefficients change somewhat as a result of allowing for more random coefficients, the price and search cost parameter estimates appear robust to changes in the number of random coefficients. We prefer a specification with less random coefficients since this increases the precision of the estimates. Moreover, with a full set of random coefficients it is very difficult to do counterfactual exercises due to numerical issues when solving for the price equilibrium (see also Skrainka, 2012).

powerful cars: consumers put a positive value on the size of a car and on horsepower per weight. The estimated coefficient of the non-European dummy is negative and highly significant, which, consistent with the results for the conditional logit model, indicates consumers on average prefer cars produced by European firms versus cars produced by non-European firms. We use cruise control as a measure of luxuriousness; as expected, consumers put a positive value on cruise control being a standard option. Except for the cost parameter on log(km per liter), all cost parameters have the expected signs. Both the distance and income-related search cost parameters are highly significant. Households with higher income have a higher opportunity cost of time and therefore likely have higher search costs. The estimates indicate search costs are indeed positively related to household income. Having children in the household does not significantly affect search costs. The estimate for ρ indicates that the expected value of not searching is relatively large, which is consistent with the high proportion of consumers not searching in any given year.

Specification (B) adds distance from the consumer to the nearest dealer to the utility function. Although not significantly different from zero, distance has a negative effect on utility, which may capture the negative effect on utility of service visits to dealers located farther away from the consumer. Most of the other parameter estimates are very similar to those for specification (A), although the estimated effect of distance on search costs decreases slightly, while having children in the household now has a larger effect on search costs that is significant at the 10 percent level. The price coefficients are similar to the estimates reported in the first column of Table 5.

The last two columns of Table 5 give parameter estimates in case we assume consumers have full information, as in BLP. The results in column (C) of this table are obtained using the same preference and cost side parameters as in specification (A) of the table, although we can no longer use the search related micro moments. The estimated price coefficients indicate the marginal utility of price has gone down substantially in comparison to the estimates for the search model. In column (D) we add distance to the utility function. The distance from a consumer to the nearest dealer of a brand has a negative marginal utility, which could be explained by the distance variable picking up some of the search frictions that exist in this market. An alternative interpretation is that distance in the utility captures the negative effect on utility of service visits to dealers located farther away from the consumer. In this regard the estimate of the coefficient of distance has the expected sign. In addition, the absolute values of the estimated price coefficients go up in comparison to the estimates in column (C), which moves the estimated price coefficients further away from those for the search model.

As explained before, our search model allows us to add more search cost covariates than distance

Table 5: Estimation results with micro moments

	Search				Full Information			
	(A)		(B)		(C)		(D)	
Variable	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
Price coefficients (price/income)								
income less than 31k	-6.417	(0.453)***	-6.437	(0.506)***	-7.477	(0.472)***	-7.994	(0.560)***
income more than 31k	-2.280	(0.174)***	-2.269	(0.176)***	-3.573	(0.287)***	-3.381	(0.263)***
Base coefficients								
constant	-13.290	(0.196)***	-13.167	(0.195)***	-29.221	(0.242)***	-30.688	(0.212)***
HP/weight	1.572	(0.087)***	1.670	(0.088)***	2.329	(0.094)***	2.495	(0.093)***
non-European	-0.885	(0.030)***	-0.879	(0.030)***	-1.051	(0.032)***	-1.030	(0.032)***
cruise control	0.191	(0.039)***	0.170	(0.039)***	0.326	(0.043)***	0.249	(0.041)***
fuel efficiency	2.558	(0.082)***	2.544	(0.081)***	2.220	(0.095)***	2.190	(0.088)***
size	8.386	(0.218)***	8.361	(0.218)***	10.016	(0.258)***	9.984	(0.239)***
family car	-0.421	(0.039)***	-0.429	(0.039)***	-0.497	(0.044)***	-0.397	(0.042)***
MPV/SUV	0.623	(0.027)***	0.623	(0.027)***	0.620	(0.030)***	0.658	(0.030)***
Random coefficients								
constant	5.330	(0.827)***	5.306	(0.965)***	10.460	(0.896)***	11.623	(1.035)***
family car × kids	0.677	(0.145)***	0.712	(0.154)***	0.559	(0.024)***	0.571	(0.024)***
distance	—		-0.018	(0.013)	—		-0.034	(0.009)***
Cost parameters								
constant	2.743	(0.009)***	2.739	(0.009)***	2.981	(0.008)***	2.943	(0.008)***
log(HP/weight)	1.071	(0.018)***	1.072	(0.018)***	1.049	(0.014)***	1.037	(0.015)***
non-European	-0.210	(0.007)***	-0.211	(0.007)***	-0.216	(0.006)***	-0.214	(0.006)***
cruise control	0.072	(0.008)***	0.067	(0.008)***	0.054	(0.007)***	0.054	(0.007)***
log(km per liter)	-0.964	(0.028)***	-0.978	(0.028)***	-1.022	(0.025)***	-1.007	(0.025)***
log(size)	0.359	(0.040)***	0.333	(0.040)***	0.625	(0.040)***	0.526	(0.039)***
Search cost parameters								
distance	0.028	(0.005)***	0.025	(0.007)***	—		—	
log(income)	0.523	(0.025)***	0.524	(0.026)***	—		—	
kids	0.263	(0.165)	0.302	(0.172)*	—		—	
Expected value of not searching								
mean	10.205	(0.425)***	10.310	(0.460)***	—		—	

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. The number of observations is 1,382. The number of simulated consumers is 2,209. Standard errors are in parenthesis.

only. For instance, in both search specifications we include household income and a dummy for whether there are children present in the household, which improves the fit of the search model. We note that while the search model offers a natural justification for adding income or other demographic information to the model, it is difficult to justify adding these variables to the utility function in a full information model, which may explain why the existing literature has not followed this approach.

Table 6 gives the estimated probabilities used for the micro moments as well as those from the survey data. All specifications are able to match the probabilities from the survey data relatively well—all estimated probabilities used as micro moments are well within 10 percent of the

Table 6: Fit micro moments

	Survey	Search (A)	Search (B)	Full information (C)	Full information (D)
Non-search related probabilities					
$E[\mathbb{1}\{i$ purchases an MPV or SUV$\} \mid \{i$ has a family size $\leq 3\}]$	0.013	0.013	0.013	0.013	0.013
$E[\mathbb{1}\{i$ purchases a family car$\} \mid \{i$ has children in the household$\}]$	0.048	0.048	0.048	0.048	0.048
$E[\mathbb{1}\{i$ purchases new vehicle$\} \mid \{y_i < \overline{y}\}]$	0.053	0.049	0.049	0.049	0.049
$E[\mathbb{1}\{i$ purchases new vehicle$\} \mid \{y_i \geq \overline{y}\}]$	0.105	0.098	0.098	0.098	0.098
Search related probabilities					
$E[\mathbb{1}\{i$ searches at least twice$\} \mid \{y_i < \overline{y}\}]$	0.061	0.060	0.060		
$E[\mathbb{1}\{i$ searches at least twice$\} \mid \{y_i \geq \overline{y}\}]$	0.096	0.096	0.096		
$E[\mathbb{1}\{i$ searches at least twice$\} \mid \{\#$ of dealers within 10km of $i < 15\}]$	0.044	0.044	0.044		
$E[\mathbb{1}\{i$ searches at least twice$\} \mid \{\#$ of dealers within 10km of $i \geq 15\}]$	0.078	0.078	0.078		

corresponding probabilities from the survey data.

