Dynamic Selection and Distributional Bounds on Search Costs in Dynamic Unit-Demand Models

Jason R. Blevins and Garrett T. Senney
The Ohio State University, Department of Economics
Working Paper 14-02
October 29, 2014

Abstract. This paper develops a dynamic model of consumer search that, despite placing very little structure on the dynamic problem faced by consumers, allows us to exploit intertemporal variation in within-period price and search cost distributions to estimate the population distribution from which consumers’ search costs are initially drawn. We show that static approaches to estimating this distribution generally suffer from a dynamic sample selection bias because forward-looking consumers with unit demand for a good may delay their purchase in a way that depends on their individual search cost. We analyze identification of the population search cost distribution using only price data and develop estimable nonparametric upper and lower bounds on the distribution function and a nonlinear least squares estimator for parametric models. We also consider the additional identifying power of weak assumptions such as monotonicity of purchase probabilities in search costs. We apply our estimators to analyze the online market for two widely used econometrics textbooks. Our results suggest that static estimates of the search cost distribution are biased upwards, in a distributional sense, relative to the true population distribution. In a small-scale simulation study, we show that this is typical in a dynamic setting where consumers with high search costs are more likely to delay purchase than those with lower search costs.

Keywords: nonsequential search, consumer search, dynamic selection, nonparametric bounds.

JEL Classification: C57, C14, D83, D43.

1. Introduction

As Benham (1972) so beautifully articulates, “the full cost of the purchase of a good . . . includes not only the cost of the item itself, but also the cost of knowledge, time, and

*We thank Stephen Cosslett for useful comments and Matthijs Wildenbeest for sharing Matlab code for maximum likelihood estimation of search cost distributions.
transportation”. This costly information acquisition generates situations of incomplete information which disrupt the law of one price, even in seemingly competitive markets for homogeneous products. The seminal work of Stigler (1961) makes one of the first theoretically founded arguments for this effect of informational asymmetry on equilibrium market outcomes. Stigler remarks that “price dispersion is a manifestation . . . of ignorance in the market” caused by a lack of information. Since his call for further work on this issue, countless studies have recognized consumer search as a significant factor in explaining the observed price dispersion.1 In this paper we propose a dynamic model with forward looking consumers and consider estimation of the population consumer search cost distribution. We show that if consumers have the option to delay purchase until a later period, as they do in our model, this results in a dynamic selection problem that will cause estimates obtained from a static model to be biased. We illustrate this both in a small-scale simulation study and in an application to online textbook markets.

Early work in the consumer search literature focused on developing models that reconciled theory with the observation that price dispersion is a stable equilibrium outcome.2 Varian (1980), Salop and Stiglitz (1982), and Burdett and Judd (1983) all derived price dispersion as a consequence of having strictly positive proportions of informed consumers and uninformed consumers in the same market. Even in markets for homogeneous products, firms possess some market power and price dispersion arises if there is some consumer heterogeneity.3 Furthermore, Burdett and Judd (1983) show that only ex post heterogeneity is required for price dispersion. This theoretical development allows for informational asymmetry to arise endogenously; for example, consumers may rationally collect different amounts of information about the market in accordance with an optimal search rule.

Some work has even provided theoretical examples of situations when markets with imperfect information could actually be harmed by competition-promoting policies. In models such as that of Varian (1980) and Stahl (1989), increasing the number of firms in the market increases average price. Moraga-González, Sandor, and Wildenbeest (2010) and Janssen and Moraga-González (2004) illustrate how profoundly the shape of the search cost distribution can affect search behavior, prices, and consumer welfare, sometimes even non-linearly. Armstrong (2008) argues that uninformed consumers can cause a strong negative externality on the rest of the market which can lead to the aforementioned coun-

---

1Refer to Pratt, Wise, and Zeckhauser (1979), Dahlby and West (1986), Clay, Krishnan, and Wolff (2001), Johnson (2002), and Moraga-González, Sandor, and Wildenbeest (2012) for examples of empirical studies.
2See McMillan and Rothschild (1994) and Baye, Morgan, and Scholten (2006) for surveys of the consumer search literature.
3Diamond (1971) showed that if all consumers have the same positive search cost then all firms will charge the monopoly price and hence there will be no price dispersion.
terintuitive effects of competition-promoting policies. It is thus apparent that knowledge of the search technology in use in a market is critical to ensuring that any government intervention or program to increase consumer welfare have the desired effects.

In many cases, researchers only have access to price data and have no information about search costs or search behavior, which are inherently difficult to measure. As a result of these data limitations, most of the empirical work on consumer search has focused on merely documenting the incidence and magnitude of price dispersion or uses indirect approaches to estimation which do not require observing search costs. Sorensen (2000) for example, makes theoretical arguments that allow one to identify markets with lower and higher average search costs, which in turn lead to measurably different levels of price dispersion.

More recent works have developed methodologies to estimate the static population search cost distribution using only the observed price distribution. Hong and Shum (2006) and Moraga-González and Wildenbeest (2008) (henceforth MGW) use the equilibrium conditions on supply and demand to consistently estimate the search cost distribution at a finite number of points which are endogenously determined. Moraga-González, Sandor, and Wildenbeest (2013) estimate the search distribution more completely on a compact interval by pooling multiple markets with the same search technology and exploiting the variation in valuations and marginal costs across markets. Wildenbeest (2011) tackles the problem when homogeneous products are sold by vertically differentiated sellers. De los Santos, Hortaçsu, and Wildenbeest (2013) estimate search costs for a differentiated product in a model where consumers learn about their utility.

However, a limitation of most of the current empirical consumer search models is that search process is static. If in reality the process is dynamic, with consumers searching for multiple periods before purchasing, then estimates from a static model will suffer from dynamic selection bias. In this paper we introduce a model with forward-looking consumers where the period search cost distribution evolves over time as new consumers, with costs drawn from a time-invariant population distribution, mix with the consumers who have not yet purchased. This dynamic model allows us to track and model how the distribution of search costs of consumers in the market evolves over time. Simulations suggest that when the consumers’ probability of purchase is monotonically decreasing in cost, selection effects will cause the period search cost distribution to be biased upwards relative to the population distribution. Additionally, this paper also proposes a procedure to estimate non-parametric bounds on the distribution of search cost from which new entrants are drawn. We also consider parametric estimation in cases where the population distribution lies in a known parametric family of distributions.

4See Moraga-González (2006) for a survey of the emerging structural consumer search literature.
Finally, we use our model to analyze the online textbook market. We use daily prices collected from a large cross-section of online retailers for new hardcover copies of two widely-used econometrics textbooks. We find that the median of the distribution of search costs for consumers in the market for these books is much lower than the estimates from a static analysis would suggest. Our estimates are also significantly lower than similar search cost estimates reported previously in the literature. We conclude that accounting for dynamic selection effects, due to consumers with low search costs making purchases and leaving the market more quickly, is important for estimating consumer search costs.

2. The Theoretical Model

2.1. Firm Pricing and Consumer Search

We examine a search model of oligopolistic firms and forward-looking consumers. There are $J$ identical firms producing a homogeneous, durable good at a constant within-period marginal cost $r_t > 0$. Following the previous literature, firms set prices in order to maximize static profits in period $t$. In equilibrium, firms set prices according to a symmetric mixed strategy represented by a cumulative distribution function $F_t$ which is absolutely continuous and assigns positive density everywhere on $[p_t^L, p_t^H]$. Moraga-González et al. (2010) prove that with some mild conditions, an equilibrium will always exist for any number of firms. They also show that for a given observed consumer search behavior, there is a unique symmetric equilibrium where the equilibrium price distribution must be atomless. Although they cannot prove a more general uniqueness proof, Moraga-González et al. (2010) run simulations that suggest that uniqueness is a more general result.

Firms choose prices simultaneously at the beginning of each period according to $F_t$.

In each period, there is a unit measure of forward looking, ex ante identical consumers with unit demand and within-period reservation value $v_t > r_t$. Each consumer’s per-firm search cost $c$ is an i.i.d. random variable with a continuous distribution function $G_t$ with positive density everywhere on the non-negative real line. Consumers retain their search cost for their entire existence as active shoppers.

We assume that consumers observe the price distribution $F_t$ and search simultaneously, receiving a chosen number of price draws from $F_t$. The early literature refers to this technology as fixed sample size search since consumers commit to sample from a predetermined number of firms before purchasing. Although Morgan and Manning (1985) prove that the optimal search strategy is generally a hybrid of both simultaneous and sequential search, Manning and Morgan (1982) demonstrate that if there exists meaningful

---

5 Baye, Kovenock, and De Vries (1992) examine asymmetric equilibria in Varian’s model and conclude that only the symmetric equilibrium survives meaningful equilibrium refinement.
economies of scale to sampling or a significant time lag in information procurement, then sampling once from a large number of firms is the optimal search strategy. Let $K \leq J$ denote the maximum number of firms a consumer can search and let $k \leq K$ denote the optimal choice of the number of firms to search (i.e., the number of firms in a consumer’s consideration set).

Each period consists of three phases: firm pricing, consumer searching, and consumer purchasing. At the beginning of each period, firms observe the distribution of search costs $G_t$ for consumers who demand the product and choose prices according to the symmetric mixed pricing strategy $F_t$ on $[p^u, p^l]$. Next, following the literature, consumers receive one free price quote from a random firm. Consumers then decide how intensely to search by choosing the optimal value of $k$ given their individual search cost $c$. They may choose to search up to $K - 1$ additional firms to receive a total of $k \leq K$ price observations per period. For each additional firm searched beyond the first, consumers incur search cost $c$. Hence, the total cost for becoming a fully informed consumer is $c(K - 1)$. Moraga-González et al. (2013, Proposition 3) shows that the optimal consumer behavior to any atomless price distribution involves ex-post information asymmetry.

After gathering prices, the consumer makes a purchase decision denoted by $d_t \in \{0, 1\}$. A consumer who purchases ($d_t = 1$) does so from the firm with the lowest price of the $k$ firms sampled. Since the product is durable and consumers have unit demand, a consumer who purchases in the current period exits the market. At the end of each period, consumers who purchase, and thus exit, are replaced by consumers with i.i.d. search costs drawn from a continuous, time-invariant population distribution $H$ with full support on the non-negative real line. Consumers who do not purchase remain, retaining their current search cost.

