# A Framework for the Estimation of Demand for Differentiated Products with Simultaneous Consumer Search*

José Luis Moraga-González†

Zsolt Sándor‡

Matthijs R. Wildenbeest§

March 2023

**Abstract**

We propose a tractable method for estimation of a simultaneous search model for differentiated products that allows for observed and unobserved heterogeneity in both preferences and search costs. We show that for type I extreme value distributed search costs, expressions for search and purchase probabilities can be obtained in closed form. We show that our search model belongs to the generalized extreme value (GEV) class, which implies that it has a full information discrete-choice equivalent, and hence search data are necessary to distinguish between the search model and the equivalent full information model. We allow for price endogeneity when estimating the model and show how to obtain parameter estimates using a combination of aggregate market share data and individual level data on search and purchases. To deal with the dimensionality problem that typically arises in search models due to a large number of consideration sets we propose a novel Monte Carlo estimator for the search and purchase probabilities. Monte Carlo experiments highlight the importance of allowing for sufficient consumer heterogeneity when doing policy counterfactuals and show that our Monte Carlo estimator is accurate and computationally fast. Finally, a behavioral assumption on how consumers search provides a micro-foundation for consideration probabilities widely used in the literature.

**Keywords:** demand estimation, price endogeneity, simultaneous search, differentiated products
**JEL Classification:** C14, D83, L13

---

*A simpler variant of the simultaneous search model proposed in this paper was first used in Moraga-González, Sándor, and Wildenbeest (2015). That paper underwent a major overhaul and the published version (Moraga-González, Sándor, and Wildenbeest, 2022) now features a sequential search model instead. For helpful comments on this paper we thank Javier Donna, Tobias Klein, Stefan Seiler, and Xuan Teng. The paper has also benefited from presentations at Tinbergen Institute, CREST (Paris), and EARIE 2022 (Vienna). Financial support from Marie Curie Excellence Grant MEXT-CT-2006-042471 and grant PN-II-ID-PCE-2012-4-0066 of the Romanian Ministry of National Education, CNCS-UEFISCDI is gratefully acknowledged.

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

‡Sapientia Hungarian University of Transylvania, E-mail: `sandorzsolt@uni.sapientia.ro`. Sándor is also affiliated with the Babeş-Bolyai University.

§University of Arizona and CEPR, E-mail: `wildenbeest@arizona.edu`.

# 1 Introduction

In a lot of markets (e.g. clothes, footwear, hotels, bicycles, motorbikes, cars, houses, insurance, etc.) consumers are often poorly informed about whether a product is a good match and have to engage in costly search in order to figure this out. Despite this obvious fact, it is surprising that the very large literature on demand estimation by and large assumes that consumers have full information. It is by now well-known that models that assume full information in settings in which consumers have limited information lead to biased estimates of demand and hence wrong policy conclusions. For example, Sovinsky Goeree (2008), Koulayev (2014), and Moraga-González, Sándor, and Wildenbeest (2022) find that demand is estimated to be more elastic than it really is when incorrectly assuming full information.[1]

We believe there are two important reasons behind the "stickiness" of the full information assumption. First, economic theory has shown that market outcomes not only depend on the magnitude of search costs (see for instance Wolinsky, 1986; Anderson and Renault, 1999) but also on the dispersion of search costs (see for instance Hortaçsu and Syverson, 2004; Moraga-González, Sándor, and Wildenbeest, 2017). Therefore, an empirically relevant consumer search model must allow for rich consumer search cost heterogeneity. However, the typical search model assumes sequential search, which quickly becomes intractable once allowing for consumer heterogeneity in both preferences and search costs. This intractability is primarily due to the many different search paths a consumer could take before deciding to purchase a product, which all need to be integrated out to derive an expression for a consumer's purchase probability (see Moraga-González, Sándor, and Wildenbeest, 2022). This often has to be done numerically, and the additional computational burden associated with this can be problematic when using a BLP-type contraction mapping (Berry, Levinsohn, and Pakes, 1995) to deal with price endogeneity, or when calculating the demand derivatives that are necessary to obtain counterfactual prices.

Second, to separately identify search costs from preferences, data on purchases is typically not enough, and needs to be supplemented with data on search behavior. For instance, without search data it is difficult to distinguish between consumers who do not buy because of a preference for the outside option, or because of refraining from searching due to high search costs. This means that in the absence of search data, the only way the model can explain non-purchases is by a low utility constant, which could lead to biases in price elasticity estimates if search frictions are in fact

[1] The bias may be in both directions depending on the correlation between prices, product characteristics, and search costs. For example, if cheaper items happen to have higher search costs, ignoring the costs of search in estimation may lead to estimates of demand that are too inelastic.

important. However, even with search data it is difficult to fully exploit this for identification due to similar tractability issues as mentioned above, since in sequential search models the probability of having searched a particular set of products is often a high-dimensional integral that cannot be reduced to a closed-form expression.

Due to these difficulties, it is not a surprise that existing search models have only been estimated in relatively simple settings while often making restrictive assumptions regarding consumer and firm behavior. For example, earlier papers such as Hong and Shum (2006) and Moraga-González and Wildenbeest (2008) did not allow for product differentiation whatsoever while Hortaçsu and Syverson (2004) did not allow for horizontal product differentiation. Although later papers did allow for richer specifications in terms of consumer preferences, search behavior was typically still modeled in a restricted way. For example, Koulayev (2014) assumes a pre-specified order of search and only allows for a maximum of six searches. Murry and Zhou (2020) restrict search to be across clusters of firms instead of individual firms, and, in the absence of search data, cannot estimate a search cost constant, which means they are unable to estimate the level of search costs. Donna, Pereira, Pires, and Trindade (2022) are able to estimate the level and variance of search costs in a setting where consumers search for a maximum of nine retailers but restrict the amount of preference and search cost heterogeneity. Moraga-González, Sándor, and Wildenbeest (2022) do estimate a relatively rich specifications of utility and search costs, but normalize the variance of the unobserved part of search costs. Finally, most search papers in the marketing literature do not allow for price endogeneity (e.g., Honka, 2014; Ursu, 2018; Yavorsky, Honka and Chen, 2021).

The purpose of this paper is to propose a rich model of consumer search that does not suffer from the above mentioned difficulties. Specifically, in Section 2 we develop a model of simultaneous search (as in Stigler, 1961; Burdett and Judd, 1983) for differentiated products in which consumers search to discover whether a product is a good fit or not. We model the search and purchase decision as a two-step procedure. In a first step, consumers decide which products to inspect, making a tradeoff between the gains from searching a subset of products and the cost of searching those products. In a second step, after having incurred search costs to find out all the relevant details, consumers choose which of the inspected products to purchase, if any.[2]

We show that if search costs follow a type I extreme value (TIEV) distribution, we can obtain a closed-form expression for an individual's probability of purchasing a specific product. This probability is the sum of the products of two logit expressions: the multinomial logit probability that

[2]For theoretical studies of models of simultaneous search for differentiated products see Section 7.6 of Anderson, de Palma, and Thisse (1992), Chade and Smith (2006), and Moraga-González, Sándor, and Wildenbeest (2021).

a group of products is searched times the multinomial logit probability that a product is chosen from the set of searched products. We show that our model belongs to the family of generalized extreme value (GEV) models, which implies that it is consistent with utility maximization (McFadden, 1978).[3] More importantly, this implies that the search model has a full information discrete choice equivalent in which consumers pick the best product out of the set of all products assuming the utility shocks are no longer TIEV distributed but distributed according to the corresponding GEV joint distribution function. This is the analog of the *eventual purchase theorem* of Armstrong (2017) and Choi, Dai, and Kim (2018) for our setting, and implies that search data are necessary to distinguish between the search model and the equivalent full information model. This is because the GEV distribution function corresponding to our model clearly shows that search costs induce a specific correlation structure among the utility shocks of the products and that such correlation structure is only affected by the mean and the variance of search costs, and not by consumer preferences.

Because in general settings the number of products may be large, the individual search and purchase probabilities may involve very large sums. We show in Section 3 that such large sums can be conveniently estimated by Monte Carlo under the natural assumption that the cost of searching a subset of firms is equal to the sum of the search costs of the firms. The relatively straightforward computation of the search and choice probabilities allows us to derive the individual likelihood of a search and purchase decision and to estimate our model following the methods proposed by Goolsbee and Petrin (2004) and Train and Winston (2007). We first construct the likelihood of our data on individuals' search and purchase behavior. Then, following Berry, Levinsohn, and Pakes (2004), we use the aggregate data and the contraction property of our system of market share equations to solve for the mean utilities of the products. These mean utilities substitute for the linear part of utility in the likelihood function, which implies that we maximize the likelihood function for the "non-linear" parameters only. This procedure is the maximum likelihood analog of the micro-moments method of Berry, Levinsohn, and Pakes (2004), which yields consistent estimates of the non-linear parameters of the model despite price endogeneity. These parameters include the variance parameters of consumers' preferences and the mean and variance parameters of search costs. Moreover, the maximum likelihood approach yields efficient estimates of such parameters (see Grieco, Murry, Pinkse and Sagl, 2022). In a final step we regress the mean utilities on prices and other exogenous characteristics using an instrument for prices to get estimates of the marginal

[3] For the case of no search cost shifters our model is equivalent to the two-stage choice set formation model in Swait (2001) (named the GenL model) in which preference parameters endogenously determine choice sets.

effects of each of these characteristics, which makes controlling for price endogeneity relatively straightforward.

We also show that our derivation of the choice-set probabilities can be used to provide a microfoundation for the consideration probabilities proposed by Sovinsky Goeree (2008), which have also been used by Abaluck and Adams-Prassl (2021) and Brand and Demirer (2022). Specifically, we show that for the special case in which we adopt the behavioral assumption that consumers base their decision about which firms to visit only on the costs of searching the various alternatives and ignore the utility consequences of their search altogether, we obtain the consideration probabilities of Sovinsky Goeree (2008). What is distinctive about this consideration set model (referred to as the alternative specific consideration (ASC) model by Abaluck and Adams-Prassl, 2021) is that the probability of a specific consideration set is the product of the consideration probabilities of each individual product, which creates cross-derivative asymmetries that, when making specific assumptions on unobserved consumer heterogeneity, can be employed to separately identify preferences from consideration without using data on consideration (see Abaluck and Adams-Prassl, 2021). The consideration set probabilities in our search model are allowed to be consumer specific and do not have the product structure of the ASC model and its associated cross-derivative asymmetries, which means search data are needed for identification in our model.

In Section 4 we provide the results of several Monte Carlo exercises that show the performance of our estimator and also illustrate the importance of estimating the mean and the variance of search costs. As explained above, search data are needed to estimate both a utility and a search cost constant—the results of our first Monte Carlo simulation indicate that failing to estimate a search cost constant leads to biased elasticity and markup estimates. The reason for this is that the search cost constant governs the extensive margin of search, so when it is high many consumers do not even start searching for a product in the first place. Failing to accurately account for the extensive margin of search inevitably causes the econometrician to wrongly conclude that the utility consumers derive from the inside products is lower than what it truly is. This leads to biases in the substitution patterns and makes demand more elastic than it really is, with too low market power estimates as a result. We show that these biases might lead to wrong policy conclusions when doing counterfactuals, for instance when studying how prices change following a negative demand shock.

In a second experiment, we study the impact of normalizing the variance of search costs. We show that normalizing the scale parameter of the search cost shock to one while the true value is three times as large leads to biases in both preference and search cost parameters. This is due to the normalization lowering the gains from search, which results in search costs and demand elasticities being underestimated. As in our first experiment, this bias carries over to the predicted price effects following a drop in aggregate demand.

Finally, we assess the performance of the Monte Carlo estimator by comparing the parameter estimates obtained using the simulated market shares and individual search and purchase probabilities to those obtained using the actual expressions when the number of firms is not too large. We find that even though our Monte Carlo estimator is slower when the number of firms is five or lower, it is about twenty times faster when there are ten firms. Moreover, in all our experiments the parameter estimates based on the Monte Carlo estimator are almost identical to those based on the actual probabilities, so the computational gains do not come at the cost of accuracy.

## Related literature

Our paper fits into the large theoretical and empirical literature on consumer search behavior. Our paper relates to a strand in the theoretical consumer search literature that focuses on search for differentiated products (Wolinsky, 1986; Anderson and Renault, 1999). Our theoretical 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 and for consumer heterogeneity in both preferences and search costs.[4] The search model in our paper is also related to the simultaneous search model in Moraga-González, Sándor, and Wildenbeest (2021), who, as in our model, allow for search cost heterogeneity and differentiated products, and provide conditions under which a price equilibrium in pure strategies exists.

The empirical literature has seen a number of recent contributions that estimate models of search for differentiated products (Kim, Albuquerque, and Bronnenberg, 2010; De los Santos, Hortaçsu, and Wildenbeest, 2012; Seiler, 2013; Dinerstein, Einav, Levin, and Sundaresan, 2018; Honka, 2014; Koulayev, 2014; Pires, 2016). Our paper is most closely related to Moraga-González, Sándor, and Wildenbeest (2022). An important difference with that paper is in the way search behavior is modeled: while in this paper we assume consumers search simultaneously, Moraga-González, Sándor, and Wildenbeest (2022) assume consumers search sequentially. With sequential search, search decisions are based on realized search outcomes, which makes it less straightforward to obtain closed-form expressions for the buying probabilities. Using recent insights from Armstrong (2017) and Choi, Dai, and Kim (2018), Moraga-González, Sándor, and Wildenbeest (2022) show how to obtain closed-form expressions for the buying probabilities. In this paper we provide an alternative approach which relies on simultaneous search combined with the search cost shocks being

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

choice-set specific. Both approaches lead to closed-form expressions for the buying probabilities when making specific assumptions about the search cost distribution. An advantage of the model we lay out in this paper is that it is easier to obtain expressions for the search probabilities, which makes it especially suitable for use with individual-specific search and choice data, as we illustrate in Section 4 using Monte Carlo experiments. Another advantage of simultaneous search is that it can easily handle settings in which firms (or retailers) sell overlapping sets of products, since search decisions are determined before any realizations of the match values are observed. In contrast, with sequential search consumers determine after each search whether to continue or not, so search decisions are conditional on observed match values. This means that if firms sell overlapping sets of products, searching one firm may give valuable information about searching another firm, which makes the optimal search rule more challenging than in the standard setting in which firms sell non-overlapping sets of products (see Weitzman, 1979).

A number of recent papers have built on earlier versions of the search model presented here. For instance, Lin and Wildenbeest (2020) develop a method to non-parametrically estimate search costs using a conditional logit version of our model, in which search costs are assumed to be consumer-specific but identical across firms. Murry and Zhou (2020) use individual-level transaction data for new products to quantify how geographical concentration among product sellers affects competition and search behavior. Donna, Pereira, Pires, and Trindade (2022) estimate the welfare effects of intermediation in the Portuguese outdoor advertising industry using a demand model that extends our search model to allow for nested logit preferences. Ershov (2018) develops a structural model of supply and demand to estimate the effects of search frictions in the mobile app market and uses our search model on the demand side. Pires (2018) studies the effect of search frictions on prices and profits in the laundry detergent market and follows our approach of using choice-set specific logit errors to smooth the choice set probabilities. Finally, De los Santos, Hortaçsu, and Wildenbeest (2012) estimate a related simultaneous search model using individual-specific data in which consumer search behavior and individual choice sets are observed—our approach uses aggregate data in addition, which allows the researcher to deal with price endogeneity. Moreover, we specifically deal with the dimensionality problem that arises when the number of possible choice sets is large, which is not necessary in their application because of the low number of choice sets.