5.3 Demand elasticities and markups

Table 7 gives demand elasticity estimates for a selection of car models sold in 2008 for both the search model (using the estimates in column (B) of Table 5) and the full information model (using the estimates in column (D) of the table). For all models, demand is estimated to be more inelastic in the search model than in the full information model. This means that assuming consumers have full information while in reality they do not, will lead to an overestimation of price sensitivity for most car models. The cross-price elasticities show a similar pattern: the percentage change in market share as a result of a percent increase in price of a rival model is in the majority of the cases smaller in the search model than in the full information model.

Table 8 compares the estimated markups between the search model and the full information model. Consistent with the elasticity patterns reported in Table 7, estimated markups in the search model are higher for all the cars. For both specifications markups are increasing in the price of the car, although as a percentage of the (pre-tax) price it is more increasing in price in the search model than in the full information specification. The estimated average (sales-weighted) percentage markup across all models in 2008 is 43 percent for the search model versus 35 percent for the full information model. These numbers are not unrealistic: BLP report an average ratio of markup to retail price of 24 percent for their main specification, whereas Petrin (2002) finds an average markup of 17 percent for the model with micro moments. Goldberg (1995), on the other hand, obtains higher markup estimates: wholesale price markups of on average 38 percent, which implies

Table 7: Demand elasticity estimates

	Mitsubishi Colt	Peugeot 207	Volkswagen Golf	Opel Astra	Nissan Qashqai	Mercedes B Class	Renault Espace	Audi A6
Search								
Mitsubishi Colt	-2.5809	0.0921	0.0453	0.0279	0.0130	0.0056	0.0026	0.0055
Peugeot 207	0.0183	-2.6205	0.0403	0.0257	0.0133	0.0066	0.0040	0.0094
Volkswagen Golf	0.0079	0.0353	-2.2828	0.0386	0.0232	0.0046	0.0035	0.0088
Opel Astra	0.0068	0.0313	0.0538	-2.2454	0.0229	0.0047	0.0038	0.0094
Nissan Qashqai	0.0044	0.0224	0.0447	0.0316	-2.0919	0.0048	0.0043	0.0110
Mercedes B Class	0.0044	0.0260	0.0208	0.0152	0.0112	-2.0748	0.0077	0.0204
Espace	0.0021	0.0160	0.0164	0.0125	0.0103	0.0079	-2.1222	0.0221
Audi A6	0.0017	0.0143	0.0152	0.0117	0.0099	0.0078	0.0083	-2.3416
Full information								
Mitsubishi Colt	-3.0797	0.1396	0.0812	0.0499	0.0227	0.0076	0.0036	0.0080
Peugeot 207	0.0278	-3.1725	0.0734	0.0468	0.0241	0.0097	0.0062	0.0148
Volkswagen Golf	0.0142	0.0642	-2.9113	0.0614	0.0364	0.0085	0.0067	0.0165
Opel Astra	0.0121	0.0570	0.0854	-2.8816	0.0359	0.0088	0.0073	0.0180
Nissan Qashqai	0.0076	0.0406	0.0701	0.0497	-2.7269	0.0093	0.0085	0.0215
Mercedes B Class	0.0060	0.0384	0.0383	0.0285	0.0219	-2.6864	0.0133	0.0346
Espace	0.0029	0.0250	0.0309	0.0241	0.0203	0.0135	-2.9271	0.0380
Audi A6	0.0024	0.0225	0.0286	0.0223	0.0193	0.0133	0.0143	-3.2616

Notes: Demand elasticities are calculated for 2008. 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 (B) in Table 5; those for the full information model are based on specification (D) in Table 5.

even larger retail price markups. In both the search model and the full information model, the Volkswagen Golf is the most profitable model among the ones listed in Table 8, and one of the most profitable models in general.

5.4 Counterfactuals

In this section we study the effects of two changes in the primitives of the model, using the search model estimates reported in column (B) of Table 5. First, we look at what happens to equilibrium prices when search costs change. Secondly, we explore what happens to prices and profits when one of the multi-product firms starts retailing all of its brands in each and everyone of its dealerships.

Change in search costs

To see how prices change if search costs decrease, we take the estimates reported in column (B) of Table 5 and simulate equilibrium prices and market shares setting the non-random part of each consumer's search cost equal to a specific percentage of her estimated search cost. Table 9 shows the effects on prices for a few selected models, for various levels of search costs. In most of these cases, prices go down for small decreases in search costs. For instance, the simulated price of a

Table 8: Markups

	price	pre-tax price	Search: markup over MC	Search: percentage markup	Search: variable profit	Full information: markup over MC	Full information: percentage markup	Full information: variable profit
Mitsubishi Colt	10,213	6,819	2,468	36.19	9.57	2,071	30.36	8.03
Peugeot 207	13,540	9,289	3,463	37.28	50.87	2,933	31.57	43.08
Volkswagen Golf	16,632	11,206	5,007	44.68	68.38	4,054	36.17	55.36
Opel Astra	17,799	11,697	5,120	43.77	46.96	4,060	34.71	37.24
Nissan Qashqai	21,083	13,559	6,508	48.00	36.43	5,071	37.40	28.39
Mercedes B Class	25,290	16,574	7,734	46.66	15.37	6,016	36.30	11.95
Renault Espace	34,516	22,061	10,407	47.17	14.86	7,645	34.65	10.92
Audi A6	41,033	25,927	11,725	45.22	37.50	8,691	33.52	27.80

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

Peugeot 207 drops from €13,540 to €13,441 when search costs are reduced to 90 percent of the estimated search costs. However, after this initial price drop, prices tend to increase for larger reductions in search costs: the simulated price of a Peugeot 207 when search costs are 50 percent of the estimated search costs is €13,843, which is higher than the price observed in the data. In fact, for more than half the models shown in Table 9, when search costs are zero (i.e., zero percent) prices are higher than prices in the original data. A similar pattern emerges when looking at the average price across all models: prices initially decrease when search costs are lowered, but for larger decreases in search costs, prices tend to go up, until they more or less stabilize for search costs lower than 30 percent of the estimated search costs.

In Table 9 we also report prices when using the estimates to simulate prices under the full information model. As we have shown in Section 3, the full information model is only equivalent to the search model when $\gamma \to -\infty$, which means that the simulated prices under the full information model are not necessarily similar to those when search costs are zero percent of the estimated search costs. However, as shown in the last column of the table, simulated prices are very similar to those in the next-to-last column of the table, and are consistent with our finding that prices tend to decrease with search costs for about half of the models. Also notice that if the deterministic part of search cost is zero then the only difference between the search model and the full information is the search cost shock λ_{iS}, so the small difference between simulated prices when search costs are zero and simulated prices under the full information model suggests the search cost shock does not have a major impact on equilibrium prices.