A rational consumer would solve this problem via backward induction by first calculating the optimal number of firms to search by minimizing their expected total expenditure:

$$k_t(c) = \arg\min_{k \leq K} \left[ c(k - 1) + \int_{p^u}^{p^l} pk (1 - F_t(p))^{k-1} f_t(p) \, dp \right].$$

The first term in the minimand is the total cost of search and the second term is the expected minimum price paid after sampling $k$ firms conditional on purchasing. For

---

6Moraga-González et al. (2013) show that in a setting with search cost heterogeneity, this assumption does not qualitatively affect the results. A costly first search will weakly decrease market size as some consumers’ costs might be large enough to prohibit participation in the market.

7We consider a model where the measure of consumers is fixed over time, however, one could also consider allowing expansion and contraction of the market size over time. This would change the relative proportion of consumers entering the market from the population distribution $H$ and would mitigate or exacerbate the dynamic selection bias effects as the case may be.
notational convenience, we rewrite the expected value of the minimum of \( k \) sampled prices in period \( t \) (the first order statistic) as \( E[p_{t,k}^{(1)}] \). The tension in the consumer’s problem is that searching additional firms lowers the expected price paid, but increases the search cost incurred. The total expenditure function is increasing in \( c \) and convex in \( k \); therefore, there is a unique solution.

Given that \( k \) is required to be a positive integer less than or equal to \( K \), consumers are partitioned by their search intensity. The cost that makes a consumer indifferent between searching \( k \) and \( k+1 \) firms is the marginal benefit of searching the \((k+1)\)-th firm;

\[
c_{t,k} = E[p_{t,k}^{(1)}] - E[p_{t,k+1}^{(1)}].
\]

(2)

The marginal benefit of search is non-increasing in \( k \), so the sequence of cutoff search costs is decreasing in \( k \): \( c_{t,k} < c_{t,k-1} < \ldots < c_{t,2} < c_{t,1} \). Let \( \mu_{t,k} \) be the measure of consumers who search exactly \( k \) firms. Combining \( \sum_{k} \mu_{t,k} = 1 \) with the ordering of the indifference costs allows the proportion of consumers searching \( k \) firms to be written as a function of the period search cost distribution:

\[
\mu_{t,k} = \begin{cases} 
1 - G_t(c_{t,1}) & k = 1, \\
G_t(c_{t,k-1}) - G(c_{t,k}) & 2 \leq k \leq K - 1, \\
G_t(c_{t,K-1}) & k = K.
\end{cases}
\]

(3)

Unfortunately, price data from a single market is insufficient to identify the upper quantiles of the search cost distribution; we can only identify the smallest cost that makes a consumer want to only search one firm. The behavior of a consumer with search cost of \( c_{t,1} \) and one with any cost larger than \( c_{t,1} \) are observationally equivalent: they both only search one firm.

Firms in our model behave as in the static models put forth by Hong and Shum (2006), Moraga-González and Wildenbeest (2008), Moraga-González et al. (2013), and others. That is, they are myopic and therefore simply choose prices to maximize profits in a single period. Given the aforementioned consumer search behavior, the firms’ optimal strategy is a symmetric mixed strategy in prices represented by the c.d.f. \( F_t \) on \([p_t, p_t']\). For the firms to be willing to mix in any one period all prices in the support must yield the same profit:

\[
\Pi_t = (p - r_t) \left[ \sum_{k=1}^{K} \frac{k \mu_{t,k}}{K} (1 - F_t(p))^{k-1} \right] \text{ for all } p \in [p_t, p_t'].
\]

(4)

Firms charging the highest price only sell to consumers who only search once. This defines the constant period profit level: \( \Pi_t = \frac{\mu_{t,1}(p_t - r_t)}{K} \). \( F_t(p) \) is then implicitly defined by (4) and
the density of the price distribution can be determined by solving the first order condition of the profit equation for $f_t(p)$:

$$ f_t(p) = \frac{\sum_{k=1}^{K} k \mu_{t,k} (1 - F_t(p))^{k-1}}{(p - r_t) \sum_{k=2}^{K} k(k-1) \mu_{t,k} (1 - F_t(p))^{k-2}}. $$

(5)

Given these density values, one can consistently estimate $\mu_t = (\mu_{t,1}, \ldots, \mu_{t,K})$ in each period using maximum likelihood (Moraga-González and Wildenbeest, 2008). By evaluating (4) at $p = p_0$, this condition can be rearranged and solved to obtain the per unit marginal cost:

$$ r_t = \frac{p_0 \cdot (\sum_{k=1}^{K} k \cdot \mu_k) - \mu_1 p_t}{\sum_{k=2}^{K} k \cdot \mu_k}. $$

(6)

2.2. Forward-Looking Consumers

From the point of view of a forward-looking consumer, next period’s price distribution depends on next period’s search cost distribution, which in turn depends on the decisions of other consumers this period (the policy function), the current search cost distribution ($G_t$), and the population search cost distribution ($H$). This suggests the possibility of a rational expectations equilibrium concept to determine the optimal policy functions $\sigma_t$, where $D_t$ is an indicator for whether the consumer purchases and $\sigma_t(c) = \Pr(D_t = 0 \mid c)$ is the conditional probability of staying in the market (not purchasing). Although we take a largely model-free approach which does not require us to know how the policies $\sigma_t$ are determined, we take a brief look at one possible model simply to provide some intuition for how $\sigma_t$ might be determined. We note that because the policy functions $\sigma_t$ are indexed by $t$, the non-purchase probability may depend on the individual search cost $c$ as well as the population distribution $H$ and the within-period search cost distributions $G_t$ for specific periods. Furthermore, this allows for very general forms of nonstationary, time- or duration-dependent purchase behavior.

Let $I_t$ denote the information set of consumers in the market in period $t$ (which contains the search cost distribution, the price distribution, etc.) and let $W_t(p, c)$ denote the value function for a consumer with an initial (costless) price quote $p$ in hand and who has search cost $c$. Assume for a moment that consumers commit to purchasing before searching and suppose that consumers’ intertemporal discount factors are equal to $\beta$. The value function $W_t(p, c)$ in this setting can be written recursively as

$$ W_t(p, c) = \max_{\max_{k \in \{1, \ldots, K\}} \left\{ p - E \left[p_{t,k}^{(1)} \mid p \right] - (k - 1)c \right\}, \beta E \left[W_{t+1}(p', c) \mid I_t\right]}.$$

This value represents the expected utility received by the consumer from behaving optimally in terms of purchase and search behavior after the initial price draw but before the
purchase and search decisions are made. The outer maximization problem represents the choice of whether to purchase now or wait. The first inner term is the value of purchasing now, either from the first firm ($k = 1$) or from the lowest price firm after searching $k - 1$ additional stores. The second inner term is the expected discounted continuation value given today’s information $I_t$. What is contained in the information set $I_t$ determines how consumers form beliefs about prices tomorrow ($p'$). For example, $I_t$ will most likely contain today’s search cost distribution ($G_t$), the population search cost distribution ($H$), and the decision rules that other consumers in the market follow in determining whether to buy or wait. Such a model quickly becomes analytically and numerically intractable, so rather than limit our analysis to any particular model we simply consider a generic model which admits some sequence of policy functions $\sigma_t : [0, \infty) \to [0,1]$ for $t = 1, 2, \ldots, T$ without placing further restrictions on them.

2.3. A Model-Free Approach to the Consumer’s Dynamic Problem

When thinking about the dynamic process across periods in this model it is helpful to consider a single consumer making a decision about whether to purchase this period or remain in the market until the next period. Any particular dynamic model for consumers will yield a policy rule determining consumer purchase behavior.

If the consumer does not purchase ($D = 0$), then she remains in the market and retains her search cost ($C' = c$). However if the consumer does purchase ($D = 1$), she leaves the market and is replaced by a consumer with a search cost drawn from the population distribution ($C' \sim H(\cdot)$). The conditional search cost distribution next period is

$$\Pr(C' \leq c' \mid C = c, D = d) = \begin{cases} 1\{c \leq c'\} & \text{if } d = 0, \\ H(c') & \text{if } d = 1. \end{cases}$$

(7) If the consumer does not purchase, then her search cost stays the same (i.e., the distribution is a mass point at $c$ with distribution function $1\{c \leq \cdot\}$). When the consumer does purchase, the search cost of the consumer who replaces her is independent of the previous cost (i.e., the cdf is $H(\cdot)$, which does not depend on $c$).

The object of interest in this model is the unconditional population search cost distribution $H$. To find this distribution we need to integrate (7) with respect to $c$ and $d$. This is slightly complicated by the fact that this is a mixed discrete-continuous distribution. For notational convenience, let $\sigma_t(c)$ be the conditional probability of not purchasing in period $t$ ($D_t = 0$) given search cost $C = c$. Then, the unconditional distribution of search costs in period $t + 1$ is

$$G_{t+1}(c') = \int_0^\infty \left[ \sigma_t(c) 1\{c \leq c'\} + (1 - \sigma_t(c)) H(c') \right] g_t(c) \, dc.$$  

(8)
In the first term the indicator function only has the effect of limiting the domain of integration. In the second term, the population distribution is independent of the search cost this period so it can be brought out of the integral. The remaining integral in the second term is over the entire support of \( C \); therefore, it becomes the unconditional probability of purchasing this period. This value is simply a constant every period. After these simplifications, (8) can be rewritten as:

\[
G_{t+1}(c') = \int_0^{c'} \sigma_t(c) g_t(c) \, dc + H(c') \int_0^\infty (1 - \sigma_t(c)) g_t(c) \, dc. \tag{9}
\]

Next period’s search cost distribution can be thought of as a mixture of the current period distribution and the population distribution, with the weights determined by the probability of a consumer purchasing and thus exiting the market. For illustration, consider the two extreme cases. If all the consumers purchase, regardless of their search costs, then next period’s distribution is simply the population distribution \( H \). On the other hand, if no consumers purchase then next period’s search cost distribution is the same as this period’s, namely, \( G_t \).

3. Identification

In this section, we consider identification of the population search cost distribution \( H \) given the sequence of observable per-period price distributions \( \{F_t\} \). We show that with only this limited data and no additional assumptions on the structure of the model (i.e., the \( \sigma_t \) functions are unrestricted), we can identify informative bounds on \( H \).