# 2 The Model

## 2.1 Utility, demand, and search costs

Consider a market that consists of $F$ different firms (indexed by $f = 1, 2, \ldots, F$), selling $J$ different products (indexed by $j = 1, 2, \ldots, J$). We assume that product $j$ is sold by a single firm $f$, but allow firms to sell multiple products. Specifically, let firm $f \in \mathcal{F}$ sell a subset of products $\mathcal{G}_f \subset \mathcal{J}$, where $\mathcal{J}$ denotes the set of products and $\mathcal{F}$ represents the set of firms.

We posit that, conditional on having inspected product $j$, consumer $i$'s indirect utility of consuming product $j$ is given by

$$u_{ij} = \alpha_i p_j + x_j' \beta_i + \xi_j + \sigma_\varepsilon \varepsilon_{ij}, \tag{1}$$

where $\alpha_i$ is consumer $i$'s price coefficient, $p_j$ is the price of product $j$, $x_j$ captures product $j$'s attributes, $\beta_i$ is a consumer-specific parameter that captures the marginal utility of each of these attributes, $\xi_j$ captures characteristics not observed by the econometrician, and $\varepsilon_{ij}$ is a consumer-product specific utility shock, with scale parameter $\sigma_\varepsilon$ and not observed by the econometrician either. We assume $\varepsilon_{ij}$ is TIEV distributed across consumers and products and captures whether product $j$ is a good match for consumer $i$. Consumers have the option of not buying any product and opt for the outside option, which gives utility

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

Before inspecting a product $j$ consumers know the characteristics $x_j$ and $\xi_j$ but are not aware of the exact match value $\varepsilon_{ij}$ and the exact price of the product $p_j$. The purpose of search is thus to figure out the realized values of the stochastic utility shocks $\varepsilon_{ij}$ and the actual prices at which products sell. The parameter $\sigma_\varepsilon$ thus captures the relative importance of the unknown part of utility in total utility.

We assume that consumers use a simultaneous search strategy, i.e., consumers choose which subset of firms to visit to maximize their expected utility; once they have visited the chosen firms and have learned all the attributes of the products they are interested in, they decide whether to buy any of the inspected products or else opt for the outside option.[5] We further assume that, when deciding which firms to visit, consumers hold correct conjectures about the (equilibrium)

[5] Although the optimal search strategy often has both simultaneous and sequential elements (Morgan and Manning, 1985), an advantage of a simultaneous search rule is that it allows the searcher to collect information quickly. Recent empirical work by De los Santos, Hortaçsu, and Wildenbeest (2012) and Honka and Chintagunta (2017) demonstrates that observed consumer search behavior on the internet is consistent with simultaneous search.

prices firms charge for their products.[6] Finally, consumers also learn the value of the outside option during the search.

Solving for a consumer's optimal search strategy is a very difficult task because the number of choice sets to be evaluated is increasing exponentially in the number of firms. Chade and Smith (2006) provide a procedure, known as the Marginal Improvement Algorithm (MIA), that finds the solution under some assumptions. One requirement is that sellers can a priori be ranked according to the first- or second-order stochastic dominance criterion. Another requirement is that search costs can only be consumer specific. In most demand estimation problems, however, these assumptions are unlikely to hold. When firms sell different numbers of products, sellers' utility distributions may not be ranked according to the first- or second-order stochastic dominance criterion.[7] Moreover, in many applications search costs will have firm-specific components.[8] In order to solve this problem, we propose to model search costs as follows. Let $\boldsymbol{S}$ be the set of all subsets of firms in $\mathcal{F}$, including the empty set, and let $S$ be an element of $\boldsymbol{S}$. We specify consumer $i$'s search cost for visiting all the firms in the subset $S$, denoted $c_{iS}$, as:

$$c_{iS} = \sum_{f \in S} \kappa \left( t'_{if} \gamma_i \right) + \sigma_\lambda \lambda_{iS}.$$

Here $\kappa$ is a known function, $t_{if}$ is a vector of search cost shifters that are consumer and/or firm specific (such as employment status and distance to the firm), $\gamma_i$ a vector of random coefficients and $\lambda_{iS}$ is a consumer-specific search cost shock for visiting a set of firms $S$, with scale parameter $\sigma_\lambda$ and not observed by the econometrician. It is the addition of this consumer choice-set specific error term to the costs of searching subsets of firms that allows us to solve the search problem. We interpret this search cost shock as choice-set specific variation in search costs that the observed search cost shifters are unable to pick up.

The choice-set specific error term allows us to compute the probability that any given search-set is chosen. This idea is analogous to adding an error term to utility in discrete-choice models. If we further assume that $-\lambda_{iS}$ follows a TIEV distribution, then we can compute the probability with which any subset of dealers is chosen in closed-form, no matter how large the number of available options is.We derive this probability in the next section.

[6] As we will be discussed later, whether consumers observe deviation prices before search or not matters for the computation of the price elasticities. However, it does not affect the estimation of the demand parameters.

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

[8] Think of a firm location or the default language of an online seller.

## 2.2 Optimal simultaneous search and purchase decision

A consumer $i$ first decides which subset of sellers to visit; then, upon visiting the chosen sellers and inspecting the products that are sold at those sellers, she makes a purchase decision. In this section we compute the probability that a given subset of alternatives is inspected by a consumer as well as the probability that a given inspected alternative is purchased.

In order to decide which (subset of) sellers to visit, consumer $i$ must compare the expected gains from searching all the possible subsets of sellers. To allow for the possibility of a non-purchase after searching, we 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 $c_{i\varnothing} = \sigma_\lambda \lambda_{i\varnothing}$.[9]

The expected gain to consumer $i$ from inspecting all the products sold by the sellers in a subset $S$ is equal to:

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

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

$$m_{iS} \equiv \mathbb{E}\left[\max_{j\in\mathcal{G}_f\cup\{0\},\ f\in S}\{u_{ij}\}\right] - \sum_{f\in S}\kappa\left(t'_{if}\gamma_i\right),$$

where, recall, $t_{if}$ is a consumer- and firm-specific vector of variables that affect the cost of visiting firm $f$ (e.g., distance of consumer's home to the firm). Letting $F_\varepsilon$ denote the CDF of $\varepsilon_{ij}$, the random variable

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

has a CDF given by $\prod_{j\in\mathcal{G}_f\cup\{0\},\ f\in S} F_\varepsilon\left((u-\delta_{ij})/\sigma_\varepsilon\right)$, where $\delta_{ij}$ is the utility consumer $i$ derives from alternative $j$, i.e.,

$$\delta_{ij} \equiv \alpha_i p_j + x'_j\beta_i + \xi_j. \tag{2}$$

Using this, we obtain

$$m_{iS} = \varsigma + \sigma_\varepsilon \log\left(1 + \sum_{j\in\mathcal{G}_f,\ f\in S}\exp[\delta_{ij}/\sigma_\varepsilon]\right) - \sum_{f\in S}\kappa\left(t'_{if}\gamma_i\right), \tag{3}$$

where $\varsigma$ is the Euler constant. 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 $\varsigma$. In what follows we drop the Euler

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

constant because it does not affect choices.

Consumer $i$ will pick the subset of sellers to visit that maximizes the expected gain $m_{iS} - \sigma_\lambda \lambda_{iS}$. Denoting consumer $i$'s optimal search-set by $S_i^*$ we have:

$$S_i^* = \arg\max_{S \in \boldsymbol{S}} [m_{iS} - \sigma_\lambda \lambda_{iS}],$$

As discussed above, conditional on the $\lambda_{iS}$'s, the explicit characterisation of the set $S_i^*$ is extremely difficult. However, because $-\lambda_{iS}$ is i.i.d. TIEV distributed, we can compute the probability $S_i^*$ takes on value $S$, which we denote $P_{iS}$:

$$P_{iS} = \frac{\exp[m_{iS}/\sigma_\lambda]}{\sum_{S' \in \boldsymbol{S}} \exp[m_{iS'}/\sigma_\lambda]}. \tag{4}$$

where the sum in the denominator is for all the possible choice sets. Note that this sum may be over a large number of consideration sets; in Section 3.1 we show how to deal with this dimensionality problem.

The expression in equation (4) is the multinomial logit probability that consumer $i$ inspects the products of the sellers contained in the set $S$. Once these products are inspected, the probability that she buys alternative $j$ (sold by one of the visited sellers) is equal to the probability that product $j$ provides the highest utility out of the products of the firms in $S$. Denoting this probability by $P_{ij|S}$, we have:

$$P_{ij|S} = \frac{\exp[\delta_{ij}/\sigma_\varepsilon]}{1 + \sum_{r \in S} \exp[\delta_{ir}/\sigma_\varepsilon]}, \tag{5}$$

where product $r$ is a product sold by one of the firms in $S$.

### 2.3 Individual purchase probabilities and aggregate market shares

Using the law of total probability, the probability that consumer $i$ purchases product $j$ that is sold by seller $f$ is

$$s_{ij} = \sum_{S \in \boldsymbol{S}_f} P_{iS} P_{ij|S}, \tag{6}$$

where $f$ denotes the firm producing $j$ and $\boldsymbol{S}_f \subset \boldsymbol{S}$ is the set of all choice sets containing firm $f$.

Using equations (3) and (4), we can write $P_{iS}$ (for $S \neq \varnothing$) as follows:

$$
\begin{aligned}
P_{iS} &= \frac{\exp[m_{iS}/\sigma_\lambda]}{\sum_{S' \in \boldsymbol{S}} \exp[m_{iS'}/\sigma_\lambda]} \\
&= \frac{\exp\left[\frac{\sigma_\varepsilon}{\sigma_\lambda} \log\left(1 + \sum_{j \in S} \exp\left[\delta_{ij}/\sigma_\varepsilon\right]\right) - \overline{c}_{iS}\right]}{1 + \sum_{S' \in \boldsymbol{S} \setminus \varnothing} \exp\left[\frac{\sigma_\varepsilon}{\sigma_\lambda} \log\left(1 + \sum_{j \in S'} \exp\left[\delta_{ij}/\sigma_\varepsilon\right]\right) - \overline{c}_{iS'}\right]} \\
&= \frac{\left(1 + \sum_{j \in S} \exp\left[\delta_{ij}/\sigma_\varepsilon\right]\right)^{\frac{\sigma_\varepsilon}{\sigma_\lambda}} \exp\left[-\overline{c}_{iS}\right]}{1 + \sum_{S' \in \boldsymbol{S} \setminus \varnothing} \left(1 + \sum_{j \in S'} \exp\left[\delta_{ij}/\sigma_\varepsilon\right]\right)^{\frac{\sigma_\varepsilon}{\sigma_\lambda}} \exp\left[-\overline{c}_{iS'}\right]},
\end{aligned}
\tag{7}
$$

where we use the notation $\overline{c}_{iS} \equiv \sum_{f \in S} \kappa\left(t'_{if}\gamma_i\right)/\sigma_\lambda$. Therefore,

$$
s_{ij} = \exp[\delta_{ij}/\sigma_\varepsilon] \sum_{S \in \boldsymbol{S}_f} \frac{\left(1 + \sum_{j \in S} \exp\left[\delta_{ij}/\sigma_\varepsilon\right]\right)^{\frac{\sigma_\varepsilon}{\sigma_\lambda}-1} \exp\left[-\overline{c}_{iS}\right]}{1 + \sum_{S' \in \boldsymbol{S} \setminus \varnothing} \left(1 + \sum_{j \in S'} \exp\left[\delta_{ij}/\sigma_\varepsilon\right]\right)^{\frac{\sigma_\varepsilon}{\sigma_\lambda}} \exp\left[-\overline{c}_{iS'}\right]}. \tag{8}
$$

Given these individual purchase probabilities, the aggregate probability that product $j$ is bought is equal to the integral:

$$
s_j = \int s_{ij} \; f_\tau(\tau_i) d\tau_i, \tag{9}
$$

where $f_\tau(\tau_i)$ is the joint density function of the random coefficients and the demographic characteristics of consumer $i$ that enter the utility and search cost specifications.

Inspection of equation (8) leads to two important conclusions. First, the scale parameters $\sigma_\varepsilon$ and $\sigma_\lambda$ cannot be separately identified from the utility and search cost coefficients. The reason is that the utility coefficients that enter $\delta_{ij}$ are all divided by $\sigma_\varepsilon$ and the search cost coefficients $\gamma_i$ are all divided by $\sigma_\lambda$. As a result, only the ratio $\sigma_\varepsilon/\sigma_\lambda$ can be separately identified from the utility and search cost coefficients. In what follows, without loss of generality, we therefore normalize $\sigma_\varepsilon$ to one:

$$
s_{ij} = \exp[\delta_{ij}] \sum_{S \in \boldsymbol{S}_f} \frac{\left(1 + \sum_{j \in S} \exp\left[\delta_{ij}\right]\right)^{\frac{1}{\sigma_\lambda}-1} \exp\left[-\overline{c}_{iS}\right]}{1 + \sum_{S' \in \boldsymbol{S} \setminus \varnothing} \left(1 + \sum_{j \in S'} \exp\left[\delta_{ij}\right]\right)^{\frac{1}{\sigma_\lambda}} \exp\left[-\overline{c}_{iS'}\right]}. \tag{10}
$$

Second, the computation of the individual purchase probabilities using equation (10) as is becomes extremely tedious in situations where there is a large number of firms. The reason is twofold: the denominator of equation (10) requires us to sum over all $2^F$ choice sets and the numerator over all $2^{F-1}$ choice sets that contain firm $f$. Depending on the number of firms, these sums may have hundreds of thousands of summands. The existing literature that built on an earlier version of this paper (Moraga-González, Sándor, and Wildenbeest, 2015) avoided this dimensionality problem in various ways. We ourselves normalized the unobserved variance of search costs to $\sigma_\lambda = 1$, in which case the individual purchase probabilities can be simplified and the dimensionality problem disappears.[10] Pires (2018) followed the same approach. Murry and Zhou (2020) did not adopt this normalization but assumed that consumers search across clusters of car sellers, rather than across sellers, and restricted the number of miles a buyer is willing to travel to visit a cluster to 40 miles.[11] Finally, Donna, Pereira, Pires, and Trindade (2022) did not adopt this normalization either but in their empirical application consumers search among only nine retailers. An important contribution of the current paper is to provide a novel Monte Carlo estimator of equation (10) that is accurate and fast even if the number of alternatives in the market is large (see Section 3.1). For example, with ten firms the Monte Carlo estimation takes twenty times less than computing the actual sum (see Section 4).