The non-monotonic relationship we find between prices and search costs deserves an explanation.

Table 9: Simulated prices for different search cost levels

Model	prices when search costs are x percent of the estimated search costs									full
	130%	110%	100%	90%	70%	50%	30%	10%	0%	info
Mitsubishi Colt	10,474	10,273	10,213	10,178	10,233	10,433	10,478	10,469	10,464	10,406
Peugeot 207	13,843	13,633	13,540	13,441	13,481	13,843	13,937	13,911	13,898	13,795
Volkswagen Golf	22,191	17,771	16,632	15,943	15,341	15,529	15,674	15,631	15,612	15,481
Opel Astra	22,458	19,054	17,799	16,959	16,128	16,159	16,265	16,223	16,205	16,073
Nissan Qashqai	24,961	22,465	21,083	19,905	18,563	18,532	18,715	18,660	18,633	18,426
Mercedes B Class	26,039	25,588	25,290	24,922	24,589	25,350	25,593	25,513	25,474	25,131
Renault Espace	34,344	34,455	34,516	34,557	34,632	34,840	34,950	34,926	34,915	34,837
Audi A6	40,865	40,961	41,033	41,077	41,184	41,463	41,613	41,587	41,575	41,521
Average (sales weighted)	20,509	19,140	18,613	18,179	17,875	18,504	18,783	18,751	18,733	18,450
Average	30,107	29,476	29,214	28,969	28,696	28,926	29,076	29,033	29,012	28,860
Share not searching	0.9740	0.9484	0.9242	0.8828	0.6703	0.1870	0.0026	0.0000	0.0000	0.0000

Notes: Prices are in Euros. The prices shown are simulated prices when search costs are x percent of the non-random part of the estimates search costs from specification (B) in Table 5, where x is between 0 and 130 percent. The last column gives simulates prices for the full information model.

As noted before, in our model an increase in search costs has three different effects. First, as can be seen in Table 9, as search costs increase the share of consumers not searching goes up. If the consumers who remain in the market are the more elastic ones, as demonstrated in Moraga-González, Sándor, and Wildenbeest (2014), firms get an incentive to lower their prices. Second, as it is standard in search models, higher search costs give firms enhanced market power over the consumers who visit, and thereby firms have an incentive to raise their prices. Finally, because we have a model in which prices are observed before search, higher search costs increase competition for visits, and this put downward pressure on prices (see also Footnote 13). The latter effect is because consumers in our model determine the optimal subset of dealers to visit by making a tradeoff between the expected utility and the search costs of visiting these dealers. Since the former depends on the observed prices of the cars, a lower price for a specific car model makes it more likely a consumer includes the dealer selling that car in her choice set, which leads to a negative relation between search costs and prices. Which of the three effects will dominate depends on the level of search costs as well as the level of competition. Both of these are determined at the model level, which explains why we find that simulated prices are increasing in search costs for some car models, while decreasing for other cars.

Selling different makes at same dealership

Most of the 16 firms in our data own several brands. For instance, the Volkswagen Group owns Audi, Seat, Skoda, and Volkswagen, and the Toyota Group owns Daihatsu, Lexus, and Toyota.

Whereas firms typically sell their brands in separate dealerships, in a setting where search frictions are important it might be beneficial to sell multiple brands at the same location.[23] Although in the Netherlands most manufacturers sell their brand at separate, single-brand dealerships, some manufacturers let some (or all) of their brands be sold under one roof. For instance, Audi and Volkswagen are typically sold at the same dealership, just as BMW and Mini (both part of the BMW Group).

To study the effects of retailing different brands within a single dealership on prices, market shares, and profits, we let all Toyota dealerships sell Lexus as well, while at the same time we eliminate the locations of the Lexus dealerships. In terms of the model this means that by visiting this combined Toyota-Lexus dealer, consumers observe relevant characteristics for all Toyota and Lexus models. Since Toyota has the most dense dealership network of the two brands, the number of dealers selling Lexus' cars goes up substantially, from 13 to 141, which means that average search costs for Lexus go down considerably.

Table 10 reports the simulated prices, market shares, and profits of the major brands in our dataset after the change in the use of the dealership networks. Average prices go down for all manufacturers, although the changes are small. The reduction in search costs is beneficial to all manufacturers since consumers are less likely to select the outside option, and, as a result of Toyota Group's change in Lexus brand retailing strategy, sales and profits go up for all manufacturers. The Toyota Group profits the most from the changes: sales go up by 1,131 cars (2.3 percent), while variable profits go up by €9.9 mln (4.6 percent). Of the three brands that are part of the Toyota Group, Lexus benefits the most: sales go up by 718 cars, which is an increase of 68.8 percent. Notice that in this calculation we have ignored changes in the fixed costs associated to this business reorganization.

6 Conclusions

In many markets consumers have imperfect information about the utility they get from the various alternatives available and have to engage in costly search to find out which products they prefer most. While the theoretical consumer search literature is well established, much less work exists trying to estimate demand and supply for environments in which consumer search is important. This paper has contributed to the literature by presenting and estimating a discrete choice model

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

Table 10: Sales, market shares, and prices for new dealer network

Group	sales (units) before	sales (units) after	market shares (%) before	market shares (%) after	prices (€) before	prices (€) after	profits (mln €) before	profits (mln €) after
BMW	18,587	18,728	3.80	3.78	50,527	50,513	162.15	164.02
Daimler-Chrysler	14,771	14,892	3.02	3.01	42,581	42,562	129.93	131.38
Fiat	24,821	25,071	5.07	5.06	18,594	18,576	84.59	85.12
Ford	68,746	69,504	14.04	14.03	38,692	38,665	476.40	481.49
Fuji	1,422	1,433	0.29	0.29	23,560	23,549	8.23	8.27
General Motors	49,962	50,506	10.20	10.20	24,096	24,076	240.62	242.75
Honda	8,479	8,555	1.73	1.73	23,017	23,002	46.14	46.29
Hyundai	17,433	17,611	3.56	3.56	16,521	16,503	64.66	65.14
Kia	12,236	12,358	2.50	2.49	19,100	19,080	49.45	49.74
Mitsubishi	7,805	7,888	1.59	1.59	27,808	27,792	33.35	33.52
Porsche	531	533	0.11	0.11	76,783	76,782	6.73	6.81
PSA Peugeot Citroen	64,389	65,109	13.15	13.14	21,520	21,497	296.75	299.20
Renault-Nissan	52,334	52,882	10.69	10.68	21,088	21,065	238.00	239.67
Suzuki	14,547	14,702	2.97	2.97	12,477	12,469	42.12	42.36
Toyota	49,227	50,358	10.05	10.17	30,018	29,987	217.61	227.51
Daihatsu	9,186	9,271	1.88	1.87	11,618	11,617	23.74	23.79
Lexus	1,044	1,762	0.21	0.36	64,262	64,118	11.98	20.27
Toyota	38,997	39,325	7.97	7.94	25,021	25,023	181.89	183.45
Volkswagen	84,294	85,250	17.22	17.21	27,945	27,909	534.99	540.95
Total	489,584	495,380	100.00	100.00	29,214	29,192	2,631.70	2,664.23

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

of demand with optimal consumer search. While doing so, we have allowed for unobserved product heterogeneity as in BLP, which sets our paper apart from recent contributions on the theme.