Moraga-González et al. (2013, Proposition 2) show that in a static model, if one observes price distributions across many markets with the same search cost distribution but with variation in valuations or marginal costs, then the search cost distribution is nonparametrically identified. However, applying their result in the dynamic setting would require that the search cost distributions are the same across markets in all periods, implying that consumer purchase decisions are deterministic. Since this is unlikely to be the case and since there is a single market for online textbooks in our application, we focus on the case of a single market.

A distinct, but related literature studies models of labor search. This literature is generally able to identify the full wage distribution, even below the workers’ reservation wage where no offers are accepted, by assuming a functional form that is recoverable from the observed, truncated distribution (the lognormal family is an example) or observing a random sample of wages offers spanning the full distribution.\(^8\) However, unlike our framework, sequential search is the standard technology for these labor models where

---

\(^8\)See Mortensen (1986) and Eckstein and van den Berg (2007) for surveys of this literature.
workers receive one wage offer at a time. Although there may exist a different distribution for unemployed and employed workers, these wage distributions are generally time-invariant. The wage offers are therefore only informative on that stationary distribution. These techniques, unfortunately, are not directly applicable to our model where consumers and firms are facing distinct price and search costs distributions, respectively, every period.

Our identification strategy is constructive in the following sense. Using the sequence of price distributions, we can first identify a finite number of points of support \( \{ c_{t,k} : t = 1, \ldots, T, k = 1, \ldots K \} \) from (2). Next, we can identify the per-period distributions \( G_t \) at these points, \( \{ G_t(c_{t,k}) : t = 1, \ldots, T, k = 1, \ldots K \} \) by Proposition 1 of Moraga-González et al. (2013). We then use the fact that the functions \( \sigma_t \) are probability policy functions to bound \( H(c) \) at arbitrary points \( c \). Additionally, we further refine the bounds by exploiting variation across periods without making further assumptions about the purchase probability functions (e.g., monotonicity) or using any additional data. Finally, we can refine the bounds more by assuming that the probability of purchase is monotonic in search cost (i.e., consumers with higher search costs are less likely to purchase).

3.1. Bounds with Unrestricted Policies

More specifically, suppose that the values \( c_{t,k} \) and \( G(c_{t,k}) \) are identified for all \( t = 1, \ldots, T \) and \( k = 1, \ldots K \). Let \( \Sigma \) be the space of all functions mapping \( [0, \infty) \to [0, 1] \). We know that \( \sigma_t \in \Sigma \), but because we have not fully specified the dynamic problem of the consumer, we cannot more fully characterize the consumers’ conditional probability of purchasing. However, we show that even knowing that each \( \sigma_t \) is a probability function, and is therefore bounded between zero and one, provides information about \( H \).

In principle, we can vary \( \sigma_t \) over the space \( \Sigma \) and determine the two functions that maximize and minimize \( G_{t+1}(c') \) at any arbitrary point \( c' \) as defined in (8). For a given cost \( c' \), let \( \sigma^L_t(\cdot; c') \) denote the function for which \( G_{t+1}(c') \) reaches the lower bound and let \( \sigma^U_t(\cdot; c') \) denote the function which achieves the upper bound. The non-purchase probabilities that generate the lower and upper bounds on \( G_{t+1}(c') \) and the implied bounds are given by the following lemma.

**Lemma 1.** When \( \sigma_t \) is unrestricted in \( \Sigma \), the conditional, non-purchase probabilities that generate the lower and upper bounds on \( G_{t+1} \) for a given cost \( c' \) are

\[
\sigma^L_t(c; c') = \begin{cases} 0 & \text{if } c \leq c', \\ 1 & \text{if } c > c', \end{cases} \quad \text{and} \quad \sigma^U_t(c; c') = \begin{cases} 1 & \text{if } c \leq c', \\ 0 & \text{if } c > c'. \end{cases}
\]

The implied bounds on \( G_{t+1}(c') \) are

\[
G^L_{t+1}(c') \equiv G_t(c')H(c') \leq G_{t+1}(c') \leq G_t(c') + [1 - G_t(c')]H(c') \equiv G^U_{t+1}(c').
\]
Proof. First, note that (8) is an integral equation of the form \( G_{t+1}(c') = \int_0^\infty L_t(c, \sigma_t(c), \sigma'_t(c)) \, dc \) where

\[
L_t(c, \sigma_t(c), \sigma'_t(c)) = [\sigma_t(c) \mathbb{1}\{c \leq c'\} + (1 - \sigma_t(c)) \, H(c')] \, g_t(c).
\]

We can use basic techniques from the calculus of variations to find policy functions \( \sigma_t \) that maximize and minimize \( G_{t+1}(c') \) at each \( c' \). Let \( L_{t,j}(c, \sigma_t(c), \sigma'_t(c)) \) denote the partial derivative of \( L_t \) with respect to the \( j \)-th argument. Then the Euler-Lagrange equation is

\[
L_{t,2}(c, \sigma_t(c), \sigma'_t(c)) - \frac{d}{dc} L_{t,3}(c, \sigma_t(c), \sigma'_t(c)) = 0
\]

and in this case we have \( L_{t,3}(c, \sigma_t(c), \sigma'_t(c)) = 0 \) and

\[
L_{t,2}(c, \sigma_t(c), \sigma'_t(c)) = \begin{cases} 
(1 - H(c'))g_t(c), & c \leq c', \\
-H(c')g_t(c), & c > c'. 
\end{cases}
\]

For all \( c' \in (0, \infty) \), \( (1 - H(c')) > 0 > -H(c') \), so we can see immediately that of functions in the space \( \Sigma \), \( \sigma^L_t(c; c') = 1\{c \leq c'\} \) will maximize \( G_{t+1}(c') \) while \( \sigma^U_t(c; c') = 1\{c \geq c'\} \) will minimize \( G_{t+1}(c') \).

Given these bounding policy functions, we can simplify (9) and evaluate the implied bounds on \( G_{t+1}(c') \) for each \( c' \). For the lower bound, the first term in the integral vanishes since \( \sigma^L_t(c; c') = 0 \) when \( c < c' \). The second term just becomes the integral of this period’s search cost density up to \( c' \) multiplied by the entry distribution evaluated at \( c' \). For the upper bound, the fact that \( \sigma^U_t(c; c') = 0 \) when \( c > c' \) effectively limits the domain of integration for the second term.

Given that we can estimate the per-period search cost distributions \( (G_t) \), we would like to use this information to learn about the population distribution \( (H) \). For a cost value \( c \) for which \( G_t(c) \) and \( G_{t+1}(c) \) are both identified and \( G_t(c) \in (0,1) \), the bounds above can be rearranged to find bounds on \( H(c) \):

\[
H^L_{t+1}(c) \equiv \frac{G_{t+1}(c) - G_t(c)}{1-G_t(c)} \leq H(c) \leq \frac{G_{t+1}(c)}{G_t(c)} \equiv H^U_{t+1}(c). 
\]

These bounds on the unknown population distribution are a function of only the period search cost distributions, which are estimable. Using a bounded approach immediately begs the question of whether the bounds \([H^L_{t+1}(c), H^U_{t+1}(c)]\) yield a uniquely identified point (i.e., \( H^L_{t+1}(c) = H^U_{t+1}(c) \)) or a proper interval (i.e., \( H^L_{t+1}(c) < H^U_{t+1}(c) \)). Note that although we use time subscripts for the bounds, this simply indicates the within-period distributions \( (G_t \text{ and } G_{t+1}) \) used to obtain the bounds; the \( H \) distribution itself is time invariant.
It is also apparent that we have only exploited information about $G_t$ in two adjacent periods ($t$ and $t+1$) so far. The following theorem shows that this is not enough information to uniquely identify $H(c)$.

**Lemma 2.** Let $c \in [0, \infty)$. If $G_t(c) \in (0, 1)$, then $H_{t+1}^L(c) < H(c) < H_{t+1}^U(c)$ for all $t$ (i.e., the upper and lower bounds are not equal).

**Proof.** Suppose to the contrary, that $H_{t+1}^L(c) = H_{t+1}^U(c)$ for some $t$. Using the definitions from (10) yields an expression for $G_{t+1}(c)$ as a function of $G_t(c)$:

$$G_{t+1}(c) = G_t(c)^2/[2G_t(c) - 1].$$  \hfill (11)

Consider three cases: $G_t(c) < 0.5$, $G_t(c) = 0.5$, and $G_t(c) > 0.5$. In the first case $2G_t(c) < 1$ and (11) imply that $G_{t+1}(c) < 0$, which is a contradiction since $G_t$ is a CDF. In the second case, imposing $H_{t+1}^L(c) = H_{t+1}^U(c)$ in (10) implies that $G_{t+1}(c) - 0.5 = G_{t+1}(c)$, which is impossible. In the third case, we have $G_t(c) < 0.5$ and $1 < 2G_t(c) < 2$ which, in light of (11), implies that $G_{t+1}(c) > 1$, which is again a contradiction since $G_t$ is a CDF. Hence, it must be the case that $H_{t+1}^L(c) < H_{t+1}^U(c)$. \hfill \(\blacksquare\)

This result implies that potentially multiple distributions might satisfy the estimated bounds on the entry distribution, but this is not particularly surprising since we have not used all of the information at hand and given the weak assumptions we make on the dynamic process. We do not impose anything other than boundedness on the consumers’ probability of purchase and are completely agnostic on the consumer’s actual dynamic decision making process. Even though the bounds are proper intervals, as will be shown next, they are still informative (i.e., they are different from the trivial bounds [0, 1]).

**Lemma 3.** Let $c \in [0, \infty)$ be such that $G_t(c) \in (0, 1)$.

a. If $G_t(c) \neq G_{t+1}(c)$, then either $H_{t+1}^L(c) > 0$ or $H_{t+1}^U(c) < 1$ (i.e., one of the bounds will always be informative).

b. If $G_t(c) = G_{t+1}(c)$, then $H_{t+1}^L(c) = 0$ and $H_{t+1}^U(c) = 1$ (i.e., the bounds are trivial).