### 2.4 Utility maximization in a GEV model of discrete choice

In this section we show that our simultaneous search model is equivalent to a standard discrete-choice model in which consumers choose from the $J$ products to maximize their utility. This is the analog of the *eventual purchase theorem* of Armstrong (2017) and Choi, Dai, and Kim (2018) for our setting. We do so by appealing to the so-called generalized extreme value (GEV henceforth) family of discrete-choice models developed by McFadden (1978). The GEV generating function that leads to our simultaneous search model is

$$\begin{aligned} G\left(y_0, y_1, \ldots, y_J\right) &= \sum_{S \in \boldsymbol{S}} \rho_S \left( \sum_{f \in S} \sum_{r \in \mathcal{G}_f} y_r \right)^{1/\sigma_\lambda} \\ &= \sum_{S \in \boldsymbol{S}} \rho_S \left( y_0 + \sum_{f \in S \setminus \{0\}} \sum_{r \in \mathcal{G}_f} y_r \right)^{1/\sigma_\lambda} \quad \text{for} \quad y_0, y_1, \ldots, y_J \geq 0, \end{aligned} \tag{12}$$

[10] In such a special case, we can integrate out the choice-set probabilities and write out the buying probabilities in equation (10) as (see Appendix B):

$$s_{ij} = \frac{\exp\left[\delta_{ij} - \ln\left(1 + \exp\left[\kappa\left(t'_{i\{f\}}\gamma_i\right)\right]\right)\right]}{1 + \sum_{k=1}^{J} \exp\left[\delta_{ik} - \ln\left(1 + \exp\left[\kappa\left(t'_{i\{g\}}\gamma_i\right)\right]\right)\right]}, \tag{11}$$

where $\kappa\left(t'_{i\{f\}}\gamma_i\right)$ contains the search cost of firm $f$ only.

[11] Murry and Zhou (2020) put it this way: "*This restriction dramatically reduces the computational burden of computing consumers' optimal search sets. Otherwise, it is computationally infeasible to compute them by allowing consumers to optimally choose their search sets among 248 clusters.*"

where $\rho_S = \exp(-\bar{c}_S)$ and $\bar{c}_S$ denotes a generic search cost corresponding to choice set $S$. Using the GEV Theorem from Swait (2001), which is slightly more general than that from McFadden (1978), we can verify that $G$ defined in equation (12) satisfies the required conditions, provided that $\sigma_\lambda \geq 1$. Indeed, it is straightforward to verify that $G$ is homogeneous of degree $1/\sigma_\lambda$ and that $\lim_{y_j \to \infty} G(y_0, y_1, \ldots, y_J) = \infty$ for any $j$. Further, using that for any $j_1, \ldots, j_k \in \{0, 1, \ldots, J\}$ that can be the outside alternative or belong to firms $f_1, \ldots, f_\ell$ we have

$$\frac{\partial^k G(y_0, y_1, \ldots, y_J)}{\partial y_{j_1} \ldots \partial y_{j_k}} = \frac{1}{\sigma_\lambda}\left(\frac{1}{\sigma_\lambda} - 1\right) \ldots \left(\frac{1}{\sigma_\lambda} - k + 1\right) \sum_{S \in \boldsymbol{S}_{f_1} \cap \ldots \cap \boldsymbol{S}_{f_\ell}} \rho_S \left(\sum_{f \in S} \sum_{r \in \mathcal{G}_f} y_r\right)^{\frac{1}{\sigma_\lambda} - k},$$

where recall that $\boldsymbol{S}_f$ denotes the set of all choice sets that contain firm $f$. One can easily see that these derivatives are nonnegative if $k$ is odd and nonpositive if $k$ is even, as required. The GEV generating function implies the following joint CDF of the $J+1$ error terms in the utility

$$\begin{aligned} F^{GEV}(\epsilon_0, \epsilon_1, \ldots, \epsilon_J) &= \exp\left[-G\left(\exp[-\epsilon_0], \exp[-\epsilon_1], \ldots, \exp[-\epsilon_J]\right)\right], \quad -\infty < \epsilon_j < \infty; \\ &= \exp\left[-\sum_{S \in \boldsymbol{S}} \left(\sum_{f \in S} \sum_{r \in \mathcal{G}_f} \exp\left[-(\epsilon_r + \sigma_\lambda \bar{c}_S)\right]\right)^{1/\sigma_\lambda}\right], \end{aligned}$$

where the choice-set specific search costs $\bar{c}_S$ act as shifters of the "mean" of the distribution and $\sigma_\lambda$ acts as a scale parameter.

Based on McFadden (1978), we can now claim that in our simultaneous search model, consumer $i$ chooses product $j$ that maximizes the utility $\bar{u}_{ij} = \delta_{ij} + \epsilon_{ij}$, where the vector of error terms $(\epsilon_{i0}, \epsilon_{i1}, \ldots, \epsilon_{iJ})$ has the CDF $F^{GEV}$ with $\bar{c}_S$ replaced by $\bar{c}_{iS}$ (defined below equation (7)). The components of the error vector $(\epsilon_{i0}, \epsilon_{i1}, \ldots, \epsilon_{iJ})$, unlike the $\varepsilon_{ij}$'s from equation (1), can be regarded as usual error terms in the sense that they are observed by consumer $i$ and they are not observed by the econometrician. The probability of choosing $j$ can be derived from $F^{GEV}$ using the following formula from McFadden (1978):

$$P(j) = \frac{\exp[\delta_{ij}]}{\mu G\left(\exp[\delta_{i0}], \exp[\delta_{i1}], \ldots, \exp[\delta_{iJ}]\right)} \frac{\partial G\left(\exp[\delta_{i0}], \exp[\delta_{i1}], \ldots, \exp[\delta_{iJ}]\right)}{\partial y_j}, \tag{13}$$

where in the definition of $G$ from equation (12) we replace $\rho_S$ by $\rho_{iS} = \exp(-\bar{c}_{iS})$. It is straightforward to verify that this is the same as the choice probability derived above.

As in Armstrong (2017) and Choi, Dai, and Kim (2018), the *eventual purchase theorem* for our simultaneous search model implies that it can be reformulated as a discrete-choice problem without search frictions. This means that in order to distinguish between a search model and a full information model, we need search data.

### 2.5 Sovinsky Goeree's (2008) choice set probabilities

Our model can be used to provide a micro-foundation for the choice set probabilities proposed by Sovinsky Goeree (2008), which have been further studied in the consideration set literature (see e.g. Abaluck and Adams-Prassl, 2021). For this purpose, let us assume that when consumers derive their optimal search strategy they do so taking into consideration the search costs and disregarding all information about the utilities associated with each search set. In such a case, the expression $m_{iS}$ from Section 2.2 becomes

$$m_{iS} = -\sum_{f\in S} \kappa\left(t_{if}'\gamma_i\right)$$

and therefore the choice set probability $P_{iS}$ becomes

$$P_{iS} = \frac{\exp\left[-\sum_{f\in S}\kappa\left(t_{if}'\gamma_i\right)/\sigma_\lambda\right]}{\sum_{S'\in\boldsymbol{S}}\exp\left[-\sum_{f\in S'}\kappa\left(t_{if}'\gamma_i\right)/\sigma_\lambda\right]} \equiv \frac{\exp[-\bar{c}_{iS}]}{\sum_{S'\in\boldsymbol{S}}\exp[-\bar{c}_{iS'}]}.$$

Denoting $\bar{c}_{i\{f\}} \equiv \kappa\left(t_{if}'\gamma_i\right)/\sigma_\lambda$, since

$$\exp\left[-\bar{c}_{iS}\right] = \prod_{f\in S}\exp\left[-\bar{c}_{i\{f\}}\right] \tag{14}$$

and

$$\sum_{S\in\boldsymbol{S}}\exp\left[-\bar{c}_{iS}\right] = \sum_{S\in\boldsymbol{S}}\prod_{f\in S}\exp\left[-\bar{c}_{i\{f\}}\right] = \prod_{f\in F}\left(1+\exp\left[-\bar{c}_{i\{f\}}\right]\right),$$

we obtain that

$$P_{iS} = \frac{\prod\limits_{f\in S}\exp\left[-\bar{c}_{i\{f\}}\right]}{\prod\limits_{f\in F}\left(1+\exp\left[-\bar{c}_{i\{f\}}\right]\right)}.$$

Note that this can be written as

$$P_{iS} = \prod_{f\in S}\phi_{if}\prod_{f\notin S}\left(1-\phi_{if}\right), \quad \text{where} \quad \phi_{if} = \frac{\exp\left[-\bar{c}_{i\{f\}}\right]}{1+\exp\left[-\bar{c}_{i\{f\}}\right]} \tag{15}$$

which has the same structure as the consideration set probability in equation (3) of Sovinsky Goeree (2008).

Identification of this model, also called the alternative specific consideration (ASC) model, is studied in Abaluck and Adams-Prassl (2021) (see also Brand and Demirer, 2022). They show that the product structure of the consideration set probabilities allows for the separate identification of preferences and consideration using only choice data. In our search model the consideration set probabilities do not have this product structure unless we adopt the behavioral assumption above that consumers ignore utility when choosing the search set. Even in such a case, the consideration set probabilities in equation (15) are consumer-specific, whereas Abaluck and Adams-Prassl (2021) focus on the case in which consideration set probabilities are product-specific.

# 3 Estimation

We estimate the model using individual-level data on search and purchases, as well as product characteristics and aggregate sales. We first discuss how to estimate the purchase probabilities in our model by Monte Carlo methods, which allows us to deal with the dimensionality problem that arises due to the number of consideration sets increasing exponentially in the number of firms. Next we discuss more generally how to estimate our model using maximum likelihood. We end this section with an informal discussion on identification of the main parameters in our model.

## 3.1 Monte Carlo estimator of the market shares

Market share expressions are needed for the BLP contraction mapping. We start by rewriting the expression for the individual purchase probabilities given in equation (10) as follows:

$$s_{ij} = \frac{\exp\left[\delta_{ij}\right]\exp\left[-\bar{c}_{i\{f\}}\right]}{\left(1+\exp\left[-\bar{c}_{i\{f\}}\right]\right)} \frac{\sum_{S\in\boldsymbol{S}_{-f}}\left(1+\bar{\delta}_{i\{f\}}+\bar{\delta}_{iS}\right)^{\frac{1}{\sigma_\lambda}-1}\frac{\exp[-\bar{c}_{iS}]}{\prod\limits_{g\in\mathcal{F}\setminus\{f\}}\left(1+\exp\left[-\bar{c}_{i\{g\}}\right]\right)}}{\sum_{S'\in\boldsymbol{S}}\left(1+\bar{\delta}_{iS'}\right)^{\frac{1}{\sigma_\lambda}}\frac{\exp[-\bar{c}_{iS'}]}{\prod\limits_{g\in\mathcal{F}}\left(1+\exp\left[-\bar{c}_{i\{g\}}\right]\right)}}, \tag{16}$$

where, for notation simplicity, we write

$$\bar{\delta}_{i\{f\}} \equiv \sum_{j\in\mathcal{G}_f}\exp\left[\delta_{ij}\right], \text{ and } \bar{\delta}_{iS} \equiv \sum_{f\in S}\bar{\delta}_{i\{f\}}.$$

We now show how to easily estimate equation (16) by Monte Carlo. Both the numerator and the denominator of the second fraction in equation (16) contain potentially very large sums. Consider the denominator of such a fraction, which contains the largest sum:

$$D_i \equiv \sum_{S' \in \boldsymbol{S}} \left(1+\bar{\delta}_{iS'}\right)^{\frac{1}{\sigma_\lambda}} \frac{\exp\left[-\bar{c}_{iS'}\right]}{\prod\limits_{g \in \mathcal{F}} \left(1+\exp\left[-\bar{c}_{i\{g\}}\right]\right)}$$

Notice now that, because $\exp\left[-\bar{c}_{iS}\right] = \prod\limits_{f \in S} \exp\left[-\bar{c}_{i\{f\}}\right]$,

$$Q_{iS} \equiv \frac{\exp\left[-\bar{c}_{iS}\right]}{\prod\limits_{g \in \mathcal{F}} \left(1+\exp\left[-\bar{c}_{i\{g\}}\right]\right)} = \prod_{g \in S} \phi_{ig} \prod_{g \notin S} (1-\phi_{ig}), \tag{17}$$

where

$$\phi_{ig} = \frac{\exp\left[-\bar{c}_{i\{g\}}\right]}{1+\exp\left[-\bar{c}_{i\{g\}}\right]},$$

which is a factorization that allows sampling for Monte Carlo estimation of the sum (similar to Sovinsky Goeree, 2008). Further, notice that $\sum_{S' \in \boldsymbol{S}} Q_{iS'} = 1$, which implies that we can regard $\left(Q_{iS'}\right)_{S' \in \boldsymbol{S}}$ as a probability mass function. As a result, we can interpret the denominator $D_i$ as the expectation of a discrete random variable taking on values $\left((1+\bar{\delta}_{iS'})^{\frac{1}{\sigma_\lambda}}\right)_{S' \in \boldsymbol{S}}$ with probabilities $\left(Q_{iS'}\right)_{S' \in \boldsymbol{S}}$.

This interpretation gives rise to our Monte Carlo estimator of $D_i$. Let $\{\mathbf{u}^r\}_{r=1}^R = \{(u_1^r, u_2^r, \ldots, u_F^r)\}_{r=1}^R$ be a sample of size $R$ of uniform $[0,1]$ vectors of dimension $F$ (the number of firms), where each $u_i^r$ is a random draw from $[0,1]$. For a given vector $\mathbf{u}^r$, let $\left(\mathbf{1}\left(u_1^r \leq \phi_{i1}\right), \mathbf{1}\left(u_2^r \leq \phi_{i2}\right), \ldots, \mathbf{1}\left(u_F^r \leq \phi_{iF}\right)\right)$ be the resulting vector of ones and zeros indicating whether a firm is included in the search set $S$ (one) or not (zero). Letting

$$\bar{\delta}_i\left(\mathbf{u}^r\right) \equiv \sum_{g=1}^{F} \mathbf{1}\left(u_g^r \leq \phi_{ig}\right) \exp\left[\delta_{ig}\right],$$

we can then rewrite $D_i$ as:

$$D_i = \sum_{S' \in \boldsymbol{S}} \left(1+\bar{\delta}_{iS'}\right)^{\frac{1}{\sigma_\lambda}} Q_{iS'} = \int_{[0,1]^F} \left(1+\bar{\delta}_i\left(\mathbf{u}\right)\right)^{\frac{1}{\sigma_\lambda}} d\mathbf{u}. \tag{18}$$

Therefore, the Monte Carlo estimator of $D_i$ is

$$\widehat{D}_i = \frac{1}{R} \sum_{r=1}^{R} \left(1+\bar{\delta}_i\left(\mathbf{u}^r\right)\right)^{\frac{1}{\sigma_\lambda}}.$$

It is well-known that the estimator $\widehat{D}_i$ is not continuous in the search costs parameters $\gamma_i$ that enter the $\phi_{ig}$'s because the indicators $\mathbf{1}\left(u_g^r \leq \phi_{ig}\right)$ may jump from 0 to 1 or from 1 to 0 as $\gamma_i$ changes slightly. This is problematic during the estimation of the model. Hence, instead of using the indicators $\mathbf{1}\left(u_g \leq \phi_{ig}\right)$, we use a smoothed version of them given by $\Phi\left(\frac{\phi_{ig}-u_g}{h}\right)$, where $\Phi$ is the standard normal CDF and $h$ is small (e.g., $h = 0.001$ or smaller).[12] Letting

$$\widetilde{\delta}_i\left(\mathbf{u}^r\right) = \sum_{g\in\mathcal{F}} \Phi\left(\frac{\phi_{ig}-u_g^r}{h}\right)\exp\left[\delta_{ig}\right],$$

the smooth Monte Carlo estimator of $D_i$ that we use in estimation is

$$\widetilde{D_i} = \frac{1}{R}\sum_{r=1}^{R}\left(1+\widetilde{\delta}_i\left(\mathbf{u}^r\right)\right)^{\frac{1}{\sigma_\lambda}}. \tag{19}$$