Specifically, in our model consumers are initially unaware of whether a given product is a good match or not. Consumer decision making consists of a search stage and a purchase stage. In the search stage, consumers optimally determine which sellers to visit in order to maximize expected utility. In making this decision, consumers take into account their preferences for the various alternatives as well as the costs of searching them. In the purchase stage, after the matching parameters of all products in their choice sets are revealed, consumers either pick the good with the highest realized utility among the products searched, or else go for the outside option. We have provided a way to estimate the model and have applied it to the Dutch market for automobiles. We use distances from consumers to the nearest dealer of a specific brand as well as household characteristics reflecting the opportunity cost of time to specify consumer search costs. Even though the model can be estimated using only aggregate data such as market shares, product characteristics, and consumer and dealer locations, we have supplemented the data with a survey on actual dealer visits for a large number of respondents to strengthen the identification and improve the precision of the estimates.

The survey reveals that consumers conduct a rather limited amount of search before buying.

Moreover, a great deal of the searches involves test-driving a car. Our estimation results have shown that search costs are both significant and economically meaningful. Assuming, instead, that search frictions are negligible and consumers have full information results in higher own- and cross-price estimates, as well as lower estimated percentage markups.

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

Finally, we have investigated the effects of changes in the usage of firm dealership networks. Intuition suggests that the effect of retailing more car brands within a dealership on prices and profits is likely to be ambiguous. On the one hand, if a firm offers more cars at a dealership, this dealership becomes more attractive for consumers because of the implied economies of search associated with one-stop shopping. This tends to relax competition and lower demand. However, if a firm chooses to offer more cars then rival firms need to be more aggressive if they wish to enter consumer search sets. This tends to lower prices and expand demand. For the case of the Dutch market, we have found that if Toyota started to sell Lexus cars in all its dealerships, then prices for all cars would go down but demand would expand. This not only benefits Toyota brands, but also leads to higher sales and profits at all other brands.

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., André de Palma, and Jacques-François Thisse: "Discrete Choice Theory of Product Differentiation," The MIT Press, Cambridge, MA, 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," Mimeo, 2014.

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

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

Berry, Steven, 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.

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

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

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 and Learning." Mimeo, 2015.

Dinerstein, Michael, Liran Einav, Jonathan Levin, and Neel Sundaresan: "Consumer Price Search and Platform Design in Internet Commerce," Mimeo, 2014.

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

Haan, Marco, and José Luis Moraga-González: "Consumer Search with Observable and Hidden Characteristics," Mimeo, 2011.

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

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.

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

Konovalov, Alexander and Zsolt Sándor: "On Price Equilibrium with Multi-Product Firms," *Economic Theory* 44, 271–292, 2010.

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

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

Moraga-González, José Luis, Zsolt Sándor, and Matthijs R. Wildenbeest: "Prices, Product Differentiation and Heterogeneous Search Costs," Tinbergen Institute Discussion Paper TI 2014-080/VII.

Morgan, Peter and Richard Manning: "Optimal Search," *Econometrica* 53, 923–944, 1985.

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

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, 2015.

Nurski, Laura and Frank Verboven: "Exclusive Dealing as a Barrier to Entry? Evidence from Automobiles," Mimeo, 2013.

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," Mimeo, 2014.

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

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

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

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

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

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

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

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

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

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

A Derivation of Buying Probabilities

Let $\bar{c}_{iS} = t'_{iS}\gamma$ is the (mean) cost of searching the dealers contained in S. Since

$$\exp[m_{iS}] = \exp\left[-t'_{iS}\gamma\right]\left(1 + \sum_{f\in S}\exp[\delta_{if}]\right),$$

we get

$$\begin{aligned} s_{ij} &= \frac{\exp[\delta_{ij}]\sum_{S\in\mathcal{S}_f}\exp\left[-\bar{c}_{iS}\right]}{\sum_{S'\in\mathcal{S}}\exp[m_{iS'}]} = \frac{\exp[\delta_{ij}]\sum_{S\in\mathcal{S}_{-f}}\exp\left[-t'_{i\{f\}}\gamma - t'_{iS}\gamma\right]}{\sum_{S'\in\mathcal{S}}\exp[m_{iS'}]} \\ &= \frac{\exp[\delta_{ij}]\exp\left[-t'_{i\{f\}}\gamma\right]\sum_{S\in\mathcal{S}_{-f}}\exp\left[-t'_{iS}\gamma\right]}{\sum_{S'\in\mathcal{S}}\exp[m_{iS'}]} \\ &= \frac{\exp[\delta_{ij}]\exp\left[-t'_{i\{f\}}\gamma\right]\sum_{S\in\mathcal{S}_{-f}}\prod_{f\in S}\exp\left[-t'_{i\{f\}}\gamma\right]}{\sum_{S'\in\mathcal{S}}\exp[m_{iS'}]}. \end{aligned}$$

Since

$$\sum_{S\in\mathcal{S}_{-f}}\prod_{f\in S}\exp\left[-t'_{i\{f\}}\gamma\right] = \prod_{g\in F\backslash\{f\}}\left(1+\exp\left[-t'_{i\{g\}}\gamma\right]\right) = \frac{\prod_{g\in F}\left(1+\exp\left[-t'_{i\{g\}}\gamma\right]\right)}{\left(1+\exp\left[-t'_{i\{f\}}\gamma\right]\right)},$$

we have that

$$s_{ij} = \frac{\exp[\delta_{ij}]\exp\left[-t'_{i\{f\}}\gamma\right]\frac{\prod_{g\in F}\left(1+\exp\left[-t'_{i\{g\}}\gamma\right]\right)}{\left(1+\exp\left[-t'_{i\{f\}}\gamma\right]\right)}}{\sum_{S'\in\mathcal{S}}\exp[m_{iS'}]} = \frac{\frac{\exp[\delta_{ij}]}{\left(1+\exp\left[t'_{i\{f\}}\gamma\right]\right)}\Pi_{iF}}{\sum_{S'\in\mathcal{S}}\exp[m_{iS'}]},$$

where $\Pi_{iF} = \prod_{g\in F}\left(1+\exp\left[-t'_{i\{g\}}\gamma\right]\right)$.