**Proof.** The lower bound, $[G_{t+1}(c) - G_t(c)]/[1 - G_t(c)]$, is informative when $G_t(c) < G_{t+1}(c)$. The upper bound, $G_{t+1}(c)/G_t(c)$, is informative when $G_{t+1}(c) < G_t(c)$. These conditions are mutually exclusive, therefore there will always be one informative bound for $H(c)$. \hfill \(\blacksquare\)

As a practical matter, when the within-period estimation is performed the cost values $c_{t,k}$ at which the period search cost distributions $G_t$ are estimated will be different across periods due to the natural variation in the observed prices over time. For example, consider
the bounds in (10) and suppose we are interested in bounding the value \( H(c_{t+1,k}) \), where \( c_{t+1,k} \) is a cost cutoff point for which we can identify \( G_{t+1}(c_{t+1,k}) \). Most likely we cannot also identify \( G_t \) at the same point \( c_{t+1,k} \). Nevertheless, we can still estimate conservative bounds by exploiting the monotonicity of \( G_t \) to evaluate it instead at nearby points. In particular, since the values \( G_t(c_{t,k'}) \) are identified we will use points \( c_{t,k'} \) for a suitable choice of \( k' \).

More specifically, for the upper bound the value \( G_t(c_{t+1,k}) \) appears in the denominator and so we can obtain a conservative upper bound by choosing \( c_{t,k'} \leq c_{t+1,k} \), which ensures that \( G_t(c_{t,k'}) \leq G_t(c_{t+1,k}) \) and hence \( H^U_t(c_{t+1,k}) = G_{t+1}(c_{t+1,k})/G_t(c_{t+1,k}) \leq G_{t+1}(c_{t+1,k})/G_t(c_{t,k'}) \). To get the best feasible bound, we should choose the maximal \( c_{t,k'} \) such that \( c_{t,k'} < c_{t+1,k} \). Hence, a conservative, feasible upper bound is

\[
\hat{H}^U_{t+1}(c_{t+1,k}) = \frac{G_{t+1}(c_{t+1,k})}{\min_{c \in \{c_{t,k'}: c_{t,k'} < c_{t+1,k}, k'=1,...,K\}} G_t(c)} \geq H^U_{t+1}(c_{t+1,k}) \geq H(c_{t+1,k}).
\]

We can construct a conservative, feasible lower bound \( \hat{H}^L_{t+1}(c_{t+1,k}) \) in a similar fashion.

Finally, as we alluded to previously, we can also refine the bounds by exploiting information across periods and imposing monotonicity across cost values. Calculating the conservative bounds on \( H(c_{t,k}) \) at all estimated cutoff costs \( c_{t,k} \) for \( t = 1, \ldots, T - 1 \) and \( k = 1, \ldots, K \), as described above, generates \( K(T-1) \) total upper and lower bounds, respectively. Yet, the only bounds that constitute meaningful restrictions on the population distribution are the innermost bounds: the smallest upper bound and the largest lower bound at a given cost value. When the most restrictive bound at a given cost is satisfied then all the other bounds will necessarily be satisfied. The population CDF is monotonic so the bounds of the population distribution should also be monotonic. Imposing monotonicity (a restriction across cost values) of the innermost bounds (which incorporate information across time) yields the following feasible, estimable bounds:

\[
\hat{H}^L(c) \equiv \max_{t \in \{2, \ldots, T\}} \max_{\hat{c} \in \{c_{t,k}: c_{t,k} \leq c, k=1,\ldots,K\}} \hat{H}^L_t(\hat{c})
\]

and

\[
\hat{H}^U(c) \equiv \min_{t \in \{2, \ldots, T\}} \min_{\hat{c} \in \{c_{t,k}: c_{t,k} \geq c, k=1,\ldots,K\}} \hat{H}^U_t(\hat{c}).
\]

These are the bounds that we estimate and report in the application described in Section 6. They are essentially envelopes of the pooled upper and lower bound values for all cost values of \( c_{t,k} \) for \( t = 2, \ldots, T \) and \( k = 1, \ldots, K \). Since these innermost bounds are aggregated across all periods and cost values, as the number of periods (observations) increases we will have informative upper and lower bounds at additional values in the support of the population distribution.
We note that although the bounds $H^L_{t+1}$ and $H^U_{t+1}$ do not meet (by Lemma 2), because we aggregate information across periods the envelope bounds $\hat{H}^L$ and $\hat{H}^U$ may meet or even cross. In other words, $H$ may be nonparametrically overidentified if there is sufficient variation in $G_t$ across periods. Indeed, in our application the estimated bounds sometimes cross. As in the case of generalized method of moments (GMM) estimation of finite-dimensional parameters (Hansen, 1982), this suggests that we should choose our estimate of $H$ to minimize some loss function. Given the two inequality constraints $H^L(c) \leq H(c) \leq H^U(c)$, we consider the directional $L^2$ loss function

$$L(H; H^L, H^U) = \int_0^\infty \left[ \left| H^L_t(c) - H(c) \right|^2_+ + \left| H^U_t(c) - H(c) \right|^2_- \right] dc,$$

where for any $z \in \mathbb{R}$ we define $|z|_+ = |z|1\{z < 0\}$ and $|z|_+ = |z|1\{z > 0\}$.

For given bounds $H^L$ and $H^U$, the average $\hat{H}(c) = (H^L(c) + H^U(c))/2$ will minimize the directional loss function. Therefore, in our application in addition to the estimated bounds we also report

$$\hat{H} = \frac{\hat{H}^L + \hat{H}^U}{2}$$

as a nonparametric estimate of $H$.

### 3.2. Improved Bounds with Monotonic Policies

Finally, we consider whether we can strengthen the bounds by imposing weak monotonicity of the policy functions $\sigma_t$. Note that without this assumption, as established in Lemma 1 above, the maximizing policy function in $\Sigma$ is weakly decreasing in $c$ meaning that consumers with higher search costs are more likely to purchase than those with lower search costs. Suppose instead that we restrict our analysis to the space of weakly increasing functions $\Sigma_M \subset \Sigma$.

**Lemma 4.** When $\sigma_t \in \Sigma_M$, the conditional, non-purchase probabilities that generate the upper bound on $G_{t+1}$ for a given cost $c'$ are

$$
\sigma^U_t(c; c') = \begin{cases} 
1, & G_t(c') > H(c'), \\
0, & G_t(c') \leq H(c'). 
\end{cases}
$$

The implied lower bound on $H$ is

$$H^L_{t+1}(c) = \begin{cases} 
\frac{G_t(c)}{G_{t+1}(c) - G_t(c)} & \text{if } G_t(c) < G_{t+1}(c), \\
\frac{1 - G_t(c)}{1 - G_{t+1}(c)} & \text{otherwise}. 
\end{cases}
$$

The upper bound given in (10) still applies.
Proof. If a weakly increasing policy function \( \sigma_t \) places non-zero probability on values of \( c \) with \( c \leq c' \), values for which \( L_{t,2} \) is positive (where \( L_{t,j} \) is defined in the proof of Lemma 1), then it necessarily places non-zero probability (at least as large) on values of \( c \) with \( c > c' \). Since \( L_{t,2} \) is negative for \( c > c' \), this strictly decreases \( G_{t+1}(c') \) relative to a decreasing policy placing zero probability on values \( c > c' \). Hence, requiring monotonicity introduces a trade-off across values of \( c \).

For maximizing \( G_{t+1}(c') \), assigning non-zero probability on \( c \leq c' \) (corresponding to the positive part of the functional derivative) must counter-balance the negative effects of necessarily assigning non-zero probability on \( c > c' \) (corresponding to the negative part of the functional derivative). Otherwise, it would be optimal to choose \( \sigma_t = 0 \). It is never optimal to increase the probability on the region \( c > c' \), so any maximizing function must be constant on \( c > c' \) and equal to the value at \( c' \). Let \( \alpha \geq 0 \) denote the constant value on \([c', \infty)\). Now, of policies that assign probability \( \alpha \geq 0 \) at \( c' \), those yielding the highest values of \( G_{t+1}(c') \) are also constant and equal to \( \alpha \) on \([0, c')\). Hence, the maximizing policy \( \sigma_t \) must be constant on \([0, \infty)\). Finally, the constant value of the maximizing function \( \alpha \) must be either 0 or 1 depending on whether the integral of \( \sigma_t(c)L_{t,2}(c, \sigma_t(c), \sigma'_t(c)) \) is negative or positive, respectively. For \( \sigma_t = 1 \), \( G_{t+1}(c') = G_t(c') \) and for \( \sigma_t = 0 \), \( G_{t+1}(c') = H(c') \). Hence, the maximizing weakly monotonic policy function is

\[
\sigma_t^U(c; c') = \begin{cases} 
1, & G_t(c') > H(c'), \\
0, & G_t(c') \leq H(c'). 
\end{cases}
\]

This yields the following upper bound on \( G_{t+1}(c') \):

\[
G_{t+1}(c') \leq G_t^U(c') = \max \{ G_t(c'), H(c') \}.
\]

In other words, if \( G_t(c') \leq H(c') \) then \( G_{t+1}(c') \leq H(c') \) and if \( G_t(c') \geq H(c') \) then \( G_{t+1}(c) \leq G_t(c') \). We can obtain a conditional lower bound on \( H(c) \) by taking the contrapositive of the second statement (and replacing \( c' \) with \( c \)): if \( G_t(c) < G_{t+1}(c) \) then \( G_t(c) < H(c) \).

Obtaining a feasible lower bound requires evaluating \( G_{t+1} \) at a period \( t + 1 \) cutoff, \( c_{t+1,k} \), evaluating the first instance of \( G_t \) at a larger value \( c_{t,k+} > c_{t+1,k} \) for which the first inequality still holds, and evaluating the second instance of \( G_t \) at a smaller value \( c_{t,k-} < c_{t+1,k} \) so that the bound is conservative. Hence, if \( G_t(c_{t,k+}) < G_{t+1}(c_{t+1,k}) \) then \( G_t(c_{t,k-}) < H(c_{t+1,k}) \).
Table 1: Representative Simulation Specifications

| Specification          | Initial Distribution ($G_0$) | Population Distribution ($H$) | Purchase Prob. ($1 - \sigma$) |
|------------------------|------------------------------|--------------------------------|-------------------------------|
| **Specification 1: Positive Bias** | $G_0 = \text{U}(0, 25)$     | $H = \text{Exponential}(4)$          | $1 - \sigma = \text{Gamma}(10, 0.5)$ |
| **Specification 2: Positive Bias with $G_0$ Similar to $H$** | $G_0 = \text{Exponential}(4.1)$ | $H = \text{Exponential}(4)$          | $1 - \sigma = \text{U}(0, 25)$ |
| **Specification 3: Both Positive and Negative Bias** | $G_0 = \text{Gamma}(25, 0.2)$ | $H = \text{Lognormal}(1.5, 1)$                      | $1 - \sigma = \text{Gamma}(20, 0.25) \text{ PDF}$ |
| **Specification 4: No Bias** | $G_0 = \text{U}(0, 25)$     | $H = \text{Exponential}(4)$          | $1 - \sigma = 0.5$ |


4. Dynamic Selection Effects: Simulation Evidence and Theoretical Results

This section investigates the effects of dynamic selection on the period search cost distribution by way of a small-scale simulation study and contains additional theoretical results which characterize the extent of the problem. We are particularly interested in how the characteristics of the consumers’ purchase policy functions are related to persistence of dynamic selection bias in terms of the difference between the search cost distribution obtained by a static analysis (the per-period distributions $G_t$) and the population search cost distribution ($H$).