The numerator of the second fraction in equation (16) is

$$N_{if} \equiv \sum_{S\in\boldsymbol{S}_{-f}} \left(1+\bar{\delta}_{i\{f\}}+\bar{\delta}_{iS}\right)^{\frac{1}{\sigma_\lambda}-1} \frac{\exp\left[-\bar{c}_{iS}\right]}{\prod_{g\in\mathcal{F}\setminus\{f\}}\left(1+\exp\left[-\bar{c}_{i\{g\}}\right]\right)}.$$

In Appendix C we show that this can be approximated by the smooth-in-parameters integral

$$\widetilde{N}_{if} = \int_{[0,1]^F}\left(1+\bar{\delta}_{i\{f\}}+\widetilde{\delta}_{if}\left(\mathbf{u}_{-f}\right)\right)^{\frac{1}{\sigma_\lambda}-1} d\mathbf{u}. \tag{20}$$

By collecting the approximations provided in equations (19) and (20) we obtain the following estimator of equation (16):

$$\widetilde{s}_{ij} = \frac{\exp\left[\delta_{ij}\right]\exp\left[-\bar{c}_{i\{f\}}\right]}{\left(1+\exp\left[-\bar{c}_{i\{f\}}\right]\right)}\frac{\widetilde{N}_{if}}{\widetilde{D}_i}. \tag{21}$$

To derive an estimator of the aggregate market share of product $j$ in equation (9) we can now use $\widetilde{s}_{ij}$ instead of $s_{ij}$. We need to integrate out the individuals' unobserved heterogeneity including the random coefficients and demographic characteristics. Hence we obtain the estimator of $s_j$

$$\begin{aligned}\widehat{s}_j &= \int \widetilde{s}_{ij} f_\tau\left(\tau_i\right)d\tau_i = \int \frac{\exp\left[\delta_{ij}\right]\exp\left[-\bar{c}_{i\{f\}}\right]}{\left(1+\exp\left[-\bar{c}_{i\{f\}}\right]\right)}\frac{\widetilde{N}_{if}}{\widetilde{D}_i} f_\tau\left(\tau_i\right)d\tau_i \\ &= \int\int_{[0,1]^F} \frac{\exp\left[\delta_{ij}\right]\exp\left[-\bar{c}_{i\{f\}}\right]}{\left(1+\exp\left[-\bar{c}_{i\{f\}}\right]\right)} \frac{\left(1+\bar{\delta}_{i\{f\}}+\widetilde{\delta}_{if}\left(\mathbf{u}_{-f}\right)\right)^{\frac{1}{\sigma_\lambda}-1}}{\widetilde{D}_i} f_\tau\left(\tau_i\right)d\mathbf{u}d\tau_i. \end{aligned} \tag{22}$$

[12]Here $\Phi$ plays the role of a kernel function and $h$ the role of a smoothing parameter similar to nonparametric kernel estimation. In Appendix F we provide Monte Carlo evidence that justifies our choice of the smoothing parameter value.

This integral can be estimated by Monte Carlo by drawing simultaneously from $\tau_i$ and $\mathbf{u}$. Denote the sample by $(\tau_i, \mathbf{u}_i)_i$, $i = 1, \ldots, ns$; we estimate $s_j$ by

$$\widetilde{s}_j = \frac{1}{ns} \sum_{i=1}^{ns} \frac{\exp[\delta_{ij}] \exp\left[-\bar{c}_{i\{f\}}\right]}{\left(1 + \exp\left[-\bar{c}_{i\{f\}}\right]\right)} \frac{\left(1 + \bar{\delta}_{i\{f\}} + \widetilde{\delta}_{if}(\mathbf{u}_{-fi})\right)^{\frac{1}{\sigma_\lambda}-1}}{\widetilde{D}_i}.$$

We note that by treating the market share integral as a joint integral in equation (22) we can reduce computing time considerably because in this way we only need to deal with two Monte Carlo sums instead of three (i.e. $\widetilde{D}_i$ and $\widetilde{s}_j$ instead of $\widetilde{D}_i$, $\widetilde{N}_{if}$ and $\widetilde{s}_j$).

### 3.2 Estimation approach

Following Goolsbee and Petrin (2004) and Train and Winston (2007), we first efficiently estimate the non-linear demand parameters (variance of preferences and mean and variance of search costs) by maximum likelihood using the individual-level data on search (firms an individual chooses to visit) and purchase (final choice an individual makes). In this step, we do not estimate the product fixed effects $\delta_j$ directly; instead, we follow Berry, Levinsohn, and Pakes (2004) and exploit the data on aggregate market shares to compute them by using the contraction mapping. In Appendix D we show that our system of market share equations is a contraction. In a second step, we estimate the effects of price and product characteristics on (mean) utility using instruments for price as in Berry (1994). This two-step procedure is the maximum likelihood analog to the two-step GMM procedure of Berry, Levinsohn, and Pakes (2004).

In order to estimate the nonlinear demand parameters, we compute the likelihood of the consumers' observed search and purchase behavior. This likelihood is based on the joint probability that consumer $i$ ($i = 1, \ldots, N$) searches the set of sellers $S$ and chooses product $j \in \mathcal{G}_f$ with $f \in S$:

$$\Pr(i \text{ searches } S \text{ and } i \text{ chooses } j) = P_{iS} P_{ij|S}.$$

The probabilities $P_{iS}$ and $P_{ij|S}$ depend on the demand and search cost parameters as well as observed individual characteristics and random coefficients. The probability that consumer $i$, whose individual characteristics are observed, searches $S$ and buys $j$ is

$$s_{iSj}(\theta) = \int P_{iS} P_{ij|S} f_v(v_i)\, dv_i, \tag{23}$$

where $f_v$ is the density function of the random coefficients only. By equations (A27) and (A28) we have

$$
\begin{aligned}
s_{iSj}\left(\theta\right) &= \int \frac{\left(1+\bar{\delta}_{iS}\right)^{\frac{1}{\sigma_\lambda}} \exp\left[-\bar{c}_{iS}\right]}{\sum_{S'\in\boldsymbol{S}} \left(1+\bar{\delta}_{iS'}\right)^{\frac{1}{\sigma_\lambda}} \exp\left[-\bar{c}_{iS'}\right]} \frac{\exp[\delta_{ij}]}{1+\bar{\delta}_{iS}} f_v\left(v_i\right) dv_i \\
&= \int \frac{\exp[\delta_{ij}] \left(1+\bar{\delta}_{iS}\right)^{\frac{1}{\sigma_\lambda}-1} \exp\left[-\bar{c}_{iS}\right]}{\sum_{S'\in\boldsymbol{S}} \left(1+\bar{\delta}_{iS'}\right)^{\frac{1}{\sigma_\lambda}} \exp\left[-\bar{c}_{iS'}\right]} f_v\left(v_i\right) dv_i \\
&= \int \frac{\exp\left[-\bar{c}_{iS}\right]}{\prod\limits_{f\in\mathcal{F}} \left(1+\exp\left[-\bar{c}_{i\{f\}}\right]\right)} \frac{\exp[\delta_{ij}] \left(1+\bar{\delta}_{iS}\right)^{\frac{1}{\sigma_\lambda}-1}}{\frac{\sum_{S'\in\boldsymbol{S}} \left(1+\bar{\delta}_{iS'}\right)^{\frac{1}{\sigma_\lambda}} \exp\left[-\bar{c}_{iS'}\right]}{\prod_{f\in\mathcal{F}} \left(1+\exp\left[-\bar{c}_{i\{f\}}\right]\right)}} f_v\left(v_i\right) dv_i,
\end{aligned}
$$

which, by equation (17), can be written as:

$$
s_{iSj}\left(\theta\right) = \int Q_{iS} \frac{\exp[\delta_{ij}] \left(1+\bar{\delta}_{iS}\right)^{\frac{1}{\sigma_\lambda}-1}}{\sum_{S'\in\boldsymbol{S}} \left(1+\bar{\delta}_{iS'}\right)^{\frac{1}{\sigma_\lambda}} Q_{iS'}} f_v\left(v_i\right) dv_i.
$$

This search and purchase probability can be estimated by Monte Carlo in the same way we estimated the individual purchase probabilities and market shares in Section 3.1. The Monte Carlo estimator is:

$$
\widetilde{s}_{iSj}\left(\theta\right) = \sum_{q=1}^{ns} Q_{iS} \frac{\exp[\delta_{ij}\left(v_q\right)] \left(1+\bar{\delta}_{iS}\left(v_q\right)\right)^{\frac{1}{\sigma_\lambda}-1}}{\widetilde{D}_i\left(v_q\right)},
$$

where $\widetilde{D}_i$ is the smooth estimator of $D_i$ given in equation (19) and the argument $v_q$ reflects explicit dependence of the underlying expressions on the draws from the random coefficients.

Let $y_{iSj}$ be 1 if consumer $i$ searches $S$ and buys $j$ and 0 otherwise. Then the log-likelihood is

$$
L_N\left(\theta\right) = \sum_{i=1}^{N} \sum_{j,S} y_{iSj} \log \widetilde{s}_{iSj}\left(\theta\right),
$$

where the summation $\sum_{j,S}$ is over all possible $j, S$. Because in practice this sum only contains the terms corresponding to $y_{iSj} = 1$, the log-likelihood can also be written as

$$
L_N(\theta) = \sum_{i=1}^{N} \log \widetilde{s}_{iSj}(\theta), \tag{24}
$$

with $S$ and $j$ being $i$'s search set and chosen product, respectively. Obviously, $S$ is allowed to be the empty set and $j$ is allowed to be the outside alternative.

As mentioned above, we do not estimate the product fixed effects $\delta_j$ that enter the likelihood directly. Instead, in every iteration of the maximum likelihood procedure we compute the $\xi_j$ that enter $\delta_{ij}$ from the aggregate data as a function of the parameters. Following BLP, the predicted market share $s_j(\theta)$ of product $j$ should match observed market shares $s_j^o$, or

$$s_j(\delta(\theta), \theta) - s_j^o = 0. \tag{25}$$

In Appendix D we show that when $\sigma_\lambda \geq 1$, this system of equations corresponds to a contraction mapping in $\delta(\theta)$, so it has exactly one solution. Hence, we can compute the components of the vector $\xi = (\xi_1, \ldots, \xi_J)$ of unobserved characteristics as

$$\xi_j = \delta_j(\theta) - \alpha p_j - x_j \beta, \tag{26}$$

where $\delta_j(\theta)$ is component $j$ of the unique fixed point $\delta(\theta)$ of the contraction mapping. It is important to note that when we substitute $\xi_j$ into $\delta_{ij}$ then $\alpha p_j$ and $x_j' \beta$ cancel. Consequently, the log-likelihood in equation (24) will depend on all the demand parameters except for $\alpha$ and $\beta$, and therefore, by maximizing $L_N(\theta) \equiv L_N(\theta_2)$ we can estimate all the demand parameters except for $\alpha$ and $\beta$ by maximum likelihood. The linear utility parameters can be estimated in a second step using either OLS or, when dealing with price endogeneity, two-stage least squares.

### 3.3 Identification

We provide an informal discussion of the identification of the model parameters. In doing so we assume that, similar to estimation, individual-level search and purchase data are available. It is important to note that in the recent literature several papers provide frameworks for identification of demand without observing choice sets (e.g., Abaluck and Adams-Prassl, 2021; Barseghyan, Coughlin, Molinari, and Teitelbaum, 2021; Lu, 2022). As we have pointed out in several places, point-identification of the parameters in our model is not possible in the absence of search data.

There are two main issues that need to be discussed regarding identification. The first is the potential endogeneity of price and search cost variables (e.g., distance from a consumer's home to the seller) while the second is the presence of common covariates in utility and search cost (e.g., the constants, which are important for both searching and buying decisions). Both of these issues can be tackled in the case when search and purchase data are available.

Consider first endogeneity. Recall from BLP that price endogeneity arises due to the fact that usually there are some unobserved characteristics that affect both utility and marginal cost and, therefore, affect also price, causing omitted variable bias. Endogeneity of a search cost variable like distance can occur in a similar way: products can have some characteristics (unobserved to the researcher) that affect both the utility that consumers derive from them and the location where they are sold. A crucial point of identification is to interpret the mean utility $\delta_j$ of product $j$ as a product-specific fixed effect parameter. These fixed effect parameters incorporate the unobserved product characteristics responsible for endogeneity, so endogeneity is eliminated (Berry, Levinsohn, and Pakes, 2004). These fixed effect parameters can be identified together with the variance parameters from the utility and search cost based on individual-level variation in choices and search. Finally, the mean utility parameters included in $\delta_j = \alpha p_j + x_j'\beta + \xi_j$ can be identified as in Berry (1994) or BLP based on the usual conditional moment restriction that the unobserved product characteristics are mean independent of the observed characteristics and excluded (from utility) marginal cost shifters.

Consider the issue of common covariates next. The most unfavorable situation for identification is when a variable that affects both the utility and search cost enters the purchase probabilities as a sum of identical functional forms. To illustrate, one such example arises when $\kappa\left(t_{if}'\gamma_i\right) = \ln\left(\exp\left(t_{if}'\gamma_i\right) - 1\right)$ and $\sigma_\lambda = 1$, in which case common covariates appear as a sum of linear terms. Using equation (A31), this specification leads to the consumer-specific buying probability

$$s_{ij} = \frac{\exp\left[\delta_{ij} - t_{if}'\gamma_i\right]}{1 + \sum_{k=1}^{J} \exp\left[\delta_{ik} - t_{ig}'\gamma_i\right]},$$

which clearly shows that if both $\delta_{ij}$ and $t_{if}'\gamma_i$ contain a constant, these cannot be identified separately based on purchase or market share data, and likewise, if there are common covariates in the utility and the search cost, the coefficients of these cannot be identified separately either.

Now suppose we also observe which firms have been searched. Using the same functional form for $\kappa$ as above as well as the normalization $\sigma_\lambda = 1$, the probability that consumer $i$ searches $S$ and chooses $j$ simplifies to

$$P_{iS}P_{ij|S} = s_{ij} \cdot \frac{\prod_{g \notin S} \left(\exp[t_{ig}'\gamma_i] - 1\right)}{\prod_{g \in \mathcal{F} \setminus f} \exp[t_{ig}'\gamma_i]}.$$

Since the fraction in the second half of this equation only contains the search cost shifters, it is straightforward to see that common covariates should now be separately identified.

# 4 Monte Carlo Experiments

In this section we use Monte Carlo experiments to illustrate the importance of using an approach that allows one to separately identify the utility and search cost constants. We also use the Monte Carlo experiments to show the importance of estimating the variance of search costs. Finally, we study the performance of the Monte Carlo estimator we use to deal with the dimensionality problem.

We use the following setup for the Monte Carlo experiments. We simulate data for 25 different markets, where in the baseline case each market has 4 different firms, each selling one product. We allow for a utility constant as well as a product attribute that is randomly drawn from a normal distribution with mean 2 and standard deviation 0.5, with parameter values as given by the first column in Table 1. The unobserved characteristic $\xi$ is drawn from a normal distribution with mean zero and standard deviation 0.1. We specify the search cost to be linear (that is, we take $\kappa$ to be the identity function). We allow for a consumer-firm specific search cost shifter that is randomly drawn from a lognormal distribution with variance of the variables' logarithm set to one and firm-specific means between 2 and 3.[13] The corresponding parameter values are shown in the first column of Table 1. For each replication we simulate prices, attributes, purchases, searches, and market shares. Prices are obtained by simulating equilibrium prices using the supply side model discussed in Appendix E, assuming consumers have to search for both price realizations and the matching term. As described in Section 3.1, we use a Monte Carlo estimator to estimate the buying probability according to equation (21), using 529 quasi-random draws and the smoothing parameter set to 0.001.