Now note that s_{i0} is given by

$$\begin{aligned}
s_{i0} &= \frac{\exp(\rho) + \sum_{S\in\mathcal{S}\setminus\{\varnothing\}} \exp\left[-\bar{c}_{iS}\right]}{\sum_{S'\in\mathcal{S}} \exp[m_{iS'}]} = \frac{\exp(\rho) + \sum_{S\in\mathcal{S}\setminus\{\varnothing\}} \exp\left[-t'_{iS}\gamma\right]}{\sum_{S'\in\mathcal{S}} \exp[m_{iS'}]} \\
&= \frac{\frac{1}{\phi} + \sum_{S\in\mathcal{S}\setminus\{\varnothing\}} \prod_{f\in S} \exp\left[-t'_{i\{f\}}\gamma\right]}{\sum_{S'\in\mathcal{S}} \exp[m_{iS'}]} = \frac{\frac{1}{\phi} - 1 + \sum_{S\in\mathcal{S}} \prod_{f\in S} \exp\left[-t'_{i\{f\}}\gamma\right]}{\sum_{S'\in\mathcal{S}} \exp[m_{iS'}]} \\
&= \frac{\frac{1}{\phi} - 1 + \prod_{f\in F}\left(1 + \exp\left[-t'_{i\{f\}}\gamma\right]\right)}{\sum_{S'\in\mathcal{S}} \exp[m_{iS'}]} = \frac{\frac{1}{\phi} - 1 + \Pi_{iF}}{\sum_{S'\in\mathcal{S}} \exp[m_{iS'}]} \\
&= \frac{\left(\frac{1-\phi}{\phi\Pi_{iF}} + 1\right)\Pi_{iF}}{\sum_{S'\in\mathcal{S}} \exp[m_{iS'}]},
\end{aligned}$$

where $\phi = \exp\left[-\rho\right]$. Since $\sum_{j=0}^{J} s_{ij} = 1$, it has to be that

$$\begin{aligned}
\sum_{S'\in\mathcal{S}} \exp[m_{iS'}] &= \left(\frac{1-\phi}{\phi\Pi_{iF}} + 1\right)\Pi_{iF} + \sum_{k=1}^{J} \frac{\exp[\delta_{ik}]}{1 + \exp\left[t'_{i\{g\}}\gamma\right]}\Pi_{iF} \\
&= \Pi_{iF}\left(\frac{1-\phi}{\phi\Pi_{iF}} + 1 + \sum_{k=1}^{J} \frac{\exp[\delta_{ik}]}{1 + \exp\left[t'_{i\{g\}}\gamma\right]}\right).
\end{aligned}$$

Consequently,

$$\begin{aligned}
s_{ij} &= \frac{\dfrac{\exp[\delta_{ij}]}{1 + \exp\left[t'_{i\{f\}}\gamma\right]}}{\dfrac{1-\phi}{\phi\Pi_{iF}} + 1 + \sum_{k=1}^{J} \dfrac{\exp[\delta_{ik}]}{1 + \exp\left[t'_{i\{g\}}\gamma\right]}}, \\
s_{i0} &= \frac{\dfrac{1-\phi}{\phi\Pi_{iF}} + 1}{\dfrac{1-\phi}{\phi\Pi_{iF}} + 1 + \sum_{k=1}^{J} \dfrac{\exp[\delta_{ik}]}{1 + \exp\left[t'_{i\{g\}}\gamma\right]}}.
\end{aligned} \tag{A13}$$

B Contraction Mapping

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

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

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

1. σ is continuously differentiable in ξ and

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

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

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

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

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

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

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

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

Then there are values $\underline{\xi}, \overline{\xi} \in \mathbb{R}$ such that the function $\overline{f} : [\underline{\xi}, \overline{\xi}]^J \to \mathbb{R}^J$ defined by $\overline{f}_j(\xi) = \min[\overline{\xi}, f_j(\xi)]$ has the property that $\overline{f}\left([\underline{\xi}, \overline{\xi}]^J\right) \subseteq [\underline{\xi}, \overline{\xi}]^J$, 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 $[\underline{\xi}, \overline{\xi}]^J$.

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 $s_j = \int s_{ij} dF_\tau(\tau_i)$, $j = 1, \ldots, J$. Note that equations

(5) and (A13) imply the derivatives

$$\begin{aligned}\frac{\partial s_{ij}}{\partial \xi_j} &= (1-s_{ij})\, s_{ij}, \quad \frac{\partial s_{ij}}{\partial \xi_k} = -s_{ik} s_{ij} \quad \text{for} \quad j \neq k, \quad \frac{\partial s_{i0}}{\partial \xi_j} = -s_{ij} s_{i0}, \\ \frac{\partial s_j}{\partial \xi_j} &= \int (1-s_{ij})\, s_{ij} dF_\tau(\tau_i), \quad \frac{\partial s_j}{\partial \xi_k} = -\int s_{ij} s_{ik} dF_\tau(\tau_i), \quad \frac{\partial s_0}{\partial \xi_j} = -\int s_{ij} s_{i0} dF_\tau(\tau_i). \quad \text{(A14)}\end{aligned}$$

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 (A14).

Condition 2. The fact that the share of the outside alternative $s_0 = 1 - \sum_{j=1}^J s_j$ is decreasing in all its arguments follows from equation (A14). 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 (A13) we see that

$$\widetilde{s}_0^j\left(x\right) = \frac{\frac{1-\phi}{\phi \Pi_{iF}} + 1}{\frac{1-\phi}{\phi \Pi_{iF}} + 1 + \frac{\exp[\delta_{ij}\left(x\right)]}{1 + \exp\left[t'_{i\{f\}} \gamma\right]}},$$

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

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

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

Further, from equation (5) we have

$$s_{ij}\exp\left(-\xi_j\right)=\frac{\dfrac{\exp[\alpha_i p_j+x_j'\beta_i]}{1+\exp\left[t_{i\{f\}}'\gamma\right]}}{\dfrac{1-\phi}{\phi\Pi_{iF}}+1+\sum_{k=1}^{J}\dfrac{\exp[\delta_{ik}]}{1+\exp\left[t_{i\{g\}}'\gamma\right]}},$$

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

$$\lim_{\xi\to-\infty}s_j\exp\left(-\xi_j\right)=\phi\Pi_{iF}\frac{\dfrac{\exp[\alpha_i p_j+x_j'\beta_i]}{1+\exp\left[t_{i\{f\}}'\gamma\right]}}{1-\phi+\phi\Pi_{iF}},$$

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

C Micro Moments

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

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

For specific examples we refer to Section 5.

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

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

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

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

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

$$\begin{aligned} E[a_{iT}|p, x, \theta, \xi, v_i] &= s_{iT}(\theta) \\ \text{var}[a_{iT}|p, x, \theta, \xi, v_i] &= s_{iT}(\theta)(1 - s_{iT}(\theta)). \end{aligned} \tag{A15}$$

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

$$g_q^R(\theta) = \frac{1}{N_R} \sum_{\substack{i=1 \\ q_i \in R}}^{N} (a_{iT} - s_{iT}(\theta)), \tag{A16}$$

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

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

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

D The Weighting Matrix

The optimal choice of Ξ is a matrix proportional to

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

where θ^0 is the true value of the parameter vector θ. Since the micro moments depend on the demand unobserved characteristics ξ, this weighting matrix is not a block diagonal in general. Several components of this matrix can be computed exactly; for example, the variances of the micro moments (see equation (A15)). However, several other components cannot be computed exactly. Therefore, in order to obtain a positive definite weighting matrix, we propose this matrix to be approximated as

$$\Xi = \begin{pmatrix} V_J\left(\hat{\theta}\right) & 0 \\ 0 & V_N\left(\hat{\theta}\right) \end{pmatrix}^{-1},$$

where $\hat{\theta}$ is a previously obtained consistent estimator of θ^0,

$$V_J\left(\hat{\theta}\right) = \frac{1}{(2J)^2} \sum_{j=1}^{2J} \left(Z_j'\psi_j\left(\hat{\theta}\right) - g_J\left(\hat{\theta}\right)\right)\left(Z_j'\psi_j\left(\hat{\theta}\right) - g_J\left(\hat{\theta}\right)\right)',$$

and $V_N\left(\hat{\theta}\right)$ is a diagonal matrix whose main diagonal contains the variances of the micro moments computed as in equation (A15).