The simulation study is structured as follows. For each simulation, we choose a starting period search cost distribution, $G_0$, a population distribution, $H$, and a time-invariant purchase probability policy function $(1 - \sigma(\cdot))$. The simulation begins with 1,000 consumers with costs drawn randomly from the starting distribution $G_0$. We then use the policy function $\sigma$ to simulate which of the 1,000 consumers purchase and which remain in the market. Those who stay in the market retain their search costs while those who purchase are replaced with consumers from the population distribution $H$. We repeat this process for 10,000 periods.

To examine the effects of the functional form of the consumer’s purchase probability on next period’s distribution we used an assortment of bounded functions that are monotonically increasing, monotonically decreasing, non-monotonic, and independent of search cost. Table 1 displays the primitives of four representative specifications which we will use to illustrate our findings. We considered all combinations of four initial distributions, three population distributions, and 20 policy functions for a total of 240 distinct simulations of 1,000 consumers over 10,000 periods.\(^9\)

The results suggest that when consumers’ purchase probabilities are monotonically decreasing in cost, selection effects cause the quantiles of the period search cost distributions to be biased upwards when compared to the population distribution, regardless of the starting or population distribution used. Specification 1, illustrated in Figure 1a, is a representative example where positive bias (at all reported quantiles) in the search cost distribution arises. The starting search cost distribution is $U(0,25)$, the population distribution is Exponential(4), and $(1 - \sigma(\cdot))$ is one minus the CDF of Gamma(10,0.5). This result is not surprising since in this scenario, consumers with larger costs are less likely to purchase each period. Thus over time the population of consumers will become more concentrated around larger costs. This result suggests that if the probability of purchasing is decreasing in search cost, then markets for durable goods should experience weakly more price dispersion over time as the remaining consumers will tend to search.

\(^9\)The appendix contains details on the full collection of distributions and policy functions used.
less intensely before purchasing.

The opposite occurs in the (counter-intuitive) situation where the consumer’s purchase probability is monotonically increasing in search cost. Regardless of the starting or population search cost distributions used, selection effects cause a downward bias relative to the population distribution.

A question raised by Specification 1 is, to what extent is the bias driven by the initial difference between $G_0$ and $H$ as opposed to dynamic selection effects? Specification 2, illustrated in Figure 1b, shows that even when the initial distribution $G_0$ is very similar to the population distribution $H$, we can see bias at all quantiles in the distribution that continues to increase over time. The starting search cost distribution in this case is Exponential(4.1), which is very close to the population distribution, Exponential(4), and $(1 - \sigma(\cdot))$ is one minus the CDF of $U(0, 25)$. More generally, we find that the rate of growth in the bias is affected by the difference between the two distributions, but not the existence of the bias.\(^{10}\)

When the probability of purchase is non-monotonic, as in Specification 3, the bias can be positive, negative, or indeterminant. As shown in Figure 1c, in some cases different quantiles might be biased in different directions. If bias does exists, then its direction appears to depend on the nature of the non-monotonicity of the purchase probability. A sizable proportion of consumers in one of the tails of the distribution who systematically choose not to purchase can shift next period’s search cost distribution in their direction. The results from the simulations with non-monotonic purchase probabilities are less similar than the monotonic case, due to the mixing that is occurring for the middle cost consumers, but the qualitative results from the monotonic cases seem to still hold. That is, regardless of the starting and population cost distributions, dynamic selection effects result in bias in the per-period search cost distribution as compared to the population distribution. The direction of the bias in the median and higher quantiles is determined by which tail of the distribution the purchase probability function puts relatively more weight on. However, when the purchase policy puts similar amounts of mass in both tails then: (a) if both high and low search cost consumers are purchasing, the bias is weakly positive, regardless of the starting and population distribution and (b) if only the middle search cost consumers are purchasing, then the bias depends on the difference between the initial and population distributions.

Furthermore, if the purchase probability is a constant then there is a closed-form

\(^{10}\)To examine this question more closely, we ran an additional 360 simulations using the same 20 policy functions as before but in cases where the initial distribution was very similar to or equal to the population distribution (e.g., by perturbing a location parameter).
Figure 1: Decile Bias of Period Search Cost Distributions ($G_t$)
solution for period \( t \)'s search cost distribution:

\[
G_t(c) = \sigma G_{t-1}(c) + (1 - \sigma) H(c).
\]

Extending this process back to the initial period reveals that period \( t \)'s search cost distribution is a simple weighted average of the initial distribution and the population distribution:

\[
G_t(c) = \sigma^t G_0(c) + (1 - \sigma) H(c) \sum_{s=1}^{t} \sigma^{s-1}.
\]

Taking the limit of this period's search cost distribution reveals that a constant purchase probability which is independent of cost causes the period search cost distribution to converge towards the population distribution. That is, \( \lim_{t \to \infty} G_t(c) = H(c) \) for all \( c \) and so there is no persistent bias. The rate of convergence depends on the consumers' purchase probability, with faster convergence as \( \sigma \) gets smaller.

The simulation results for Specification 4 are consistent with this analysis. Figure 1d shows an example of this convergence for the case where the starting distribution is \( U(0, 25) \), the population distribution is \( \text{Exponential}(4) \), and the purchase probability is constant at \( 1 - \sigma = 0.5 \).

The presence of bias can be further characterized as the following theorems describe. Let the unconditional probability of not purchasing in period \( t \) for a given policy function \( \sigma_t \) be \( R_t \equiv \int_{0}^{\infty} \sigma_t(c) g_t(c) \, dc \). Note that this quantity is constant within a period and is independent of the search cost \( c \). Lemma 5 shows that for a family of policy functions, if \( G_t \) is biased for \( H \) in the current period, then \( G_{t+1} \) will be biased for \( H \) in the next period as well.

**Lemma 5.** If \( G_t \neq H \), then \( G_{t+1} \neq H \) for any policy function \( \sigma_t \) such that \( R_t > \inf_{c'} \frac{g_t(c')}{h(c')} \).

**Proof.** Suppose to the contrary that there exists a policy \( \sigma_t \) such that \( R_t > \inf_{c'} \frac{g_t(c')}{h(c')} \) and that \( G_{t+1} = H \) (i.e., there is no bias). From (9), for all \( c' \)

\[
H(c') \int_{0}^{\infty} \sigma_t(c) g_t(c) \, dc - \int_{0}^{c'} \sigma_t(c) g_t(c) \, dc = 0.
\]

Differentiating with respect to \( c' \) and rearranging yields

\[
\sigma_t(c') = \frac{h(c')}{g_t(c')} R_t.
\]

But \( \sigma_t \) is a probability-valued function and is bounded: \( \sigma_t(c') \leq 1 \) for all \( c' \). This implies that \( R_t \leq \frac{g_t(c')}{h(c')} \) for all \( c' \), which is a contradiction.
The following corollary shows that if the within-period search cost distribution has reached an unbiased steady state (i.e., $G_t = G_{t+1} = H$ for some $t$), then the policy function ($\sigma_t(\cdot)$) must be constant and equal to the unconditional probability of not purchasing.

**Corollary.** If $G_t = G_{t+1} = H$, then the policy function $\sigma_t$ is constant with $\sigma_t(c) = R_t$ for all $c$.

**Proof.** If $G_t = G_{t+1} = H$, then $g_t = g_{t+1} = h$ and it follows from (12) that $\sigma_t(c') = R_t$ for all $c'$.

### 5. Estimation

In this section, we propose a method to estimate the model described above using only data on prices over time. In particular, suppose we observe $N_t$ prices in periods $t = 1, \ldots, T$ and without loss of generality suppose that the prices are ordered from smallest to largest in each period (i.e., $p_{t,1} < \ldots < p_{t,N_t}$ for all $t$). The estimation method consists of two or three stages: 1) nonparametric estimation of the within-period search cost distributions, 2) nonparametric estimation of bounds on the entry distribution, and 3) optionally using the nonparametric bounds to estimate a parametric entry distribution.

First, we follow the MGW approach to estimate the within-period search cost distributions at the cutoff points using nonparametric maximum likelihood. We use the minimum and maximum observed prices each period as estimates of the support of the price distribution, $p_{t,1} = \underline{p}_t$ and $p_{t,N_t} = \overline{p}_t$.

The ML estimation problem in each period $t$ is:

$$\max_{\mu_t} \sum_{i=2}^{N_t-1} \log f_t(p_{i}; \mu_t)$$

subject to $(p - r) \sum_{k=1}^{K} \frac{k\mu_{t,k}}{K} (1 - F_t(p))^{k-1} = \frac{\mu_{t,1}(\overline{p}_t - r)}{K}$ for all $p \in [\underline{p}_t, \overline{p}_t]$.

where $\mu_t = (\mu_{t,1}, \ldots, \mu_{t,K})$ and $f_t(p; \mu_t)$ is defined in (5).

Recall that the price distribution is strictly monotonically increasing, so the inverse of $F_t$ exists. The set of search cost cutoffs $\{c_{t,k}\}_{k=1}^{K}$ are defined by the following integral:

$$c_{t,k} = \int_{\underline{p}_t}^{\overline{p}_t} F_t^{-1}(z) ((k+1)z - 1) (1-z)^{k-1} dz \text{ for all } k = 1, 2 \ldots K - 1.$$  

Monte Carlo techniques can be used to evaluate this integral to obtain the indifference costs each period.\footnote{Moraga-González et al. (2010, Proposition 5) prove that in a symmetric equilibrium, the series of critical cutoff costs is the solution to a system of equations that, for a fixed $v_t$, $n_t$, and $G_t(c)$, are guaranteed to be numerically solvable.}

\footnote{Refer to Kiefer and Neumann (1993) and Donald and Paarsch (1993) regarding the use of order statistics to estimate bounds.}
search each number of firms \( k = 1, \ldots, K \), and the cutoff search costs \( c_{t,k} \) for each \( k \), we use (3) to obtain the values of the search cost CDF \( G_t(c_{t,k}) \) at each of cutoff search cost values.