In our first experiment we focus on the separate identification of the utility constant and the search cost constant and show that separate identification of the constants is important for estimates of elasticities and markups, as well as for counterfactuals. We do this in the simplest possible setting, in which the scale parameter $\sigma_\lambda$ is set to one. We estimate the model using a combination of aggregate data and individual data.

The results in column (A) of Table 1 are for when using individual purchase data, whereas column (B) of the table uses search and purchase data. Since it is not possible to separate the search cost constant from the utility constant when not using search data, we estimate only a utility cost constant in column (A), while we estimate both constants when supplementing the individual purchase data with search data. Not surprisingly, the effect of the search cost constant is absorbed

[13]Specifically, the firm-specific means of the logarithm depend on the number of firms and are equally spread between 2 and 3, i.e., with four firms the means are 2, 2.333, 2.666, and 3.

Table 1: Results Monte Carlo Experiments

| Variable | true coeff. | (A) purchase data | (B) search data |
|---|---|---|---|
| *Preference parameters* | | | |
| Constant | -1.000 | -2.102 | -1.029 |
| | | (0.115) | (0.049) |
| Attribute 1 | 2.000 | 1.998 | 1.999 |
| | | (0.018) | (0.018) |
| Price | -2.000 | -1.999 | -1.999 |
| | | (0.011) | (0.010) |
| *Search cost parameters* | | | |
| Constant | 1.500 | | 1.473 |
| | | | (0.059) |
| Shifter | 1.000 | 1.166 | 0.994 |
| | | (0.172) | (0.069) |
| *Elasticities and markups* | | | |
| Own-price elasticity | -5.141 | -5.524 | -5.153 |
| | | (0.207) | (0.225) |
| Markup | 0.617 | 0.539 | 0.613 |
| | | (0.007) | (0.013) |
| *Counterfactual effects of drop in utility* | | | |
| Price (% change) | -4.828 | -1.602 | -4.692 |
| | | (0.243) | (0.544) |

*Notes*: Standard deviations are in parentheses. Specification (A) uses a combination of aggregate data and individual purchase data, whereas specification (B) uses a combination of aggregate data and individual search and purchase data. data are generated for 4 firms and 25 markets. The total number of individual-specific observations used for estimation is 2,500. The number of quasi-random draws for the Monte Carlo estimator is 529 and smoothing parameter value 0.001.

by the utility constant when the search cost constant is not estimated, which in both cases is estimated to be lower than the true value of -1. However, although not being able to estimate the search cost constant affects the estimate of the utility constant, it does not affect the other utility parameters. The results in the table also suggest that when using search data, we can pin down both utility and search cost constants.

Not being able to pin down the search cost constant affects estimates of the own-price elasticity as well as markups. As shown in the table, average own-price elasticities are estimated to be larger in magnitude in the absence of a search cost constant, and average markups are lower. The reason for this is that ignoring the search cost constant is equivalent to assuming that search costs are lower than they really are. Consumers who do not search (and hence do not buy) are incorrectly interpreted as consumers who do search but do not buy. To accommodate this, the utility constant is estimated to be much smaller than it really is. This increases competition between inside products and demand becomes more elastic (and markups lower). These biases may affect counterfactuals such as the impact of an overall decrease in demand. For instance, prices are expected to go down

Table 2: Results Scale Parameter Estimated

| Variable | true coeff. | (A) scale parameter normalized | (B) scale parameter estimated |
|---|---|---|---|
| *Preference parameters* | | | |
| Constant | -1.000 | -0.979 | -0.998 |
| | | (0.053) | (0.047) |
| Attribute 1 | 2.000 | 1.903 | 2.000 |
| | | (0.022) | (0.024) |
| Price | -2.000 | -1.954 | -1.999 |
| | | (0.014) | (0.012) |
| *Search cost parameters* | | | |
| Shifter | 1.000 | 0.339 | 1.070 |
| | | (0.009) | (0.315) |
| Scale parameter | 0.300 | 1.000 | 0.305 |
| | | | (0.084) |
| *Elasticities and markups* | | | |
| Own-price elasticity | -5.172 | -5.097 | -5.168 |
| | | (0.225) | (0.222) |
| Markup | 0.607 | 0.605 | 0.607 |
| | | (0.010) | (0.012) |
| *Counterfactual effets of drop in utility* | | | |
| Price (% change) | -4.349 | -3.817 | -4.387 |
| | | (0.384) | (0.489) |

*Notes*: Standard deviations are in parentheses. Both specifications use a combination of aggregate data and individual search and purchase data. data are generated for 4 firms and 25 markets. The total number of individual-specific observations used for estimation is 2,500. The number of quasi-random draws for the Monte Carlo estimator is 529 and smoothing parameter value 0.001.

by close to 5 percent as a result of a drop in the utility constant of 2 (down from $-1$ to $-3$), but when using the estimates in column (A) simulate the expected price change, the percentage decline in prices is around 1.6 percent. A decrease in the utility constant makes that people substitute from inside products to the outside option. This increases competition between inside products which reduces prices. When we do not estimate a search cost constant, the market is more competitive that it really is. Reducing the utility constant has then less of an impact on prices.

Table 2 gives results for the case in which the true value of the scale parameter $\frac{\sigma_\varepsilon}{\sigma_\lambda}$ is 0.3. In Column (A) of the table, we estimate the model normalizing the scale parameter equal to one, while in Column (B) we estimate the scale parameter. The results shown in columns (A) and (B) are for the case in which individual-level purchase and search data are used to estimate the model. The results in Column (A) show that when the scale parameter is normalized to 1 rather than estimated, both preference and search cost parameters are biased. This bias carries over to the predicted price effects of a drop in aggregate demand. Specifically, we observe demand elasticity estimates that are somewhat lower in magnitude than their true counterparts, and price effects that

Table 3: Results Monte Carlo Experiments

| | | (A) 3 firms | | (B) 5 firms | | (C) 10 firms | |
|---|---|---|---|---|---|---|---|
| Variable | true coeff. | MC est | actual | MC est | actual | MC est | actual |
| *Preference parameters* | | | | | | | |
| Constant | -1.000 | -0.974 | -0.974 | -1.015 | -1.016 | -1.010 | -1.010 |
| | | (0.070) | (0.070) | (0.050) | (0.050) | (0.046) | (0.046) |
| Attribute 1 | 2.000 | 1.995 | 1.996 | 2.000 | 2.002 | 2.004 | 2.006 |
| | | (0.052) | (0.051) | (0.038) | (0.038) | (0.031) | (0.031) |
| Price | -2.000 | -1.997 | -1.998 | -1.999 | -1.999 | -2.001 | -2.002 |
| | | (0.027) | (0.026) | (0.020) | (0.020) | (0.013) | (0.013) |
| *Search cost parameters* | | | | | | | |
| Constant | 1.500 | 1.531 | 1.542 | 1.486 | 1.496 | 1.544 | 1.555 |
| | | (0.351) | (0.349) | (0.239) | (0.237) | (0.187) | (0.186) |
| Shifter | 1.000 | 1.043 | 1.052 | 1.024 | 1.032 | 1.029 | 1.038 |
| | | (0.272) | (0.270) | (0.198) | (0.197) | (0.158) | (0.157) |
| Scale parameter | 1.000 | 1.036 | 1.026 | 1.022 | 1.013 | 0.985 | 0.976 |
| | | (0.228) | (0.222) | (0.175) | (0.171) | (0.133) | (0.130) |
| Estimation time (minutes) | | 3:55 | 1:15 | 6:21 | 5:58 | 12:19 | 253:13 |

Standard deviations are in parentheses. All specifications use a combination of aggregate data and individual search and purchase data. data are generated for 25 markets. The total number of individual-specific observations used for estimation is 2,500. The number of quasi-random draws for the Monte Carlo estimator is 529 and smoothing parameter value 0.001. Estimation time is the average time (in minutes) it takes for one iteration to converge, and is obtained using Matlab 2020a on a Cray XC40 supercomputer.

are underestimated. The intuition for this is as follows. A low variance of search costs is equivalent to a high variance of the portion of utility that is not observable before search, which increases the gains from search and makes the market more competitive. Hence, normalizing the scale parameter to 1 results in lower gains from search. To accommodate the observed intensity of search, search costs have to be estimated much lower than what they truly are, preferences are downward biased and demands are less elastic than what they really are. The results in Column (B) indicate that the scale parameter can be successfully estimated, with more accurate counterfactual results.

Finally, we study the performance of the Monte Carlo estimator for the market shares as well as the individual search and purchase probabilities. This estimator allows us to deal with the dimensionality problem that arises because one has to sum over all possible choice sets to get these probabilities. However, with a small number of firms it is feasible to proceed without using the estimator and use the actual expressions for the market shares and individual search and purchase probabilities. In fact, as shown in column (A) of Table 3, when there are only 3 firms in the market it takes on average 1:15 minutes for a single iteration to converge, which is more than three times faster than the Monte Carlo estimator. Estimation time is about the same when there are 5 firms, but when there are 10 firms the Monte Carlo estimator is about 20 times faster than when using the actual probabilities. This is in line with our expectations since for the Monte Carlo estimator the computing time is expected to increase linearly in the number of firms while for the estimator that uses the actual probabilities the computing time increases exponentially. Moreover, in all three cases the estimates when using the Monte Carlo estimator are almost identical to when using the actual probabilities, so the computational gains do not come at the cost of accuracy.

## 5 Conclusions

This paper has proposed a framework for the estimation of demand for differentiated products with simultaneous consumer search that allows for observed and unobserved heterogeneity in both preferences and search costs. 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 shown that imposing a behavioral assumption on how consumers search in our framework provides a micro foundation for consideration probabilities widely used in the literature.

Demand modeling with simultaneous search for differentiated products poses a few challenges for the modeler. First, finding a consumer's optimal search set is a very complicated task when the number of alternatives is large and the existing methods do not generalize to common choice situations. We have proposed a fix to this problem that consists of adding a TIEV distributed shock to the costs of searching a choice set. We have shown that our model belongs to the GEV class of discrete choice models, which implies that it is consistent with utility maximization and that search data are necessary to distinguish between the search model and the equivalent full information model. Second, estimation of product demand suffers from a dimensionality problem because the product of a firm may be part of a very large number of search sets. To deal with this problem, we have provided a novel Monte Carlo approach and have shown that our estimator of the purchase probabilities is accurate and computationally fast. Finally, we have proposed a maximum likelihood approach to estimate demand and search costs using individual-level data on search and purchases. Monte Carlo experiments have highlighted the importance of allowing for sufficient consumer heterogeneity (in preferences and search costs) when doing policy counterfactuals.

# A Derivation of Expression (16) for the Individual Purchase Probabilities

Based on the discussion from Section 2 and using the notation from Section 2.3, we have $\bar{c}_{i\varnothing} = 0$. Recall from Section 4 that

$$\bar{\delta}_{iS} \equiv \sum_{j \in S} \exp\left[\delta_{ij}\right].$$

and notice that it can be decomposed as follows:

$$\bar{\delta}_{iS} = \bar{\delta}_{i\{f\}} + \bar{\delta}_{iS\setminus\{f\}},$$

where $\bar{\delta}_{i\{f\}} = \sum_{j \in \mathcal{G}_f} \exp\left[\delta_{ij}\right]$ and $\bar{\delta}_{iS\setminus\{f\}} = \sum_{j \in S\setminus\{f\}} \exp\left[\delta_{ij}\right]$. Notice that for $S = \varnothing$, we have $\bar{\delta}_{i\varnothing} = 0$.

Using this notation, we write $P_{iS}$ and $P_{ij|S}$ as:[14]

$$P_{iS} = \frac{\left(1 + \bar{\delta}_{iS}\right)^{\frac{1}{\sigma_\lambda}} \exp\left[-\bar{c}_{iS}\right]}{\sum_{S' \in \boldsymbol{S}} \left(1 + \bar{\delta}_{iS'}\right)^{\frac{1}{\sigma_\lambda}} \exp\left[-\bar{c}_{iS'}\right]}, \tag{A27}$$

$$P_{ij|S} = \frac{\exp[\delta_{ij}]}{1 + \bar{\delta}_{iS}}. \tag{A28}$$

Therefore:

$$\begin{aligned}
s_{ij} &= \sum_{S \in \boldsymbol{S}_f} P_{iS} P_{ij|S} \\
&= \sum_{S \in \boldsymbol{S}_f} \frac{\left(1 + \bar{\delta}_{iS}\right)^{\frac{1}{\sigma_\lambda}} \exp\left[-\bar{c}_{iS}\right]}{\sum_{S' \in \boldsymbol{S}} \left(1 + \bar{\delta}_{iS'}\right)^{\frac{1}{\sigma_\lambda}} \exp\left[-\bar{c}_{iS'}\right]} \frac{\exp[\delta_{ij}]}{1 + \bar{\delta}_{iS}} \\
&= \exp\left[\delta_{ij}\right] \frac{\sum_{S \in \boldsymbol{S}_f} \left(1 + \bar{\delta}_{iS}\right)^{\frac{1}{\sigma_\lambda} - 1} \exp\left[-\bar{c}_{iS}\right]}{\sum_{S' \in \boldsymbol{S}} \left(1 + \bar{\delta}_{iS'}\right)^{\frac{1}{\sigma_\lambda}} \exp\left[-\bar{c}_{iS'}\right]} && \text{(A29)} \\
&= \exp\left[\delta_{ij}\right] \exp\left[-\bar{c}_{i\{f\}}\right] \frac{\sum_{S \in \boldsymbol{S}_{-f}} \left(1 + \bar{\delta}_{i\{f\}} + \bar{\delta}_{iS}\right)^{\frac{1}{\sigma_\lambda} - 1} \exp\left[-\bar{c}_{iS}\right]}{\sum_{S' \in \boldsymbol{S}} \left(1 + \bar{\delta}_{iS'}\right)^{\frac{1}{\sigma_\lambda}} \exp\left[-\bar{c}_{iS'}\right]}, && \text{(A30)}
\end{aligned}$$

where $f$ is the firm producing $j$ and $\boldsymbol{S}_f$ is the set of all choice sets containing firm $f$. The last

[14] The denominator of the fraction in $P_{iS}$ in equation (7) is

$$1 + \sum_{S' \in \boldsymbol{S}\setminus\varnothing} \left(1 + \bar{\delta}_{iS'}\right)^{\frac{1}{\sigma_\lambda}} \exp\left[-\bar{c}_{iS'}\right] = \sum_{S' \in \boldsymbol{S}} \left(1 + \bar{\delta}_{iS'}\right)^{\frac{1}{\sigma_\lambda}} \exp\left[-\bar{c}_{iS'}\right].$$

equality follows from the fact that the numerator of the fraction in $s_{ij}$ in equation (A29) is

$$\sum_{S \in \boldsymbol{S}_f} \left(1 + \bar{\delta}_{iS}\right)^{\frac{1}{\sigma_\lambda} - 1} \exp\left[-\bar{c}_{iS}\right] = \exp\left[-\bar{c}_{i\{f\}}\right] \sum_{S \in \boldsymbol{S}_{-f}} \left(1 + \bar{\delta}_{i\{f\}} + \bar{\delta}_{iS}\right)^{\frac{1}{\sigma_\lambda} - 1} \exp\left[-\bar{c}_{iS}\right],$$

where $\boldsymbol{S}_{-f}$ is the set of all choice sets that do not contain firm $f$.