Alternative choices for Ξ are $\Xi = I$ (the identity matrix) or

$$\Xi = \begin{pmatrix} (Z'Z)^{-1} & 0 \\ 0 & I_M \end{pmatrix}$$

where I_M is the identity matrix of dimension M, which is the number of micro moments.

E Importance Sampling and Estimation of the Scale Parameter

We start by succinctly extending the model for the case where the scale parameter is not fixed to 1. Recall that the utility consumer i derives from car j and the outside alternative are given by:

$$\begin{aligned} u_{ij} &= \alpha_i p_j + x_j'\beta_i + \xi_j + \sigma_\varepsilon \varepsilon_{ij}, \\ u_{i0} &= \sigma_\varepsilon \varepsilon_{i0}. \end{aligned}$$

Consumer i's search cost for visiting all the dealerships in S is

$$c_{iS} = t_{iS}'\gamma + \sigma_\lambda \lambda_{iS},$$

where $(-\lambda_{iS})$ are i.i.d. type I extreme value distributed across consumers and subsets of dealers. Those consumers who pick a choice set that only includes the outside good face the search cost $c_{i\varnothing} = \lambda_{i\varnothing}$.

Recall that the expected gains to consumer i from searching the dealerships in a subset S are

$$m_{iS} - \sigma_\lambda \lambda_{iS},$$

where

$$m_{iS} = E\left[\max_{j \in \mathcal{G}_f \cup \{0\},\ f \in S} \{u_{ij}\}\right] - t_{iS}'\gamma.$$

If F denotes the CDF of ε_{ij}, the random variable $\max_{j\in\mathcal{G}_f\cup\{0\},\ f\in S}\{u_{ij}/\sigma_\varepsilon\}$ has CDF $\prod_{j\in\mathcal{G}_f\cup\{0\},\ f\in S} F(u-\delta_{ij}/\sigma_\varepsilon)$, where δ_{ij} is the mean utility consumer i derives from alternative j, i.e., $\delta_{ij}=\alpha_i p_j+x_j'\beta_i+\xi_j$. Using this, we obtain

$$\begin{aligned} m_{iS} &= \sigma_\varepsilon \log\left(1+\sum_{f\in S}\exp[\Delta_{if}]\right)-t_{iS}'\gamma, \\ m_{i\varnothing} &= 0, \end{aligned}$$

where $\Delta_{if}=\log\left(\sum_{j\in\mathcal{G}_f}\exp[\delta_{ij}/\sigma_\varepsilon]\right)$.[24] Consumer i will pick the subset S_i that maximizes the expected gain $m_{iS}-\sigma_\lambda\lambda_{iS}$, i.e.,

$$\begin{aligned} S_i &= \arg\max_{S\in\mathcal{S}}[m_{iS}-\sigma_\lambda\lambda_{iS}] \\ &= \arg\max_{S\in\mathcal{S}}\left[\sigma_\varepsilon\log\left(1+\sum_{f\in S}\exp[\Delta_{if}]\right)-t_{iS}'\gamma-\sigma_\lambda\lambda_{iS}\right]. \end{aligned}$$

Since we assume $(-\lambda_{iS})$ is i.i.d. type I extreme value distributed, the probability that consumer i finds it optimal to sample the set of dealers S_i is P_{iS_i} where

$$P_{iS}=\frac{\exp[m_{iS}/\sigma_\lambda]}{\sum_{S'\in\mathcal{S}}\exp[m_{iS'}/\sigma_\lambda]}.$$

Given that consumer i searches the set S_i, the probability that consumer i buys j is equal to the probability that j is purchased out of the products of the firms in S_i is $P_{ij|S_i}$ where

$$P_{ij|S}=\frac{\exp[\delta_{ij}/\sigma_\varepsilon]}{1+\sum_{r\in S}\exp[\delta_{ir}/\sigma_\varepsilon]}.$$

In order to see which parameters can be identified, we can write P_{iS} out:

$$P_{iS}=\frac{\exp\left[\frac{\sigma_\varepsilon}{\sigma_\lambda}\log\left(1+\sum_{f\in S}\sum_{j\in\mathcal{G}_f}\exp\left[\frac{\delta_{ij}}{\sigma_\varepsilon}\right]\right)-t_{iS}'\frac{\gamma}{\sigma_\lambda}\right]}{\sum_{S'\in\mathcal{S}}\exp\left[\frac{\sigma_\varepsilon}{\sigma_\lambda}\log\left(1+\sum_{f\in S'}\sum_{j\in\mathcal{G}_f}\exp\left[\frac{\delta_{ij}}{\sigma_\varepsilon}\right]\right)-t_{iS'}'\frac{\gamma}{\sigma_\lambda}\right]}. \tag{A17}$$

Note that the vector γ only appears in the model in these expressions, that is, in the fractions γ/σ_λ. Therefore, only these fractions can be identified. The parameter σ_λ also appears in the fraction $\sigma_\varepsilon/\sigma_\lambda$, which can be identified. Elsewhere, σ_ε only appears in the fractions $\delta_{ij}/\sigma_\varepsilon$, which can be identified, but are not sufficient to identify the utility parameters together with σ_ε and σ_λ.

[24] Note that we have set the value of the outside option when not searching to $\rho=0$. This can easily be relaxed.

In conclusion, we can only identify the fraction $\sigma_\varepsilon/\sigma_\lambda$.

Based on these arguments, we can reparametrize the model. One possible reparametrization is to assume $\sigma_\varepsilon = \mu$, $\sigma_\lambda = 1$, and preserve the same notation for the utility parameters, even though they are divided by μ. Then

$$P_{iS} = \frac{\exp\left[\mu \log\left(1 + \sum_{f\in S}\sum_{j\in\mathcal{G}_f} \exp\left[\Delta_{ij}\right]\right) - t'_{iS}\gamma\right]}{\sum_{S'\in\mathcal{S}} \exp\left[\mu \log\left(1 + \sum_{f\in S'}\sum_{j\in\mathcal{G}_f} \exp\left[\Delta_{ij}\right]\right) - t'_{iS'}\gamma\right]} \tag{A18}$$

and

$$\begin{aligned} m_{iS} &= \mu \log\left(1 + \sum_{f\in S}\sum_{j\in\mathcal{G}_f} \exp\left[\Delta_{ij}\right]\right) - t'_{iS}\gamma, \\ m_{i\varnothing} &= 0. \end{aligned}$$

In this case

$$P_{ij|S} = \frac{\exp[\Delta_{ij}]}{1 + \sum_{r\in S}\exp[\Delta_{ir}]}.$$

Then the interpretation of the parameter μ is that it measures the magnitude in utility terms of the "search-like" product attributes. We note that a value $\mu > 1$ also reduces the effect of the unobserved cost shocks λ_{iS}. For practical reasons it is better to use equation (A18); denote the variables involved in $\Delta_{ij} = \frac{\delta_{ij}}{\sigma_\varepsilon}$ by $a_i = \frac{\alpha_i}{\sigma_\varepsilon}, \mathbf{b}_i = \frac{\beta_i}{\sigma_\varepsilon}, \eta_j = \frac{\xi_j}{\sigma_\varepsilon}$, that is, $\Delta_{ij} = a_i p_j + x'_j \mathbf{b}_i + \eta_j$.