Next, the CDF values \( G_t(c_{t,k}) \) allow us to estimate the \( 2K(T - 1) \) bounds on the population distribution \( H \). We can obtain nonparametric upper and lower bounds on \( H \) by using the innermost bounds as detailed in Section 3.

Finally, if the population distribution is a member of a parametric family of distributions, then one can use the nonparametric bounds to estimate the finite-dimensional parameters of the distribution using nonlinear least squares (NLS). That is, if \( H(\cdot) = H(\cdot; \theta) \) then we can construct a NLS criterion function based on squared directional violations of the nonparametric bounds that arise for a given value of \( \theta \):

\[
Q_T(\theta) = \sum_{t=1}^{T-1} \sum_{k=1}^{K} \left( |\hat{H}^L_t(c_{t,k}) - H(c_{t,k}; \theta)|^2 + |\hat{H}^U_t(c_{t,k}) - H(c_{t,k}; \theta)|^2 \right)
\]  

(13)

In words, for a given value of \( c \) if \( H(c; \theta) \) falls within the bounds, then the bounds are satisfied and the contribution to the criterion function \( Q_T \) is zero. On the other hand, for values of \( c \) for which \( H(c; \theta) \) violates one or both of the bounds, then the squared of the distances by which each bound is violated are added.

6. An Application to the Online Market for Econometrics Textbooks

In this section, we apply our dynamic model and estimation procedure using data from the online market for two widely-used graduate econometrics textbooks. A recent representative survey of American households finds that 87% of all adults 18 years and older access and use the internet on a regular basis (that jumps to 97% for people aged 18 to 49), Zickhur (2012). Furthermore, a recent study found that the average age of a graduate student at an American institution is 33 while 49% of students are under the age of 29, Bell (2009). We can infer from this two surveys that the actual market participants for these textbooks would actually be using the internet to conduct their search. The market for online books is a mature and stable part of the overall publishing industry; a recent study found that the book industry had the second highest penetration rate among domestic internet users (De los Santos, 2010).\(^{13}\) Furthermore, the 2013 publishing industry’s annual review found that 44% of all of Americans’ expenditures on books went to online retailers, Bowker and Publishers’ Weekly (2013). Many studies—Bailey (1998), for instance—find that price dispersion is a persistent feature for this market. Scholten and Smith (2002), and Pan, Ratchford, and Shankar (2003a,b) find that price dispersion is generally larger for books sold online than in traditional brick and mortar stores.

\(^{13}\)Refer to Clay et al. (2001) for a review of the online book industry.
Previous research on consumer search in both physical and electronic markets has reached the consensus that price search is particularly limited. Typical results from the pre-internet literature find that roughly 40-60% of consumers visit only one firm prior to purchase. More recently, De los Santos (2010) uses detailed individual level online browsing and purchasing behavior data and finds that in a quarter of all transactions the consumer only searches one firm, while the average consumer searches only 1.29 firms. For consumers searching more than one firm, almost half exercised their recall option and did not purchase from the last firm visited. Using similar data, Johnson, Moe, Fader, Bellman, and Lohse (2004) also find that the average online book shopper only searched 1.2 firms before purchasing. Further, De los Santos, Hortacsu, and Wildenbeest (2012) conclude that a simultaneous search strategy can explain observed online search patterns better than sequential search.

6.1. Data

Our dataset contains daily price observations collected from nine online retail outlets from March to November 2006 (236 days) for new hardcover editions of Advanced Econometrics by Amemiya (1985) and Microeconometrics by Cameron and Trivedi (2005). For the purposes of this analysis, a seller is classified as a verified retailer if the seller ID attached to the listing could be matched to a legitimate external physical or digital firm. The first column of Table 2 contains general summary statistics on prices across all periods for both books. The Amemiya sample contains prices for 79 unique sellers for a total of 11,475 observations. 74.9% of those listings are from verified retailers and 9.3% are from major retailers. The Cameron and Trivedi sample has 110 unique sellers and a total of 15,791 observations, with 62.8% of those listings being from verified retailers and 7.9% from major retailers. In both cases, a sizable proportion of the listings are from individual, non-commercial entities who listed books for sale on the marketplace areas of the websites considered.

Table 2 also contains the summary statistics for prices by seller type, for both books. There is less variation in prices offered by major retailers than those for the other sellers. Non-verified retailers generally offer both the highest price and nearly the lowest price in every period. As a group, they generate most of the price dispersion observed. This result is unsurprising given that this group is the most heterogeneous, being composed of non-commercial individuals and small sellers with an online presence only on a single

---

14Refer to Newman (1977) for an extensive review of the literature for search in non-electronic markets.
15The online retailers are Abebooks, Alibris, Amazon, Barnes & Noble, Half.com, Overstock, Powell’s, Super Book Deals, and Walmart.
16Powell’s, Super Book Deals, Walmart, Barnes & Noble, Amazon, and Overstock are considered major retailers for this application.
### Table 2: Price Summary Statistics

|                     | All Retailers | Non-Verified Retailers | Verified Retailers | Major Retailers |
|---------------------|---------------|------------------------|--------------------|-----------------|
| **Amemiya (1985)**  |               |                        |                    |                 |
| Minimum             | 34.99         | 35.00                  | 34.99              | 52.50           |
| Maximum             | 179.21        | 179.21                 | 166.88             | 79.00           |
| Mean                | 84.00         | 93.17                  | 80.93              | 68.84           |
| Median              | 83.50         | 84.62                  | 83.35              | 71.23           |
| St. Dev.            | 19.96         | 30.41                  | 13.59              | 6.60            |
| Observations        | 11,457        | 2,874                  | 8,583              | 1,018           |
| **Cameron and Trivedi (2005)** |       |                        |                    |                 |
| Minimum             | 53.01         | 53.01                  | 56.25              | 56.25           |
| Maximum             | 193.15        | 193.15                 | 160.26             | 82.25           |
| Mean                | 86.26         | 91.65                  | 82.99              | 70.17           |
| Median              | 82.00         | 81.95                  | 82.44              | 82.44           |
| St. Dev.            | 22.72         | 30.91                  | 14.93              | 4.46            |
| Observations        | 15,791        | 5,951                  | 9,841              | 1,163           |

website. Given that a rather sizable proportion of the listings in our sample are from non-verified retailers, a group that is generally characterized by high price dispersion, the results of Scholten and Smith (2002) and others are not surprising in this light. The standard deviation for prices from non-verified retailers is around $30 for both books, while it is less than $7 for the major retailers. We can see that the observed online price dispersion is mostly being driven by the presence of these non-major retailers.

Figure 2 depicts the stability in the minimum and maximum prices for both books over the sample. For the Amemiya data, the maximum price per period is within $10 of the overall maximum 87.7% of the time, has a standard deviation of $18.64, and only once in the sample is it offered by a verified retailer. The minimum price per period is within $10 of the overall minimum 33.9% of the time, but 91.1% of the period minimums are within $20. There is more persistence in the minimums, they have a standard deviation of $7.85. Verified retailers offer the minimum price for 94.9% of the periods.

For the Cameron and Trivedi data, maximum per period price is within $10 of the overall maximum 70.3% of the time, has a standard deviation of $23.53, and is offered by a verified retailer 6.7% of the time. The minimum price per period is within $10 of the overall minimum 71.6% of the time, while all the prices are within $20. The minimum
price per period has a standard deviation of $3.54, and is offered by a verified retailer 69.4% of the periods. For both books, the minimum per period price demonstrates less movement on average, period to period, than the per period maximum. The per period minimum is concentrated around the overall minimum; the difference between the largest and smallest per period minimum is less $24 for both books. Conversely, the per period maximum has a relative large range, more than $40 for both books. This is most likely due to the composition of the sellers offering the different extreme prices. As mentioned above, a vast majority of the sellers offering the maximum observed price are non-commercial sellers with no other presence except on that one listing website. On the other hand, a large proportion of the sellers offering the minimum are verified commercial sellers.
As is clearly visible from Figure 2, the market for both books are characterized by persistent and significant price dispersion. The central quantiles seem to be relatively stable over the summer months. However, the medians are significantly more volatile in the autumn for both books, with Cameron and Trivedi having a similar movement in the spring. This movement seems to be driven by changes in the composition of sellers. Over these time periods the proportion of listings from non-verified retailers increases, the proportion of listings from major retailers increases, the number of observations per period decreases, and there is an increase in individual, non-commercial first time sellers starting to sell at lower prices. Since there are less observations in these periods, each listings has a larger effect on the quantiles. The major retailers as a group offer the books at consistently lower prices than the other groups on average, but at the same time the non-verified retailers have significant variation in their prices. An increase in both of these groups causes the prices to be more polarized.

6.2. Results

The first stage of our procedure involves estimating the period search cost distributions, \( G_t \), for each \( t \), at the search cost cutoffs, \( \{c_{t,k}\}_{k=1}^{K} \). We set the maximum number of firms that a consumer can search at \( K = 15 \). This choice was made because over the course of our sample, the online marketplaces average roughly 12 listings per day coupled with the fact that Johnson et al. (2004) and De los Santos (2010) find that the average online book shopper searches 1.2 and 1.29 online retailers, respectively (\( 12 \times 1.2 = 14.4 \) and \( 12 \times 1.29 = 15.48 \)). With \( K = 15 \), becoming a fully informed consumer (searching all \( K \) firms) is generally searching a marketplace plus most of the major retailers. The choice of the maximum number of firms searched is consistent with findings of the current behavioral research on online shopping. Additionally, Hong and Shum (2006) and Moraga-González and Wildenbeest (2008) both show that to consistently estimate the period search cost distributions requires at least as many moment conditions—generated by observed prices—as the number of firms consumers can search. For some periods in the Amemiya sample we only have 19 observations, which restricts our choice to \( K \leq 19 \). Furthermore, Moraga-González and Wildenbeest (2008) perform simulations and find that misspecifying the number of firms in the market by less than 20% only has minor effects and does not qualitatively affect the shape of the search cost distribution. The estimation was also carried out using \( K = 10 \) and \( K = 19 \) and the results were qualitatively similar to the results presented here with \( K = 15 \).\(^{17}\)

Traditionally the literature on search distinguishes between two types of consumers:

\(^{17}\)Results from these robustness tests are available from the authors upon request.
fully informed and fully uninformed consumers (see, e.g., Widle and Schwartz (1979), Varian (1980), and Stahl (1989)). This emphasis is understandable since it simplifies the computation while still being a reasonable proxy for reality. Consumer heterogeneity is still driving the predictions that search frictions result in price dispersion. For comparability, we can examine the evolution in the proportion of both of these extreme information consumers over time using the first stage results. In our notation, the proportion of uninformed consumers in period $t$ is $\mu_{t,1}$ and the proportion of fully informed consumers is $\mu_{t,15}$. Hence, the proportion of partially informed consumers is $\sum_{k=2}^{14} \mu_{t,k}$.