Now divide both the numerator and denominator of equation (A30) by

$$\sum_{S' \in \boldsymbol{S}} \exp\left[-\bar{c}_{iS'}\right]$$

Noticing that

$$\sum_{S \in \boldsymbol{S}} \exp\left[-\bar{c}_{iS}\right] = \prod_{f \in \mathcal{F}} \left(1 + \exp\left[-\bar{c}_{i\{f\}}\right]\right) = \left(1 + \exp\left[-\bar{c}_{i\{f\}}\right]\right) \prod_{g \in \mathcal{F} \setminus \{f\}} \left(1 + \exp\left[-\bar{c}_{i\{g\}}\right]\right),$$

and

$$\sum_{S \in \boldsymbol{S}_{-f}} \exp\left[-\bar{c}_{iS}\right] = \prod_{g \in \mathcal{F} \setminus \{f\}} \left(1 + \exp\left[-\bar{c}_{i\{g\}}\right]\right)$$

gives expression (16) for the individual purchase probabilities.

# B Derivation of Buying Probabilities under Normalization

Using the normalizations $\sigma_\lambda = \sigma_\varepsilon = 1$ in equation (8) gives

$$s_{ij} = \exp\left[\delta_{ij}\right] \frac{\sum_{S \in \boldsymbol{S}_f} \exp\left[-\bar{c}_{iS}\right]}{\sum_{S' \in \boldsymbol{S}} (1 + \bar{\delta}_{iS'}) \exp[-\bar{c}_{iS'}]}.$$

We get

$$\begin{aligned} s_{ij} &= \exp[\delta_{ij}] \frac{\sum_{S \in \boldsymbol{S}_{-f}} \exp\left[-\bar{c}_{i\{f\}} - \bar{c}_{iS}\right]}{\sum_{S' \in \boldsymbol{S}} (1 + \bar{\delta}_{iS'}) \exp[-\bar{c}_{iS'}]} \\ &= \exp[\delta_{ij}] \frac{\exp\left[-\bar{c}_{i\{f\}}\right] \sum_{S \in \boldsymbol{S}_{-f}} \exp\left[-\bar{c}_{iS}\right]}{\sum_{S' \in \boldsymbol{S}} (1 + \bar{\delta}_{iS'}) \exp[-\bar{c}_{iS'}]} \\ &= \exp[\delta_{ij}] \frac{\exp\left[-\bar{c}_{i\{f\}}\right] \sum_{S \in \boldsymbol{S}_{-f}} \prod_{f \in S} \exp\left[-\bar{c}_{i\{f\}}\right]}{\sum_{S' \in \boldsymbol{S}} (1 + \bar{\delta}_{iS'}) \exp[-\bar{c}_{iS'}]}. \end{aligned}$$

Since

$$\sum_{S\in\boldsymbol{S}_{-f}}\prod_{f\in S}\exp\left[-\bar{c}_{i\{f\}}\right]=\prod_{g\in F\setminus\{f\}}\left(1+\exp\left[-\bar{c}_{i\{f\}}\right]\right)=\frac{\prod\limits_{g\in F}\left(1+\exp\left[-\bar{c}_{i\{g\}}\right]\right)}{\left(1+\exp\left[-\bar{c}_{i\{f\}}\right]\right)},$$

we have that

$$s_{ij}=\exp[\delta_{ij}]\frac{\exp\left[-\bar{c}_{i\{f\}}\right]\frac{\prod_{g\in F}\left(1+\exp\left[-\bar{c}_{i\{g\}}\right]\right)}{\left(1+\exp\left[-\bar{c}_{i\{f\}}\right]\right)}}{\sum_{S'\in\boldsymbol{S}}(1+\bar{\delta}_{iS'})\exp[-\bar{c}_{iS'}]}=\frac{\exp\left[\delta_{ij}-\ln\left(1+\exp\left[\bar{c}_{i\{f\}}\right]\right)\right]\Pi_i}{\sum_{S'\in\boldsymbol{S}}(1+\bar{\delta}_{iS'})\exp[-\bar{c}_{iS'}]},$$

where $\Pi_i=\prod\limits_{g\in F}\left(1+\exp\left[-\bar{c}_{i\{g\}}\right]\right)$.

Now note that $s_{i0}$ is given by

$$\begin{aligned}s_{i0}&=\frac{1+\sum_{S\in\boldsymbol{S}\setminus\{\varnothing\}}\exp\left[-\bar{c}_{iS}\right]}{\sum_{S'\in\boldsymbol{S}}(1+\bar{\delta}_{iS'})\exp[-\bar{c}_{iS'}]}=\frac{\sum_{S\in\boldsymbol{S}}\prod\limits_{f\in S}\exp\left[-\bar{c}_{i\{f\}}\right]}{\sum_{S'\in\boldsymbol{S}}(1+\bar{\delta}_{iS'})\exp[-\bar{c}_{iS'}]}\\&=\frac{\prod\limits_{f\in F}\left(1+\exp\left[-\bar{c}_{i\{f\}}\right]\right)}{\sum_{S'\in\boldsymbol{S}}(1+\bar{\delta}_{iS'})\exp[-\bar{c}_{iS'}]}=\frac{\Pi_i}{\sum_{S'\in\boldsymbol{S}}(1+\bar{\delta}_{iS'})\exp[-\bar{c}_{iS'}]}.\end{aligned}$$

Since $\sum_{j=0}^{J}s_{ij}=1$, it has to be that

$$\begin{aligned}\sum_{S'\in\boldsymbol{S}}(1+\bar{\delta}_{iS'})\exp[-\bar{c}_{iS'}]&=\Pi_i+\sum_{k=1}^{J}\exp\left[\delta_{ik}-\ln\left(1+\exp\left[\bar{c}_{i\{g\}}\right]\right)\right]\Pi_i\\&=\Pi_i\left(1+\sum_{k=1}^{J}\exp\left[\delta_{ik}-\ln\left(1+\exp\left[\bar{c}_{i\{g\}}\right]\right)\right]\right).\end{aligned}$$

Consequently,

$$\begin{aligned}s_{ij}&=\frac{\exp\left[\delta_{ij}-\ln\left(1+\exp\left[\bar{c}_{i\{f\}}\right]\right)\right]}{1+\sum_{k=1}^{J}\exp\left[\delta_{ik}-\ln\left(1+\exp\left[\bar{c}_{i\{g\}}\right]\right)\right]}\\&=\frac{\exp\left[\delta_{ij}-\ln\left(1+\exp\left[\kappa\left(t'_{i\{f\}}\gamma_i\right)\right]\right)\right]}{1+\sum_{k=1}^{J}\exp\left[\delta_{ik}-\ln\left(1+\exp\left[\kappa\left(t'_{i\{g\}}\gamma_i\right)\right]\right)\right]}.\end{aligned}\tag{A31}$$

and

$$s_{i0}=\frac{1}{1+\sum_{k=1}^{J}\exp\left[\delta_{ik}-\ln\left(1+\exp\left[\kappa\left(t'_{i\{g\}}\gamma_i\right)\right]\right)\right]}.$$

## C Monte Carlo Estimation of the Individual Purchase Probabilities

To complete the explanation on how to estimate the individual purchase probabilities by Monte Carlo, we now provide details on how to estimate the numerator of the second fraction in equation (16):

$$N_{if} \equiv \sum_{S \in \boldsymbol{S}_{-f}} \left(1 + \bar{\delta}_{i\{f\}} + \bar{\delta}_{iS}\right)^{\frac{1}{\sigma_\lambda}-1} \frac{\exp\left[-\bar{c}_{iS}\right]}{\prod\limits_{g \in \mathcal{F}\setminus\{f\}} \left(1 + \exp\left[-\bar{c}_{i\{g\}}\right]\right)}.$$

Again, because $\exp\left[-\bar{c}_{iS}\right] = \prod\limits_{f \in S} \exp\left[-\bar{c}_{i\{f\}}\right]$, for $S$ not containing $f$ we have:

$$Q_{ifS} \equiv \frac{\exp\left[-\bar{c}_{iS}\right]}{\prod\limits_{g \in \mathcal{F}\setminus\{f\}} \left(1 + \exp\left[-\bar{c}_{i\{g\}}\right]\right)} = \prod_{g \in S} \frac{\exp\left[-\bar{c}_{i\{g\}}\right]}{1 + \exp\left[-\bar{c}_{i\{g\}}\right]} \prod_{\substack{g \notin S, \\ g \neq f}} \frac{1}{1 + \exp\left[-\bar{c}_{i\{g\}}\right]} = \prod_{g \in S} \phi_{ig} \prod_{\substack{g \notin S, \\ g \neq f}} \left(1 - \phi_{ig}\right),$$

where

$$\phi_{ig} \equiv \frac{\exp\left[-\bar{c}_{i\{g\}}\right]}{1 + \exp\left[-\bar{c}_{i\{g\}}\right]}.$$

Hence, we can write $N_{if}$ as:

$$N_{if} = \sum_{S \in \boldsymbol{S}_{-f}} \left(1 + \bar{\delta}_{i\{f\}} + \bar{\delta}_{iS}\right)^{\frac{1}{\sigma_\lambda}-1} Q_{ifS}.$$

Like it was the case for the denominator, $\sum_{S \in \boldsymbol{S}_{-f}} Q_{ifS} = 1$ so $\left(Q_{ifS}\right)_{S \in \boldsymbol{S}_{-f}}$ can be regarded as a probability mass function. Therefore, $N_{if}$ can be interpreted as the expected value of a discrete random variable taking on values $\left[\left(1 + \bar{\delta}_{i\{f\}} + \bar{\delta}_{iS}\right)^{\frac{1}{\sigma_\lambda}-1}\right]_{S \in \boldsymbol{S}_{-f}}$ with probabilities $\left(Q_{ifS}\right)_{S \in \boldsymbol{S}_{-f}}$. Hence, similar to equation (18), by letting

$$\bar{\delta}_{if}\left(\mathbf{u}_{-f}^{r}\right) = \sum_{g \in \mathcal{F}\setminus\{f\}} \mathbf{1}\left(u_g \leq \phi_{ig}\right) \exp\left[\delta_{ig}\right],$$

we can then rewrite $N_{if}$ as:

$$N_{if} = \int_{[0,1]^{F-1}} \left(1 + \bar{\delta}_{i\{f\}} + \bar{\delta}_{if}\left(\mathbf{u}_{-f}\right)\right)^{\frac{1}{\sigma_\lambda}-1} d\mathbf{u}_{-f}.$$

Note that we can also write $N_{if}$ as

$$N_{if} = \int_{[0,1]^{F}} \left(1 + \bar{\delta}_{i\{f\}} + \bar{\delta}_{if}\left(\mathbf{u}_{-f}\right)\right)^{\frac{1}{\sigma_\lambda}-1} d\mathbf{u},$$

which holds due to the fact that $\int_{[0,1]} 1du = 1$. This way of writing $N_{if}$ is useful because it allows the Monte Carlo estimates of the market shares to depend on the same draws. Like before, to obtain a smooth-in-parameters estimator of $N_{if}$, we use $\Phi\left(\frac{\phi_{ig} - u_g}{h}\right)$ rather than $\mathbf{1}\left(u_g \leq \phi_{ig}\right)$, where $\Phi$ is standard normal CDF and $h$ is a smoothing parameter (see Section 3.1). Let

$$\widetilde{\delta}_{if}\left(\mathbf{u}_{-f}\right) = \sum_{g \in \mathcal{F} \setminus \{f\}} \Phi\left(\frac{\phi_{ig} - u_g}{h}\right) \exp\left[\delta_{ig}\right]$$

and the corresponding smooth-in-parameters version of $N_{if}$

$$\widetilde{N}_{if} = \int_{[0,1]^F} \left(1 + \bar{\delta}_{i\{f\}} + \widetilde{\delta}_{if}\left(\mathbf{u}_{-f}\right)\right)^{\frac{1}{\sigma_\lambda} - 1} d\mathbf{u}.$$

# D The Contraction Property

Here we present a version of the Contraction Theorem from BLP also used by Moraga-González, Sándor, and Wildenbeest (2022). This Contraction Theorem is essentially the same as BLP's but here the conditions are specified on the market share function instead of the contraction mapping itself. We maintain the normalization $\sigma_\varepsilon = 1$. From equation (6), the first order derivatives of $s_{ij}$ with respect to $\delta_j = \alpha p_j + x_j'\beta + \xi_j$ and $\delta_h$ when $j, h \in \mathcal{G}_f$ and when $h \in \mathcal{G}_g,\ g \neq f$, are:

$$\frac{\partial s_{ij}}{\partial \eta_j} = \left(\frac{1}{\sigma_\lambda} - 1\right) \sum_{S \in \boldsymbol{S}_f} P_{iS} P_{ij|S}^2 + \left(1 - \frac{1}{\sigma_\lambda} s_{ij}\right) s_{ij}, \tag{A32}$$

$$\frac{\partial s_{ij}}{\partial \eta_h} = \left(\frac{1}{\sigma_\lambda} - 1\right) \sum_{S \in \boldsymbol{S}_f} P_{iS} P_{ij|S} P_{ih|S} - \frac{1}{\sigma_\lambda} s_{ih} s_{ij} \quad \text{for} \quad j, h \in \mathcal{G}_f, \tag{A33}$$

$$\frac{\partial s_{ij}}{\partial \eta_h} = \left(\frac{1}{\sigma_\lambda} - 1\right) \sum_{S \in \boldsymbol{S}_f \cap \boldsymbol{S}_g} P_{iS} P_{ij|S} P_{ih|S} - \frac{1}{\sigma_\lambda} s_{ih} s_{ij} \quad \text{for} \quad h \in \mathcal{G}_g,\ g \neq f. \tag{A34}$$

The first order derivatives of the choice probability of the outside option $s_{i0}$ with respect to $\eta_j'$ is

$$\frac{\partial s_{i0}}{\partial \eta_j} = \left(\frac{1}{\sigma_\lambda} - 1\right) \sum_{S \in \boldsymbol{S}} P_{iS} P_{i0|S} P_{ij|S} - \frac{1}{\sigma_\lambda} s_{ij} s_{i0}, \tag{A35}$$

where

$$s_{i0} = \sum_{S \in \boldsymbol{S}} P_{iS} P_{i0|S} = \sum_{S \in \boldsymbol{S}} \frac{\exp[m_{iS}/\sigma_\lambda]}{\sum_{S' \in \boldsymbol{S}} \exp[m_{iS'}/\sigma_\lambda]} \frac{1}{1 + \sum_{r \in S} \exp[\delta_{ir}]}. \tag{A36}$$

### D.1 Contraction theorem

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

$$f_j(\delta) = \delta_j + \log s_j^o - \log s_j(\delta), \quad j = 1, \ldots, J,$$

where $s^o = (s_1^o, \ldots, s_J^o)$ is the vector of observed market shares and suppose that the market share vector $s(\delta)$ as a function of $\delta = (\delta_1, \ldots, \delta_J) \in \mathbb{R}^J$ satisfies the following conditions.