Integrating out S_i from the probability that consumer i purchases product j, we obtain:

$$\begin{aligned} s_{ij} &= \sum_{S\in\mathcal{S}_f} P_{iS} P_{ij|S} \\ &= \sum_{S\in\mathcal{S}_f} \frac{\exp[m_{iS}]}{\sum_{S'\in\mathcal{S}}\exp[m_{iS'}]} \frac{\exp[\Delta_{ij}]}{1 + \sum_{r\in S}\exp[\Delta_{ir}]}. \end{aligned}$$

Let $\tau_i := (\alpha_i, \beta_i)$ be the vector of all random variables that capture consumer heterogeneity. Then the probability that product j is purchased is the integral

$$s_j = \int s_{ij} f_\tau(\tau_i) d\tau_i. \tag{A19}$$

We now provide a procedure to simulate these market shares by importance sampling methods.

Simulation of market shares

We use the empirical distribution of consumer demographics, including distances to dealers to proxy for search costs, which means the predicted market shares given by equation (A19) do not have an analytical solution and need to be simulated. Implementation of the simulations is not straightforward due to two problems. The first problem arises from the sum over 2^{F-1} choice sets involved in s_{ij}. The sum can be viewed as an expected value of a discrete random variable and estimated by Monte Carlo simulations by sampling from the discrete distribution. However, Monte Carlo simulations applied directly lead to estimators that are not continuous in the model parameters, which causes problems when we use them in optimization algorithms to find the estimates of the parameters. An additional complexity is that for a large number of firms, as in our empirical application, the number of choice sets from which one needs to sample is extremely large.

We can use importance sampling to tackle these two computational problems together. For this we first construct importance sampling probabilities. For an arbitrary choice set S, let

$$Q_{iS}(\theta) = \prod_{g \in S} \phi_{ig}(\theta) \prod_{h \notin S} (1 - \phi_{ih}(\theta)),$$

where ϕ_{ig} are probabilities that will be specified below in (A21). We define the importance sampling probabilities as the set $\{Q_{iS}\}_{S \in \mathcal{S}_{-f}}$, where S_{-f} denotes the set of all subsets of $F \backslash \{f\}$, $Q_{iS} = Q_{iS}(\theta_0)$, and θ_0 is the initial value of the parameters used in the numerical search for the parameter estimates. We note that these probabilities have a structure similar to those in Sovinsky Goeree (2008).

In order to estimate s_j first we rewrite $s_{ij} = \sum_{S \in \mathcal{S}_f} P_{iS} P_{ij|S}$ as

$$s_{ij} = \sum_{S \in \mathcal{S}_{-f}} P_{i\{f\} \cup S} P_{ij|\{f\} \cup S},$$

Clearly, for $S \in \mathcal{S}_{-f}$

$$P_{i\{f\} \cup S} = \frac{\exp\left[\mu \log\left(1 + \exp[\Delta_{if}] + \sum_{g \in S} \exp[\Delta_{ig}]\right) - \left(t'_{if} + \sum_{g \in S} t'_{ig}\right)\gamma\right]}{\sum_{S' \in \mathcal{S}} \exp[m_{iS'}]}$$

and

$$P_{ij|\{f\} \cup S} = \frac{\exp(\Delta_{ij})}{1 + \exp(\Delta_{if}) + \sum_{g \in S} \exp(\Delta_{ig})}.$$

Now, rewrite s_{ij} in the importance sampling form

$$s_{ij} = \sum_{S \in \mathcal{S}_{-f}} Q_{iS} \frac{P_{i\{f\} \cup S}}{Q_{iS}} P_{ij|\{f\} \cup S},$$

where $\sum_{S \in \mathcal{S}_{-f}} Q_{iS} = 1$ holds. A set S drawn randomly from $\mathcal{S}_{-f}$ can be represented as the vector of $[0,1]$ i.i.d. uniform random variables $u_{i,-f} = (u_{i1}, \ldots, u_{if-1}, u_{if+1}, \ldots, u_{iF})$ because according to the importance sampling probabilities we can draw S by drawing $u_{i,-f}$ such that $g \in S$ iff $u_{ig} \leq \phi_{ig}$ for all $g \in \mathcal{F} \backslash \{f\}$ (we omit the argument θ_0 from $\phi_{ig}(\theta_0)$). So we can use the argument $u_{i,-f}$ in the expressions involved in s_{ij}. We introduce

$$\begin{aligned}
Q_{if}(u_i) &= \prod_{g \in \mathcal{F} \backslash \{f\}} \phi_{ig}^{\mathbf{1}(u_{ig} \leq \phi_{ig})} (1 - \phi_{ig})^{\mathbf{1}(u_{ig} > \phi_{ig})} \\
P_{if}(u_i) &= \frac{\exp\left[\log\left(1 + \exp[\Delta_{if}] + \sum_{g \in \mathcal{F} \backslash \{f\}} \mathbf{1}\left(u_{ig} \leq \phi_{ig}\right) \exp[\Delta_{ig}]\right) - \left(t'_{if} + \sum_{g \in \mathcal{F} \backslash \{f\}} \mathbf{1}(u_{ig} \leq \phi_{ig}) t'_{ig}\right)\gamma\right]}{\sum_{S' \in \mathcal{S}} \exp[m_{iS'}]} \\
P_{ij|f}(u_i) &= \frac{\exp(\Delta_{ij})}{1 + \exp(\Delta_{if}) + \sum_{g \in \mathcal{F} \backslash \{f\}} \mathbf{1}\left(u_{ig} \leq \phi_{ig}\right) \exp(\Delta_{ig})},
\end{aligned}$$

where $\mathbf{1}(u_{ig} \leq \phi_{ig})$ is the indicator of the event $(u_{ig} \leq \phi_{ig})$. These correspond to Q_{iS}, $P_{i\{f\} \cup S}$ and $P_{ij|\{f\} \cup S}$, respectively. Then

$$s_{ij} = \int\limits_{[0,1]^{F-1}} \frac{P_{if}(u_i)}{Q_{if}(u_i)} P_{ij|f}(u_i) du_{i,-f} = \int\limits_{[0,1]^{F}} \frac{P_{if}(u_i)}{Q_{if}(u_i)} P_{ij|f}(u_i) du_i,$$

which yields

$$s_j = \int\limits_{\mathbb{R}^D} \int\limits_{[0,1]^F} \frac{P_{if}(u_i)}{Q_{if}(u_i)} P_{ij|f}(u_i) f_\tau(\tau_i) du_i d\tau_i, \tag{A20}$$

where D is the dimension of the random vector τ_i.