Starting with Amemiya, Figure 3a clearly shows a large amount of daily fluctuation in the proportion of uninformed consumers. However, out of this volatility a general upward
Table 3: Estimated Search Cost Quantiles, Static Model

| Specification    | Amemiya | Cameron and Trivedi |
|------------------|---------|---------------------|
| Mean across periods | 1.20    | 4.45                |
|                   | (0.16)  | (0.14)              |
| Median across periods | 0.32    | 5.57                |
|                   | (0.01)  | (0.15)              |

Note: Bootstrap standard errors are reported in parentheses using 1000 replications.

The trend is still present according to a Mann-Kendall (MK) test for a nonparametric trend (Mann, 1945; Kendall, 1975). The p-value for the MK test with the null hypothesis of no trend is approximately equal to zero (smaller than $10^{-16}$). Hence, the share of consumers who choose to not search before purchasing appears to be increasing.

Figure 3a also shows that the proportion of fully informed consumers has a statistically significant downward trend (MK test, $p \approx 0$), meaning a decreasing share of consumers search all firms before purchasing. However, it should be noted that there is a sharp increase in the first half of June before it falls back to around its pre-June level. Data covering multiple years would be necessary to determine if this drop is a seasonal effect, but given the timing it is reasonable to conclude that this is driven by a change in demand due to school effects. June is the traditional end of the academic year for American universities, which is also the time when students tend to sell textbooks that they no longer need. The pressure from these used books, which consumers could consider as a close substitute for the new editions, might induce consumers to have more incentive to search during these periods. Given that the dramatic changes in the composition of the market coincides with the ending of the summer term, school effects seem reasonable.

Similar to the previous case, there is a significant amount of movement over time in the Cameron and Trivedi data, as shown in Figure 3b. There is a structural break in the level of uninformed consumers around late May where it decreases sharply. However dividing the sample had this break point, both subsamples possess an upward trend (MK test, $p \approx 0$ both before and after). On average, the share of the market that is not searching before purchasing is increasing over time. The proportion of informed consumers has an upward trend over the whole sample (MK test, $p = 0.0025$). However, the proportion of informed consumers also has a structure break where it increases dramatically at the same period as the proportion of uninformed consumers. Before this break there is a significant time trend (MK test, $p = 0.0189$), but following the break, there is no trend (MK test, $p = 0.3422$).

From the first stage, the estimated per-period search cost distributions are also of
interest. Table 3 summarizes the estimated quantiles of these distributions over our sample. For each of the 236 days, we estimate the within-period distribution of search costs the 25th, 50th, and 75th percentiles of the distribution. We report the mean and median values of each quantile across periods. For Amemiya, we see that the median search cost is about $9 on average. For Cameron and Trivedi, it is about $8 on average. However, these estimates do not account for entry and exit of consumers from the population of consumers who are active in the market. Our simulations and estimated bounds provide evidence that these selection effects may cause the period distribution to systematically shift away from the population distribution.

The second stage of the estimation procedure uses the period search cost distributions to calculate the bounds on $H$ at the estimated search cost cutoffs. Figure 4 shows the estimated bounds on the population distribution for Amemiya (1985) and Cameron and Trivedi (2005), respectively, along with pointwise 95% confidence intervals for the functional values. Displayed for each book are the estimated bounds $\hat{H}^L$ and $\hat{H}^U$ for the case where $\sigma_t$ is unrestricted and the case where it is assumed to be monotonic. The figures graphically illustrate the uncertainty about the estimated bounds on the population distribution at the higher quantiles. As discussed in Section 3, since the bounds occasionally cross there is some evidence of possible nonparametric overidentification of $H$ and so we also report estimates below using the average bound $\hat{H}$ as a nonparametric “point” estimate of the function $H$ which minimizes the directional $L^2$ loss function.

The nonparametric estimates $\hat{H}^L$, $\hat{H}^U$, and $\hat{H}$ can be used to directly compare bounds on quantiles of the population distribution with quantiles estimated using a static approach. For both books, the estimated medians of the per-period distributions are typically around $8-$10. The first two rows of Table 4 contain estimates of three quantiles, including the median, from the non-parametric upper and lower bounds for each book. The third row corresponds to the average of the bounds. Based on these bounds, the medians of the population distributions for both books are smaller than for the per-period distributions with the exception of the nonparametric lower bound for Amemiya. As can be seen in Figure 4a, as a consequence of the weak assumptions we work under there is not much variation in the lower bound over the $1-$16 range. In the lower panel of Table 4, we report estimates of $\hat{H}^L$, $\hat{H}^U$, and $\hat{H}$ which exploit monotonicity of $\sigma_t$. For Amemiya in particular, this results in a higher estimated lower bound and a substantially lower estimated median.

We also compare the estimated nonparametric quantile bounds with the quantiles obtained by estimating a parametric model where $H(\cdot) = H(\cdot; \theta)$ is the CDF of the exponential distribution with mean $\theta$ and variance $\theta^2$. In this case, we estimate $\theta$ using the NLS approach outlined in Section 5. Without monotonicity, we find that $\hat{\theta}_A = 4.13 (0.27)$ for Amemiya and $\hat{\theta}_{CT} = 5.08 (0.28)$ for Cameron and Trivedi. After imposing monotonicity
Figure 4: Estimated Upper and Lower Bounds for $H$ with 95% Confidence Bands

(a) Amemiya (1985), Unrestricted $\sigma$

(b) Amemiya (1985), Monotonic $\sigma$

(c) Cameron and Trivedi (2005), Unrestricted $\sigma$

(d) Cameron and Trivedi (2005), Monotonic $\sigma$
of \( \sigma_t \) the estimated means are lower, with \( \hat{\theta}_A = 1.98 \) (0.18) and \( \hat{\theta}_{CT} = 3.00 \) (0.24). (In both cases we report bootstrap standard errors in parentheses calculated using 1000 replications.) Figure 5 plots the estimated CDFs of the parametric population distributions for both books. As shown in the last row of each panel in Table 4, these estimated parameters yield medians for \( H \) of $2.86 ($0.19) for Amemiya and $3.52 ($0.29) for Cameron and Trivedi. Even without using monotonicity, both of these medians are significantly lower than the typical estimated median of the period search cost distributions. After using monotonicity, the estimated medians drop to $1.37 ($0.13) and $2.08 ($0.17), respectively.

Both the nonparametric and parametric results suggest that in this market it is important to account for dynamic selection effects. That is, we should distinguish between the distribution of search costs of the current consumers over time and the distribution of search costs of new entrants. The distributions of search costs of consumers who are active in the market do appear to differ substantively from the population distribution of search costs of newly entering consumers.

Using the simulation results as a guide we can infer some of the characteristics of the consumers’ purchase probability policy function, \((1 - \sigma_t)\). The purchase probabilities are likely not monotonically decreasing in cost as this would cause the period medians to be biased downward (as opposed to upward) relative to the population median. It also does not appear that the purchase probability policy is constant (i.e., independent of cost), since in this case we would expect to see the period medians converging to the median of the population distribution. Hence, there is evidence that consumers’ purchase policies for these markets place relatively more probability weight in the right tail than the left tail, which leads to the observed persistent upward bias in the period search cost distribution.
Table 4: Estimated Search Cost Quantiles, Dynamic Model

| Specification                      | Amemiya 0.25 | 0.50 | 0.75 | Cameron and Trivedi 0.25 | 0.50 | 0.75 |
|-----------------------------------|--------------|------|------|-------------------------|------|------|
| **Unrestricted σ**                |              |      |      |                         |      |      |
| Nonparametric $H$, Lower Bound    | 0.20         | 16.68| –    | 0.09                    | 6.58 | –    |
|                                  | (0.01)       | (0.02)| –    | (0.02)                  | (0.02)| –    |
| Nonparametric $H$, Upper Bound    | 0.88         | 5.50 | 6.81 | 2.95                    | 4.95 | 5.59 |
|                                  | (0.89)       | (0.53)| (3.08)| (0.39)                  | (0.38)| (1.56)|
| Nonparametric $H$, Average        | 0.20         | 6.48 | 18.70| 0.13                    | 5.46 | 9.84 |
|                                  | (0.01)       | (2.00)| (0.57)| (0.02)                  | (1.21)| (1.38)|
| Parametric $H$, Exponential       | 1.19         | 2.86 | 5.72 | 1.46                    | 3.52 | 7.04 |
|                                  | (0.08)       | (0.19)| (0.37)| (0.08)                  | (0.19)| (0.38)|
| **Monotonic σ**                   |              |      |      |                         |      |      |
| Nonparametric $H$, Lower Bound    | 0.20         | 0.23 | 11.90| 0.09                    | 6.00 | 10.75|
|                                  | (0.01)       | (0.01)| (0.08)| (0.02)                  | (0.02)| (0.17)|
| Nonparametric $H$, Upper Bound    | 0.88         | 5.50 | 6.81 | 2.95                    | 4.95 | 5.59 |
|                                  | (0.89)       | (0.53)| (3.08)| (0.39)                  | (0.38)| (1.56)|
| Nonparametric $H$, Average        | 0.20         | 1.01 | 10.61| 0.13                    | 5.19 | 10.51|
|                                  | (0.01)       | (2.01)| (0.57)| (0.02)                  | (1.21)| (1.54)|
| Parametric $H$, Exponential       | 0.57         | 1.37 | 2.75 | 0.86                    | 2.08 | 4.17 |
|                                  | (0.05)       | (0.13)| (0.25)| (0.07)                  | (0.17)| (0.33)|

*Note: Bootstrap standard errors are reported in parentheses using 1000 replications. In cases indicated by "–" the 0.75 quantile was higher than the largest estimated search cost cutoff and could not be determined.*

7. Conclusion

As illustrated by Armstrong (2008), understanding the nature of the search technology used in a market is important for evaluating and forecasting outcomes due policy interventions designed to improve consumer welfare. The previous literature on estimating consumer search cost distributions has relied primarily on static models. However, when consumers shop for a durable good over multiple periods when search is costly, selection effects may cause search cost distributions estimated using static methods to be biased relative to the actual population search cost distribution. This bias might prevent the competition promoting program to from actually improving consumer welfare as intended.