1. $s$ is continuously differentiable in $\delta$ and

$$\frac{\partial s_j}{\partial \delta_j}(\delta) \leq s_j(\delta), \quad \frac{\partial s_j}{\partial \delta_k}(\delta) \leq 0 \text{ for any } j,\ k \neq j \text{ and } \delta \in \mathbb{R}^J,$$

(the former is equivalent to the fact that the function $\overline{s}_j : \mathbb{R}^J \to \mathbb{R}$, $\overline{s}_j(\delta) = s_j(\delta)\exp(-\delta_j)$ is decreasing in $\delta_j$) and

$$\sum_{k=1}^{J} \frac{\partial s_j}{\partial \delta_k}(\delta) > 0 \text{ for any } \delta \in \mathbb{R}^J.$$

2. The function $\overline{s}_j$ defined in Condition 1 satisfies

$$\lim_{\delta \to -\infty} \overline{s}_j(\delta) > 0.$$

3. The share of the outside alternative $s_0(\delta) = 1 - \sum_{j=1}^{J} s_j(\delta)$ is decreasing in all its arguments and it satisfies that for any $j$ and $x \in \mathbb{R}$ the limit

$$\lim_{\delta_{-j} \to -\infty} s_0(\delta_1, \ldots, \delta_{j-1}, x, \delta_{j+1}, \ldots, \delta_J) \equiv \widetilde{s}_0^j(x)$$

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

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

where $\delta_{-j} \to -\infty$ means that $\delta_1 \to -\infty, \ldots, \delta_{j-1} \to -\infty, \delta_{j+1} \to -\infty, \ldots, \delta_J \to -\infty$.

Then there are values $\underline{\delta}, \overline{\delta} \in \mathbb{R}$ such that the function $\overline{f} : \left[\underline{\delta}, \overline{\delta}\right]^J \to \mathbb{R}^J$ defined by $\overline{f}_j(\delta) = \min\left[\overline{\delta}, f_j(\delta)\right]$ has the property that $\overline{f}\left(\left[\underline{\delta}, \overline{\delta}\right]^J\right) \subseteq \left[\underline{\delta}, \overline{\delta}\right]^J$, is a contraction with modulus less than

1 with respect to the sup norm $\|(x_1, \ldots, x_J)\| = \max_j |x_j|$, and, in addition, $f$ has no fixed point outside $\left[\underline{\delta}, \overline{\delta}\right]^J$.

## D.2 Verifying the contraction theorem conditions

We next verify the conditions for our market share vector function $s = (s_1, \ldots, s_J)$ for $\sigma_\lambda \geq 1$, where $s_j = \int s_{ij} f_\tau(\tau_i) d\tau_i$, $j = 0, 1, \ldots, J$.

**Condition 1**. The market share vector $s$ is obviouly continuously differentiable in $\delta$.

The inequality $\frac{\partial s_j}{\partial \delta_j} \leq s_j$ holds if (see equation (A32))

$$\left(\frac{1}{\sigma_\lambda} - 1\right) \sum_{S \in \boldsymbol{S}_f} P_{iS} P_{ij|S}^2 + \left(1 - \frac{1}{\sigma_\lambda} s_{ij}\right) s_{ij} - s_{ij} \leq 0,$$

i.e.,

$$\left(\frac{1}{\sigma_\lambda} - 1\right) \sum_{S \in \boldsymbol{S}_f} P_{iS} P_{ij|S}^2 - \frac{1}{\sigma_\lambda} s_{ij}^2 \leq 0.$$

This is clearly true for $\sigma_\lambda \geq 1$.

The inequality $\frac{\partial s_j}{\partial \delta_k} < 0$ holds if (see equations (A33) and (A34))

$$\left(\frac{1}{\sigma_\lambda} - 1\right) \sum_{S \in \boldsymbol{S}_f} P_{iS} P_{ij|S} P_{ih|S} - \frac{1}{\sigma_\lambda} s_{ih} s_{ij} < 0 \quad \text{and}$$
$$\left(\frac{1}{\sigma_\lambda} - 1\right) \sum_{S \in \boldsymbol{S}_f \cap \boldsymbol{S}_g} P_{iS} P_{ij|S} P_{ih|S} - \frac{1}{\sigma_\lambda} s_{ih} s_{ij} < 0 \quad \text{for} \quad h \in \mathcal{G}_g,\ g \neq f.$$

These clearly hold for $\sigma_\lambda \geq 1$.

Finally, to prove $\sum_{k=1}^J \frac{\partial s_j}{\partial \delta_k} > 0$, first note that $\frac{\partial s_j}{\partial \delta_k} = \frac{\partial s_k}{\partial \delta_j}$ (see equations (A33) and (A34)). Then

$$\sum_{k=1}^J \frac{\partial s_j}{\partial \delta_k} = \sum_{k=1}^J \frac{\partial s_k}{\partial \delta_j} = -\frac{\partial s_0}{\partial \delta_j}.$$

If $\sigma_\lambda \geq 1$ then (see equation (A35))

$$\frac{\partial s_{i0}}{\partial \delta_j} = \left(\frac{1}{\sigma_\lambda} - 1\right) \sum_{S \in \boldsymbol{S}} P_{iS} P_{i0|S} P_{ij|S} - \frac{1}{\sigma_\lambda} s_{ij} s_{i0} < 0 \tag{A37}$$

and therefore $-\frac{\partial s_0}{\partial \delta_j} > 0$. This completes the proof of Condition 1.

**Condition 2**. We have

$$\lim_{\delta \to -\infty} s_j \exp\left[-\delta_j\right] = \int \lim_{\delta \to -\infty} s_{ij} \exp\left[-\delta_j\right] f_\tau(\tau_i) d\tau_i.$$

Further,

$$s_{ij} \exp\left[-\delta_j\right] = \sum_{S \in \boldsymbol{S}_f} \frac{\exp[m_{iS}/\sigma_\lambda]}{\sum_{S' \in \boldsymbol{S}} \exp[m_{iS'}/\sigma_\lambda]} \frac{\exp[d_{ij}]}{1 + \sum_{g \in S} \exp[\delta_{ig}]},$$

where $d_{ij} = \delta_{ij} - \delta_j$ does not depend on $\delta_j$. Since by equation (3)

$$\lim_{\delta \to -\infty} \exp[m_{iS}/\sigma_\lambda] = \exp\left(-\bar{c}_{iS}\right), \tag{A38}$$

$$\lim_{\delta \to -\infty} \frac{1}{1 + \sum_{g \in S} \sum_{h \in \mathcal{G}_g} \exp[\delta_{ih}]} = 1, \tag{A39}$$

we get that

$$\lim_{\delta \to -\infty} s_{ij} \exp\left[-\delta_j\right] = \sum_{S \in \boldsymbol{S}_f} \frac{\exp\left[-\bar{c}_{iS}\right]}{\sum_{S' \in \boldsymbol{S}} \exp\left[-\bar{c}_{iS'}\right]} \exp[d_{ij}] > 0.$$

So Condition 2 holds.

**Condition 3**. The fact that the share of the outside alternative $s_0 = 1 - \sum_{j=1}^{J} s_j$ is decreasing in all its arguments follows from (A37) for $\sigma_\lambda \geq 1$.

In order to see that $\widetilde{s}_0^j(x)$ finite, first notice that the limit $\lim_{\delta_{-j} \to -\infty} s_0$ is (see equation (A36))

$$\int \lim_{\delta_{-j} \to -\infty} \sum_{S \in \boldsymbol{S}} \frac{\exp[m_{iS}/\sigma_\lambda]}{\sum_{S' \in \boldsymbol{S}} \exp[m_{iS'}/\sigma_\lambda]} \frac{1}{1 + \sum_{g \in S} \sum_{h \in \mathcal{G}_g} \exp[\delta_{ih}]} f_\tau(\tau_i) d\tau_i.$$

If $j \in \mathcal{G}_f$,

$$\lim_{\delta_{-j} \to -\infty} \exp[m_{iS}/\sigma_\lambda] = \begin{cases} \exp\left[\log\left(1 + \exp[\delta_{ij}]\right)/\sigma_\lambda - \bar{c}_{iS}\right] & \text{if } f \in S \\ \exp\left(-\bar{c}_{iS}\right) & \text{if } f \notin S \end{cases}$$

and

$$\lim_{\delta_{-j} \to -\infty} \frac{1}{1 + \sum_{g \in S} \sum_{h \in \mathcal{G}_g} \exp[\delta_{ih}]} = \begin{cases} \dfrac{1}{1 + \exp[\delta_{ij}]} & \text{if } f \in S \\ 1 & \text{if } f \notin S. \end{cases}$$

So

$$\lim_{\delta_{-j} \to -\infty} \sum_{S \in \boldsymbol{S}} \exp[m_{iS}/\sigma_\lambda] = \sum_{S \in \boldsymbol{S}_f} \exp\left[\log\left(1 + \exp[\delta_{ij}]\right)/\sigma_\lambda - \bar{c}_{iS}\right] + \sum_{S \notin \boldsymbol{S}_f} \exp\left(-\bar{c}_{iS}\right).$$

Since all these limits exist and are strictly positive, $\lim_{\delta_{-j} \to -\infty} s_0$ will be finite.

The limit $\lim_{x\to-\infty} \widetilde{s}_0^j(x)$ is

$$\int \lim_{\delta\to-\infty} \sum_{S\in\boldsymbol{S}} \frac{\exp[m_{iS}/\sigma_\lambda]}{\sum_{S'\in\boldsymbol{S}} \exp[m_{iS'}/\sigma_\lambda]} \frac{1}{1+\sum_{g\in S}\sum_{h\in\mathcal{G}_g} \exp[\delta_{ih}]} f_\tau(\tau_i) d\tau_i.$$

From equations (A38) and (A39)

$$\lim_{\delta\to-\infty} \sum_{S\in\boldsymbol{S}} \frac{\exp[m_{iS}/\sigma_\lambda]}{\sum_{S'\in\boldsymbol{S}} \exp[m_{iS'}/\sigma_\lambda]} \frac{1}{1+\sum_{g\in S}\sum_{h\in\mathcal{G}_g} \exp[\delta_{ih}]} = \lim_{\delta\to-\infty} \sum_{S\in\boldsymbol{S}} \frac{\exp[m_{iS}/\sigma_\lambda]}{\sum_{S'\in\boldsymbol{S}} \exp[m_{iS'}/\sigma_\lambda]} = 1.$$

The limit $\lim_{x\to\infty} \widetilde{s}_0^j(x)$ is

$$\int \lim_{\delta_j\to\infty} \left( \lim_{\delta_{-j}\to-\infty} \sum_{S\in\boldsymbol{S}} \frac{\exp[m_{iS}]}{\sum_{S'\in\boldsymbol{S}} \exp[m_{iS'}]} \frac{1}{1+\sum_{g\in S}\sum_{h\in\mathcal{G}_g} \exp[\delta_{ih}]} \right) f_\tau(\tau_i) d\tau_i.$$

Note that with $j \in \mathcal{G}_f$

$$\lim_{\delta_j\to\infty} \left( \lim_{\delta_{-j}\to-\infty} \exp[m_{iS}/\sigma_\lambda] \right) = \begin{cases} \infty & \text{if } f \in S \\ \exp\left(-\bar{c}_{iS}\right) & \text{if } f \notin S, \end{cases}$$

so

$$\lim_{\delta_j\to\infty} \left( \lim_{\delta_{-j}\to-\infty} \sum_{S\in\boldsymbol{S}} \exp[m_{iS}/\sigma_\lambda] \right) = \infty.$$

Also

$$\lim_{\delta_j\to\infty} \left( \lim_{\delta_{-j}\to-\infty} \frac{1}{1+\sum_{g\in S}\sum_{h\in\mathcal{G}_g} \exp[\delta_{ih}]} \right) = \begin{cases} 0 & \text{if } f \in S \\ 1 & \text{if } f \notin S, \end{cases}$$

so

$$\lim_{\delta_{-j}\to-\infty} \sum_{S\in\boldsymbol{S}} \frac{\exp[m_{iS}/\sigma_\lambda]}{\sum_{S'\in\boldsymbol{S}} \exp[m_{iS'}/\sigma_\lambda]} \lim_{\delta_j\to\infty} \frac{1}{1+\sum_{g\in S}\sum_{h\in\mathcal{G}_g} \exp[\delta_{ih}]} = \lim_{\delta_{-j}\to-\infty} \sum_{S\notin\boldsymbol{S}_f} \frac{\exp[m_{iS}/\sigma_\lambda]}{\sum_{S'\in\boldsymbol{S}} \exp[m_{iS'}/\sigma_\lambda]}.$$

Therefore,

$$\begin{aligned}
&\lim_{\delta_j\to\infty} \left( \lim_{\delta_{-j}\to-\infty} \sum_{S\in\boldsymbol{S}} \frac{\exp[m_{iS}/\sigma_\lambda]}{\sum_{S'\in\boldsymbol{S}} \exp[m_{iS'}/\sigma_\lambda]} \frac{1}{1+\sum_{g\in S}\sum_{h\in\mathcal{G}_g} \exp[\delta_{ih}]} \right) \\
&= \lim_{\delta_j\to\infty} \left( \lim_{\delta_{-j}\to-\infty} \sum_{S\notin\boldsymbol{S}_f} \frac{\exp[m_{iS}/\sigma_\lambda]}{\sum_{S'\in\boldsymbol{S}} \exp[m_{iS'}/\sigma_\lambda]} \right) \\
&= \sum_{S\notin\boldsymbol{S}_f} \frac{\exp\left(-\bar{c}_{iS}\right)}{\infty} = 0.
\end{aligned}$$

So Condition 3 is satisfied. □

# E Supply Side

We assume manufacturers maximize profits in a pricing game. Assuming a Nash equilibrium exists for this game, any product sold should have prices that satisfy the first order conditions

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

We assume deviation prices are *not* not observable. Since this implies $\partial P_{iS}/\partial p_k = 0$, we get

$$\frac{\partial s_j}{\partial p_k} = \int \sum_{S \in \mathcal{S}_j} \left( P_{iS} \frac{\partial P_{ij|S}}{\partial p_k} \right) f(\tau_i) d\tau_i.$$

This means

$$\frac{\partial s_j}{\partial p_k} = \begin{cases} \int -\alpha_i \left( s_{ij} - \sum_{S \in \mathcal{S}_j} P_{iS} P_{ij|S}^2 \right) f(\tau_i) d\tau_i & \text{if } k = j; \\ \int -\alpha_i \left( - \sum_{S \in \mathcal{S}_j} P_{iS} P_{ij|S} P_{ik|S} \right) f(\tau_i) d\tau_i & \text{if } k \neq j. \end{cases} \tag{A40}$$

To estimate the marginal costs, the first order conditions can be rewritten as in BLP

$$p - \Delta(p)^{-1} s(p) = mc, \tag{A41}$$

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

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

# F The Smoothing Parameter for the Monte Carlo Estimator

In this section we present Monte Carlo evidence on how the Monte Carlo estimation error depends on the smoothing parameter. One of the main questions here is whether the performance of the Monte Carlo estimator is sensitive to the choice of the smoothing parameter. To answer this, we study the performance of the estimator in (19) for estimating sums of the type (18). Specifically, we estimate

$$D = \sum_{S \in \boldsymbol{S}} \left( 1 + \sum_{f \in S} \delta_f \right)^{\frac{1}{\sigma_\lambda}} \prod_{f \in S} \phi_f \prod_{f \notin S} (1 - \phi_f) = \int_{[0,1]^F} \left( 1 + \sum_{f=1}^{F} \mathbf{1} \left( u_f \leq \phi_f \right) \exp\left[ \delta_f \right] \right)^{\frac{1}{\sigma_\lambda}} d\mathbf{u},$$

where $\phi_f = \frac{\exp[-c_f]}{1+\exp[-c_f]}$, by

$$\widetilde{D} = \frac{1}{R} \sum_{r=1}^{R} \left( 1 + \sum_{f \in \mathcal{F}} \Phi \left( \frac{\phi_f - u_f^r}{h} \right) \exp[\delta_f] \right)^{\frac{1}{\sigma_\lambda}}, \tag{A42}$$

where $\Phi$ is the standard normal CDF, $h$ is a smoothing parameter, and $\{(u_1^r, u_2^r, \ldots, u_F^r)\}_{r=1}^R \subset [0,1]^F$ is a quasi-random sample of size $R$ of type $(0, 2, s)$-net (e.g., Sándor and András, 2004).[15]

For different values of $F$ (reported in Table A1) we draw vectors $(\delta_1, \ldots, \delta_F)$ and $(c_1, \ldots, c_F)$ randomly such that $\delta_f \sim N(0, 25)$ and $c_f \sim N(0, 1)$. We compute (A42) 100 times by using different randomized versions of the quasi-random sample; we also calculate the actual value of $D$ and we multiply the ratio of the two by 1000. This way we normalize $D$ to 1000, so the Monte Carlo estimation errors can be compared across the different cases. Based on the estimates computed 100 times we calculate the root-mean-squared-error (hereafter RMSE). Finally, we replicate this procedure 10 times and report the means and standard deviations (Stds) of the 10 RMSE values in Table A1.