This latter formula is convenient because it shows how to estimate s_j by Monte Carlo. We can simply draw a random sample $(u_i, \tau_i)_{i=1}^{ns}$ jointly from their distribution and compute the Monte Carlo estimate of s_j as

$$\widetilde{s}_j = \frac{1}{ns} \sum_{i=1}^{ns} \left[\frac{P_{if}(u_i)}{Q_{if}(u_i)} P_{ij|f}(u_i)\right].$$

Note that although u_i is F-dimensional, we only use the $(F-1)$-dimensional $u_{i,-f}$ to compute $\widetilde{s}_j$, as equation (A20) also suggests, so there is a kind of redundancy of draws. The draw u_{if} is used for computing $\widetilde{s}_r$ for products r belonging to firms rival to f.

Algorithm 1 (Importance Sampling) *The algorithm consists of the following steps:*

1. *For each $i = 1, \ldots, ns$ draw $u_i \sim U[0,1]^F$ and $\tau_i \sim f_\tau$;*

2. *For each $f \in F$ compute ϕ_{if} and $Q_{if}(u_i)$; this implicitly determines the choice set $S_{i0} \subset F \backslash \{f\}$ of i as*
$$S_{i0} = \{g \in \mathcal{F} \backslash \{f\} : u_{ig} \leq \phi_{ig}\}$$
(note that the choice set for computing s_j for $j \in G_f$ will be $\{f\} \cup S_{i0}$, so always contains f);

3. *For each f compute $P_{if}(u_i)$ assuming for the moment that $\sum_{S' \in \mathcal{S}} \exp[m_{iS'}]$ is known;*

4. *For each j compute $P_{ij|f}(u_i)$;*

5. *For each j compute $\widetilde{s}_j$.*

In order to specify $\phi_{if}(\theta)$, the first idea that comes to mind is to use the criterion that the two sets of probabilities are proportional at the singleton subsets of firms $\{f\}$, $f = 1, \ldots, F$, i.e.,

$$\frac{Q_{i\{f\}}}{Q_{i\varnothing}} = \frac{P_{i\{f\}}}{P_{i\varnothing}},$$

which implies that

$$\frac{\phi_{if}}{1 - \phi_{if}} = \exp[m_{i\{f\}}],$$

so

$$\phi_{if} = \frac{\exp[m_{i\{f\}}]}{1 + \exp[m_{i\{f\}}]} = \frac{\exp[\log(1 + \exp[\delta_{if}]) - d_{if}\gamma_i]}{1 + \exp[\log(1 + \exp[\delta_{if}]) - d_{if}\gamma_i]}. \tag{A21}$$

Simulation experiments based on these importance sampling probabilities show that they are not satisfactory.

We can exploit more information on the structure of m_{iS} and incorporate it in the ϕ's by using the criterion that the two sets of probabilities are proportional at subsets of firms $\{f, g_1, \ldots, g_H\}$ and $\{g_1, \ldots, g_H\}$ for $f = 1, \ldots, F$. More precisely,

$$\frac{Q_{i\{f,g_1,\ldots,g_H\}}}{Q_{i\{g_1,\ldots,g_H\}}} = \frac{P_{i\{f,g_1,\ldots,g_H\}}}{P_{i\{g_1,\ldots,g_H\}}},$$

which implies

$$\frac{\phi_{if}}{1-\phi_{if}} = \exp[m_{i\{f,g_1,\ldots,g_H\}} - m_{i\{g_1,\ldots,g_H\}}].$$

If we assume that this holds for all $\{f, g_1, \ldots, g_H\} \subset \mathcal{F}$ we get that

$$\frac{\phi_{if}}{1-\phi_{if}} = \exp[\overline{m}_{if,H}],$$

where $\overline{m}_{if,H}$ denotes the mean of $m_{i\{f,g_1,\ldots,g_H\}} - m_{i\{g_1,\ldots,g_H\}}$ over all $\{g_1, \ldots, g_H\} \subset \mathcal{F}\backslash\{f\}$ (this implicitly assumes that $g_h \neq g_k$ for all $h \neq k,\ h,\ k = 1, \ldots, H$). This yields

$$\phi_{if} = \frac{\exp[\overline{m}_{if,H}]}{1+\exp[\overline{m}_{if,H}]},\ f = 1, \ldots, F.$$

The number of terms involved in $\overline{m}_{if,H}$ is the number of subsets of $\mathcal{F}\backslash\{f\}$ having H elements, that is, combinations $\binom{F-1}{H}$. Intuitively, the larger the subsets $\{g_1, \ldots, g_H\}$ involved, the more information on the structure of m_{iS} is captured by the ϕ's. The computational burden for computing the ϕ's for large F and H is high, but they only need to be computed for the starting value of the parameters.

We still need to find an estimator for $M_i = \sum_{S\in\mathcal{S}} \exp[m_{iS}]$, the denominator of $P_{if}(u_i)$. We can again use importance sampling based on the probabilities $\{Q_{iS}\}_{S\in\mathcal{S}}$ defined above. For this, write

$$M_i = \sum_{S\in\mathcal{S}} Q_{iS} \frac{\exp[m_{iS}]}{Q_{iS}} = \int\limits_{[0,1]^F} \frac{m_i(w)}{Q_i(w)} dw,$$

where

$$\begin{aligned} m_i(w) &= \exp\left[\log\left(1 + \sum_{f\in\mathcal{F}} \mathbf{1}(w_f \leq \phi_{if}) \exp[\Delta_{if}]\right) - \sum_{f\in\mathcal{F}} \mathbf{1}(w_f \leq \phi_{if}) t'_{if}\gamma\right], \\ Q_i(w) &= \prod_{f\in\mathcal{F}} \phi_{if}^{\mathbf{1}(w_f\leq\phi_{if})} (1-\phi_{if})^{\mathbf{1}(w_f>\phi_{if})} \quad \text{with} \quad w = (w_1, \ldots, w_F). \end{aligned}$$

The Monte Carlo estimator of M_i is

$$\widetilde{M}_i = \frac{1}{N} \sum_{n=1}^{N} \frac{m_i(w_n)}{Q_i(w_n)},$$

where $\{w_n\}_{n=1}^N$ is a set of i.i.d. draws from the uniform distribution on $[0, 1]^F$.

Some simulation experiments show that this Monte Carlo importance sampling estimator based on quasi-random samples works well for $H = 3$ and $Q_{iS} = Q_{iS}(\theta)$, that is, if we use the current parameter values. It can fail in some cases for $Q_{iS} = Q_{iS}(\theta_0)$ when the current parameter values θ are far from the starting values θ_0. Due to this it may be worthwhile to find good starting values by minimizing the objective function through a large scale grid search where now the importance sampling probabilities $Q_{iS}(\cdot)$ depend on the current value of θ.