This paper serves as the first step to extending methods to empirically estimate search cost distribution in a dynamic setting. Our simulations have demonstrated the possibility of significant distributional bias if the dynamics aspects of the search process is ignored.
Even though we use a largely model-free approach to estimate bounds on the consumer search cost distribution, our estimates take such dynamic selection effects into account. We demonstrate our method by applying it to a longitudinal dataset of online prices for two econometrics textbooks in order to highlight the potential problem, which was also explored in our theoretical results and in a simulation study.

References

Amemiya, T. (1985). *Advanced Econometrics*. Cambridge, MA: Harvard University Press.

Armstrong, M. (2008). The interactions between competition and competitive policy. *Competition Policy International* 4, 97–147.

Bayley, J. (1998). Electronic commerce: Prices and consumer issues from three products: Books, compact discs, and software. *Organization for Economics Co-Operation and Development* 98(4), 1–22.

Baye, M., D. Kovenock, and C. De Vries (1992). It takes two to tango: Equilibria in a model of sales. *Games and Economic Behavior* 4, 493–510.

Baye, M., J. Morgan, and P. Scholten (2006). Information, search, and price dispersion. In T. Hendershott (Ed.), *Handbook of Economics and Information Systems*, Volume 1, pp. 1–60. Emerald Group.

Bell, N. (2009). Research report on students in graduate education. Technical report, Council of Graduate School.

Benham, L. (1972). The effect of advertising on the price of eyeglasses. *Journal of Law and Economics* 15(2), 337–352.

Bowker and Publishers’ Weekly (2013). 2013 *U.S. Book Consumer Demographic & Buying Behaviors Annual Review*. Bowker and Publishers’ Weekly.

Burdett, K. and K. Judd (1983). Equilibrium price dispersion. *Econometrica* 51(4), 955–969.

Cameron, A. C. and P. K. Trivedi (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press.

Clay, K., R. Krishnan, and E. Wolff (2001). Prices and price dispersion on the web: Evidence from the online book industry. *The Journal of Industrial Economics* 49(4), 521–539.

Dahlby, B. and D. West (1986). Price dispersion in an automobile insurance market. *Journal of Political Economy* 94(2), 418–438.
De los Santos, B. (2010). Consumer search on the internet. Working Paper.

De los Santos, B., A. Hortaçu, and M. Wildenbeest (2012). Testing models of consumer search using data on web browsing and purchasing behavior. *American Economic Review* 102(6), 2955–2980.

De los Santos, B., A. Hortaçu, and M. Wildenbeest (2013). Search with learning. Working paper, Indiana University.

Diamond, P. (1971). A model of price adjustment. *Journal of Economic Theory* 3, 156–168.

Donald, S. and H. Paarsch (1993). Piecewise pseudo-maximum likelihood estimation in empirical models of auctions. *International Economic Review* 34(1), 121–148.

Eckstein, Z. and G. van den Berg (2007). Empirical labor search: A survey. *Journal of Econometrics* 136(2), 531–564.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica* 50, 1029–1054.

Hong, H. and M. Shum (2006). Using price distributions to estimate search costs. *RAND Journal of Economics* 37(2), 257–275.

Janssen, M. and J. L. Moraga-González (2004). Strategic pricing, consumer search and the number of firms. *The Review of Economic Studies* 71(4), 1089–1118.

Johnson, E., W. Moe, P. Fader, S. Bellman, and G. Lohse (2004). On the depth and dynamics of online search behavior. *Management Science* 50(3), 299.

Johnson, R. (2002). Search costs, lags and prices at the pump. *Review of Industrial Organization* 20(1), 33–50.

Kendall, M. G. (1975). *Rank Correlation Methods* (4 ed.). London: Griffin.

Kiefer, N. and G. Neumann (1993). Wage dispersion with homogeneity: The empirical equilibrium search model. In H. Bunzel, P. Jensen, and C. Westergård-Nielsen (Eds.), *Panel Data and Labour Market Dynamics*, Contributions to Economic Analysis, New York, pp. 57–74. North-Holland.

Mann, H. B. (1945). Nonparametric tests against trend. *Econometrica* 13, 245–259.

Manning, R. and P. Morgan (1982). Search and consumer theory. *The Review of Economic Studies* 49(2), 203–216.
McMillan, J. and M. Rothschild (1994). Search. In R. Aumann and S. Hart (Eds.), Handbook of Game Theory with Economic Applications, Volume 2, pp. 905–927. North Holland.

Moraga-González, J. L. (2006). Estimation of search costs. Working Paper.

Moraga-González, J. L., Z. Sandor, and M. Wildenbeest (2010). Nonsequential search equilibrium with search cost heterogeneity. Working Paper.

Moraga-González, J. L., Z. Sandor, and M. Wildenbeest (2012). Consumer search and prices in the automobile market. Working Paper.

Moraga-González, J. L., Z. Sandor, and M. Wildenbeest (2013). Semi-nonparametric estimation of consumer search costs. Journal of Applied Econometrics 28(7), 1205–1223.

Moraga-González, J. L. and M. Wildenbeest (2008). Maximum likelihood estimation of search costs. European Economic Review 52(5), 820–847.

Morgan, P. and R. Manning (1985). Optimal search. Econometrica 53(4), 923–944.

Mortensen, D. (1986). Job search and labor markets. In O. Ashenfelter, R. Layard, and D. Card (Eds.), Handbook of Labor Economics, Volume 2, Chapter 15, pp. 849–987. New York: North-Holland.

Newman, J. (1977). Consumer external search: Amount and determinants. In A. Woodside, J. Sheth, and P. Bennett (Eds.), Consumer and Industrial Buying Behavior, Chapter 6, pp. 79–94. New York: North-Holland.

Pan, X., B. Ratchford, and V. Shankar (2003a). Can price dispersion in online markets be explained by differences in e-tailer service quality? Journal of the Academy of Marketing Science 30(4), 443–456.

Pan, X., B. Ratchford, and V. Shankar (2003b). The evolution of price dispersion in internet retail markets. Advances in Microeconomics: Economics of the Internet and e-Commerce 12, 85–105.

Pratt, J., D. Wise, and R. Zeckhauser (1979). Price differences in almost competitive markets. Quarterly Journal of Economics 93, 189–211.

Salop, S. and J. Stiglitz (1982). The theory of sales: A simple model of equilibrium price dispersion with identical agents. American Economic Review 72(5), 1121–1130.

Scholten, P. and A. Smith (2002). Price dispersion then and now: Evidence from retail and e-tail markets. Advances in Microeconomics: Economics of the Internet and e-Commerce 11, 63–88.
Sorensen, A. (2000). Equilibrium price dispersion in retail markets for prescription drugs. *Journal of Political Economy* 108(4), 833–850.

Stahl, D. (1989). Oligopolistic pricing with sequential consumer search. *American Economic Review* 79(4), 700–712.

Stigler, G. (1961). The economics of information. *Journal of Political Economy* 69(3), 213–225.

Varian, H. (1980). A model of sales. *American Economic Review* 70(4), 651–659.

Widle, L. and A. Schwartz (1979). Equilibrium comparison shopping. *Review of Economic Studies* 46(3), 543–553.

Wildenbeest, M. (2011). An empirical model of search with vertically differentiated products. *The RAND Journal of Economics* 42(4), 729–757.

Zickhur, K. (2012). Digital differences. Technical report, Pew Research Center’s Internet & American Life Project.

### A. Simulation Study Details

The starting period distributions \((G_0)\) that were used in the simulation study were: Gamma\((25, 0.2)\), Lognormal\((1.5, 1)\), \(U(0,25)\), and an Empirical CDF estimated from last period of the Amemiya (1985) data. The entry distributions \((H)\) used were Exponential\((4.12676)\), Gamma\((25, 0.2)\), and Lognormal\((1.5, 1)\).

The non-purchase policy functions \((\sigma)\) used were \(\sigma(c) = \Pr(\text{Buy} = 0 \mid c) = \{0,0.25,0.5,0.75\}\), the CDF of \(U(0,25)\), one minus the CDF of \(U(0,25)\), the CDF of Lognormal\((2.5,1)\), one minus the CDF of Lognormal\((2.5,1)\), the CDF of \(\Gamma(10,0.5)\), one minus the CDF of \(\Gamma(10,0.5)\), the absolute value of \(\sin(c)\), the absolute value of \(\cos(c)\), an increasing step function \(\Pr(\text{Buy} = 0 \mid c) = \{0,0.5,1\}\) when \(\{c \leq 5, 5 < c < 20, 20 \leq c\}\), a decreasing step function \(\Pr(\text{Buy} = 0 \mid c) = \{1,0.5,0\}\) when \(\{c \leq 5, 5 < c < 20, 20 \leq c\}\), the PDF of \(\Gamma(20,0.25)\), the triangle distribution \(\Pr(\text{Buy} = 0 \mid c) = \{c/\beta, 2-c/\beta\}\) when \(\{c \leq \beta, c > \beta\}\), and the double triangle distribution \(\Pr(\text{Buy} = 0 \mid c) = \{1-c/\beta, c/\beta - 1\}\) when \(\{c \leq \beta, c < \beta\}\) for each \(\beta \in \{5,9,12.5\}\).

Additionally, we examined cases where the initial distribution is very similar to the population distribution. For these simulations, the starting period distributions \((G_0)\) were: Gamma\((25, 0.2)\), Lognormal\((1.5, 1)\), Exponential\((4.1)\), The entry distributions \((H)\) used were Gamma\((25, 0.2)\), Gamma\((23,0.1)\), Exponential\((4.12676)\), Exponential\((4.25)\), Lognormal\((1.5, 1)\), and Lognormal\((2,1)\). The same non-purchase policy functions \((\sigma)\) from above were used to maintain internal comparability.