We do these computations for various values of $\sigma_\lambda$, $h$, and $R$ (see Table A1; the $h$ values appear in the first line called "Bandwidths"). In the table we can see that for cases in which the smoothing parameter is lower than $10^{-2}$ the RMSE's are rather low. This implies that both the bias and the standard deviation is small in these cases. The smoothing parameter values that yield the lowest mean RMSE values are the two middle values, that is, $h = 10^{-3}$ and $3 \cdot 10^{-4}$. In the vast majority of the cases taking smoothing parameter values lower than these does not increase the mean RMSE's significantly.

We can also notice that in most cases the RMSE's are lower for higher sample sizes $R$, which is something we would expect. Changing the scale parameter $\sigma_\lambda$ does not seem to alter significantly the magnitudes of the RMSE's. We can notice a similar phenomenon when we look at changes in the number of firms $F$. In conclusion, based on the results from Table A1, when we take a smoothing parameter lower than or equal to $10^{-3}$, the Monte Carlo estimator (A42) is expected to have both bias and standard deviation lower than 0.4% of the true value of $D$. This remarkable precision explains why in Table 3 the parameter estimates obtained when using the Monte Carlo estimator are almost identical to those when using the actual probabilities.

[15] Using this type of quasi-random sample turns out to reduce computing time at least by a factor of 40 with respect to a pseudo-random sample because even the lowest sample (of size 256) yields more precise estimates than a pseudo-random sample of size $10,000$.

Table A1: Means and Standard Deviations of RMSE's

| $\sigma_\lambda$ | $R$ | Bandwidths | $3 \cdot 10^{-2}$ | $1 \cdot 10^{-2}$ | $3 \cdot 10^{-3}$ | $1 \cdot 10^{-3}$ | $3 \cdot 10^{-4}$ | $1 \cdot 10^{-4}$ | $3 \cdot 10^{-5}$ | $1 \cdot 10^{-5}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Number of firms $F = 5$ | | | | | |
| 5 | 256 | **Means** | **30.59** | **10.47** | **3.68** | **2.11** | **2.00** | **2.21** | **2.16** | **2.19** |
| | | Stds | 25.42 | 8.55 | 2.74 | 1.24 | 1.01 | 1.33 | 1.23 | 1.31 |
| | 529 | **Means** | **26.65** | **9.06** | **2.91** | **1.40** | **1.07** | **1.04** | **1.00** | **0.99** |
| | | Stds | 24.79 | 8.24 | 2.43 | 0.86 | 0.63 | 0.65 | 0.69 | 0.61 |
| | 1024 | **Means** | **28.71** | **9.72** | **2.97** | **1.16** | **0.65** | **0.60** | **0.61** | **0.66** |
| | | Stds | 21.00 | 7.08 | 2.14 | 0.80 | 0.51 | 0.41 | 0.42 | 0.45 |
| 2 | 256 | **Means** | **17.15** | **5.79** | **2.05** | **1.53** | **1.64** | **1.69** | **1.75** | **1.85** |
| | | Stds | 11.26 | 3.71 | 1.07 | 0.68 | 0.86 | 1.01 | 1.03 | 1.06 |
| | 529 | **Means** | **17.31** | **5.90** | **1.99** | **1.16** | **1.04** | **1.09** | **1.24** | **1.23** |
| | | Stds | 9.31 | 3.07 | 0.86 | 0.38 | 0.37 | 0.34 | 0.38 | 0.42 |
| | 1024 | **Means** | **26.58** | **8.88** | **2.71** | **0.96** | **0.55** | **0.63** | **0.68** | **0.69** |
| | | Stds | 19.17 | 6.39 | 1.90 | 0.59 | 0.25 | 0.38 | 0.44 | 0.44 |
| 1.25 | 256 | **Means** | **9.55** | **3.35** | **1.69** | **2.03** | **2.16** | **2.29** | **2.33** | **2.37** |
| | | Stds | 9.63 | 3.31 | 1.35 | 1.48 | 1.68 | 1.94 | 1.90 | 1.82 |
| | 529 | **Means** | **11.19** | **3.77** | **1.36** | **1.05** | **1.45** | **1.73** | **1.79** | **1.80** |
| | | Stds | 6.52 | 2.16 | 0.70 | 0.46 | 0.67 | 0.81 | 0.86 | 0.82 |
| | 1024 | **Means** | **8.78** | **2.95** | **0.91** | **0.46** | **0.50** | **0.60** | **0.70** | **0.69** |
| | | Stds | 5.93 | 2.00 | 0.59 | 0.29 | 0.25 | 0.28 | 0.34 | 0.33 |
| | | | | | Number of firms $F = 10$ | | | | | |
| 5 | 256 | **Means** | **9.29** | **3.19** | **1.16** | **0.81** | **0.81** | **0.75** | **0.72** | **0.76** |
| | | Stds | 12.02 | 4.10 | 1.21 | 0.64 | 0.50 | 0.43 | 0.40 | 0.42 |
| | 529 | **Means** | **9.60** | **3.28** | **1.10** | **0.59** | **0.52** | **0.55** | **0.57** | **0.57** |
| | | Stds | 12.17 | 4.12 | 1.25 | 0.47 | 0.30 | 0.35 | 0.35 | 0.33 |
| | 1024 | **Means** | **14.81** | **4.95** | **1.51** | **0.55** | **0.29** | **0.28** | **0.29** | **0.28** |
| | | Stds | 15.03 | 5.01 | 1.51 | 0.51 | 0.24 | 0.23 | 0.24 | 0.24 |
| 2 | 256 | **Means** | **22.91** | **7.84** | **2.80** | **2.25** | **2.63** | **2.79** | **2.90** | **2.86** |
| | | Stds | 23.13 | 7.98 | 2.80 | 2.17 | 2.58 | 2.56 | 2.69 | 2.52 |
| | 529 | **Means** | **16.45** | **5.51** | **1.68** | **0.77** | **0.65** | **0.62** | **0.65** | **0.60** |
| | | Stds | 26.10 | 8.72 | 2.56 | 0.99 | 0.65 | 0.60 | 0.57 | 0.59 |
| | 1024 | **Means** | **10.86** | **3.63** | **1.12** | **0.44** | **0.34** | **0.36** | **0.35** | **0.38** |
| | | Stds | 14.78 | 4.93 | 1.48 | 0.53 | 0.33 | 0.34 | 0.32 | 0.33 |
| 1.25 | 256 | **Means** | **5.96** | **2.10** | **1.21** | **1.62** | **1.93** | **2.00** | **2.07** | **2.14** |
| | | Stds | 9.02 | 3.07 | 1.49 | 1.97 | 2.33 | 2.26 | 2.59 | 2.65 |
| | 529 | **Means** | **4.17** | **1.37** | **0.50** | **0.47** | **0.65** | **0.82** | **0.90** | **0.90** |
| | | Stds | 4.60 | 1.56 | 0.49 | 0.35 | 0.47 | 0.58 | 0.65 | 0.72 |
| | 1024 | **Means** | **4.77** | **1.60** | **0.51** | **0.28** | **0.31** | **0.39** | **0.43** | **0.45** |
| | | Stds | 6.37 | 2.12 | 0.62 | 0.26 | 0.26 | 0.35 | 0.40 | 0.40 |
| | | | | | Number of firms $F = 15$ | | | | | |
| 5 | 256 | **Means** | **20.05** | **7.08** | **3.04** | **2.36** | **2.29** | **2.33** | **2.33** | **2.46** |
| | | Stds | 10.44 | 3.48 | 1.18 | 0.65 | 0.74 | 0.78 | 0.80 | 0.81 |
| | 529 | **Means** | **29.40** | **10.16** | **3.50** | **1.84** | **1.67** | **1.63** | **1.65** | **1.66** |
| | | Stds | 21.73 | 7.48 | 2.32 | 1.05 | 1.11 | 1.00 | 1.07 | 1.01 |
| | 1024 | **Means** | **14.99** | **5.24** | **1.76** | **1.00** | **0.85** | **0.83** | **0.89** | **0.82** |
| | | Stds | 8.84 | 2.98 | 0.86 | 0.36 | 0.35 | 0.39 | 0.43 | 0.32 |
| 2 | 256 | **Means** | **23.28** | **8.06** | **3.09** | **2.22** | **2.68** | **2.85** | **3.11** | **3.02** |
| | | Stds | 15.55 | 5.02 | 1.23 | 0.70 | 1.16 | 1.19 | 1.24 | 1.22 |
| | 529 | **Means** | **22.35** | **7.63** | **2.59** | **1.36** | **1.23** | **1.33** | **1.43** | **1.50** |
| | | Stds | 11.46 | 3.67 | 0.91 | 0.39 | 0.51 | 0.54 | 0.46 | 0.49 |
| | 1024 | **Means** | **16.32** | **5.56** | **1.91** | **1.00** | **0.77** | **0.82** | **0.84** | **0.85** |
| | | Stds | 14.32 | 4.74 | 1.31 | 0.36 | 0.30 | 0.31 | 0.34 | 0.32 |
| 1.25 | 256 | **Means** | **5.45** | **1.92** | **1.17** | **1.50** | **1.81** | **1.93** | **2.00** | **2.11** |
| | | Stds | 3.21 | 1.07 | 0.48 | 0.63 | 1.01 | 1.04 | 1.04 | 0.96 |
| | 529 | **Means** | **10.69** | **3.62** | **1.31** | **0.95** | **1.19** | **1.28** | **1.34** | **1.37** |
| | | Stds | 8.73 | 2.95 | 1.03 | 0.59 | 0.67 | 0.75 | 0.95 | 0.86 |
| | 1024 | **Means** | **5.25** | **1.78** | **0.62** | **0.39** | **0.45** | **0.53** | **0.60** | **0.60** |
| | | Stds | 3.15 | 1.05 | 0.30 | 0.13 | 0.13 | 0.14 | 0.17 | 0.16 |

## References


Abaluck, Jason and Abi Adams-Prassl: "What Do Consumers Consider Before they Choose? Identification from Asymmetric Demand Responses," *The Quarterly Journal of Economics* 136, 1611–1663, 2021.

Amemiya, Takeshi: "Qualitative Response Models: A Survey," *Journal of Economic Literature* 19, 1483–1536, 1981.

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 André de Palma: "Multiproduct Firms: A Nested Logit Approach," *The Journal of Industrial Economics* 40, 261–176, 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, Mark: "Ordered Consumer Search," *Journal of European Economic Association* 15, 989–1024, 2017.

Barseghyan, Levon, Maura Coughlin, Francesca Molinari, and Joshua C. Teitelbaum: "Heterogeneous Choice Sets and Preferences," *Econometrica* 89, 2015–2048, 2021.

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

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

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

Brand, James and Mert Demirer: "Approximation-Free Estimation of Preferences and Consideration with Many Products," Mimeo, 2022.

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

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

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

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

Donna, Javier D., Pedro Pereira, Tiago Pires, and André Trindade: “Measuring the Welfare of Intermediaries,” *Management Science* 68, 8083-8115, 2022.

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

Ershov, Daniel: “The Effects of Consumer Search Costs on Entry and Quality in the Mobile App Market,” Mimeo, 2018.

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

Grieco, Paul L. E., Charles Murry, Joris Pinkse, and Stephan Sagl: “Conformant and Efficient Estimation of Discrete Choice Demand Models,” Mimeo, 2022.

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

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

Honka, Elisabeth and Pradeep Chintagunta: “Simultaneous or Sequential? Search Strategies in the US Auto Insurance Industry,” *Marketing Science* 36, 21–40, 2017.

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

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

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

Lin, Haizhen and Matthijs Wildenbeest: "Non-Parametric Estimation of Search Costs for Differentiated Products: Evidence from Medigap," *Journal of Business & Economic Statistics* 38, 754–770, 2020.

Lu, Zhentong: "Estimating Multinomial Choice Models with Unobserved Choice Sets," *Journal of Econometrics* 226, 368–398, 2022.

McFadden, Daniel: "Modeling the Choice of Residential Location," in A. Karlqvist, L. Lundqvist, F. Snickars, and J. Weibull, eds., *Spatial Interaction Theory and Planning Models*, North-Holland, Amsterdam, 75–96, 1978.

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, Zsolt Sándor, and Matthijs R. Wildenbeest: "Consumer Search and Prices in the Automobile Market," CEPR Discussion Paper No. 10487, 2015. (Accessible at https://cepr.org/publications/dp10487.)

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

Moraga-González, José Luis, Zsolt Sándor, and Matthijs R. Wildenbeest: "Simultaneous Search for Differentiated Products: the Impact of Search Costs and Firm Prominence," *The Economic Journal* 131, 1308–1330, 2021.

Moraga-González, José Luis, Zsolt Sándor, and Matthijs R. Wildenbeest: "Consumer Search and Prices in the Automobile Market," forthcoming in the *Review of Economic Studies*.

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

Murry, Charles and Yiyi Zhou: "Consumer Search and Automobile Dealer Co-Location," *Management Science* 66, 1909–1934, 2020.

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

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

Pires, Tiago: "Measuring the Effects of Search Costs on Equilibrium Prices and Profits," *International Journal of Industrial Organization* 60 179–205, 2018.

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

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

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

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

Swait, Joffre: "Choice Set Generation within the Generalized Extreme Value Family of Discrete Choice Models," *Transportation Research Part B* 35, 643–666, 2001.

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

Ursu, Raluca M.: "The Power of Rankings: Quantifying the Effect of Rankings on Online Consumer Search and Purchase Decisions." *Marketing Science* 37, 530–552, 2018.

Yavorsky, Dan, Elisabeth Honka, and Keith Chen: "Consumer Search in the U.S. Auto Industry: The Role of Dealership Visits," *Quantitative Marketing and Econonomics* 19, 1–52, 2021.

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

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