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
Office of the Comptroller of the Currency

August 29, 2018†

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 price distributions to estimate the distribution from which consumer search costs are initially drawn. We show that static approaches to estimating this distribution may suffer from dynamic sample selection bias. This can happen if consumers are forward-looking and delay their purchases in a way that systematically depends on their individual search costs. We consider identification of the population search cost distribution using only price data and develop estimable nonparametric upper and lower bounds on the distribution function, as well as a nonlinear least squares estimator for parametric models. We also consider the additional identifying power of weak, theoretical 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. We illustrate this and other forms of bias in a small-scale simulation study.

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

JEL Classification: C57, C14, D83, D43.

---

*We thank Stephen Cosslett, Alessandro Gavazza, Han Hong, Matthew Lewis, Rosa Matzkin, Lindsey Novak, Marc Rysman, Matthew Shum, and Andrew Sweeting for useful comments. We are especially grateful to Matthijs Wildenbeest for feedback and for sharing his Matlab code for maximum likelihood estimation of search cost distributions. The views in this paper are those of the authors and do not reflect those of the Office of the Comptroller of the Currency or the Department of the Treasury.

†This paper is a revision of a working paper whose first draft appeared on October 14, 2014.

1. Introduction

As Benham (1972) articulates so clearly, "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 transportation." This costly information acquisition generates situations of incomplete information, which disrupts the law of one price, even in seemingly competitive markets for homogeneous products. In a seminal paper, Stigler (1961) 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 costs as a significant factor in explaining price dispersion.[1]

In this paper we propose a dynamic model with forward-looking consumers to estimate the population consumer search cost distribution. Like many other recent studies, such as Hong and Shum (2006), Moraga-González and Wildenbeest (2008) (henceforth MGW), Wildenbeest (2011), Moraga-González, Sandor, and Wildenbeest (2013), and Sanches, Silva Junior, and Srisuma (2017), we focus on the case where only the distribution of prices in each period is observed, but no individual-level data on consumers is available. We show that if consumers have the option to delay purchase until a later period, a dynamic selection problem exists that will cause estimates obtained from a static model to be biased. We illustrate this through an application focusing on online textbook markets in Section 6 as well as a small-scale simulation study in Section 4.

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 and uninformed consumers in the same market. Even for homogeneous products, firms will possess some market power and price dispersion will arise if there is at least 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. In more recent work, Janssen and Moraga-González (2004) and Moraga-González, Sandor,

[1] Refer to Pratt, Wise, and Zeckhauser (1979), Dahlby and West (1986), Clay, Krishnan, and Wolff (2001), Johnson (2002), Hortaçsu and Syverson (2004), Hong and Shum (2006), and Moraga-González, Sandor, and Wildenbeest (2012) for previous empirical studies.

[2] See McMillan and Rothschild (1994) and Baye, Morgan, and Scholten (2006) for surveys of the consumer search literature.

[3] Diamond (1971) showed that if all consumers have the same positive search cost, then all firms will charge the monopoly price, thus there will be no price dispersion.

and Wildenbeest (2017) illustrate how profoundly the shape of the search cost distribution can affect search behavior, prices, and consumer welfare, sometimes even nonlinearly.

Although several papers in the literature have data on individual-level search intensities (De los Santos, Hortaçsu, and Wildenbeest, 2012; Honka, 2014; Gautier, Moraga-Gonzalez, and Wolthoff, 2016), 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. For example, in our application we can easily observe posted prices but it would be prohibitively expensive to obtain daily data on quantities of textbooks sold by the nearly two hundred sellers in our dataset or daily information about the characteristics or habits of individual consumers in the market. As a result of such data limitations, many empirical studies on consumer search have focused on simply documenting the incidence and magnitude of price dispersion, or using 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 papers have developed methodologies to estimate the static search cost distribution for a single period using only the observed price distribution.[4] Hong and Shum (2006) and Moraga-González and Wildenbeest (2008) use the equilibrium conditions on supply and demand to consistently estimate the search cost distribution within one period at a finite number of points. Moraga-González et al. (2013) estimate the search distribution more completely on an interval by pooling multiple markets with the same search technology and exploiting the variation in valuations and marginal costs across markets. Sanches et al. (2017) use a minimum distance approach, modifying Hong and Shum (2006) to create a consistent and asymptotically normal estimator. Wildenbeest (2011) tackles the problem when homogeneous products are sold by vertically differentiated sellers, while De los Santos, Hortaçsu, and Wildenbeest (2017) estimate search costs for a differentiated product in a model with learning.

Another part of the literature, including Gowrisankaran and Rysman (2012) and Melnikov (2013), is concerned with dynamic models of demand for differentiated durable goods where consumers are forward looking and expect quality to improve over time, but there are no search frictions. Product quality is not changing in our model, but rather the composition of consumers in the market and the prices charged by firms are changing over time. Our model also accounts for consumer dynamics, but the model is for homogeneous goods where firms set prices in the presence of search frictions.

However, a limitation of most current empirical consumer search models is that the search process is static: consumers and firms are both myopic and play a one-shot game.

[4]See Moraga-González (2006) for a survey of the emerging structural consumer search literature.

In reality, if consumers remain in the market for multiple periods before purchasing, then using a static model will only yield estimates of the distribution of search costs for active consumers in a particular period. To a researcher, the population distribution from which the search costs of new consumers are drawn may be of interest instead of or in addition to the cross-sectional distribution in a given period. Consumers' decisions to purchase each period may be functions of their search costs; for example, consumers who do not purchase may have systematically higher or lower search costs. In such cases, static estimation techniques will suffer from dynamic selection bias. The result is that the estimated per-period search cost distributions will be biased vis-á-vis the time invariant population distribution.

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 existing consumers in the market who have not yet purchased. As in previous models, firms choose prices to maximize per-period profits taking as given the current distribution of consumer search costs in the market.[5] We can allow for firm heterogeneity in our model along the lines of Wildenbeest (2011). This dynamic model allows us to track and model how the distribution of search costs of consumers in the market evolves over time. Simulations discussed in Section 4 suggest that when consumers' purchase probabilities are increasing or decreasing in search cost, selection effects will cause the within period search cost distribution to be biased downwards or upwards, respectively, relative to the population distribution. We propose a procedure to estimate non-parametric bounds on the population search cost distribution. Our method can be easily be extended to the parametric cases where the population distribution is thought to belong to a parametric family.

We apply the proposed procedures to analyze the online markets for new hardcover copies of two popular econometrics textbooks using a dataset of daily prices collected from a large cross-section of online retailers. 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. Relative to our estimated bounds on the population distributions for the two books we consider, the medians of the static search cost distributions are 34–98% and 72–301% higher, respectively. In light of these findings, we conclude that accounting for dynamic selection effects is

[5] Allowing agents on both sides of the market to be forward looking would require us to impose substantially more structure and may still yet be intractable. However, investigating modeling and computational strategies for dynamic pricing with forward-looking consumers in the presence of informational frictions is an interesting direction for future work.

important for estimating consumer search costs.

2. The Theoretical Model

We consider a model where prices for a homogeneous, durable good are chosen by J firms in an oligopoly market, and where price acquisition is costly to a measure of consumers, each with unit demand for the product. Each period in our model consists of a firm pricing phase, followed by a consumer search and purchase phase. At the beginning of each period t, firms observe the current distribution of consumer search costs G_t, which has support on the nonnegative real line. Firms then choose prices according to a unique, symmetric mixed pricing strategy F_t with support $[\underline{p}_t, \overline{p}_t]$. Given the distribution of prices F_t, consumers then decide how intensely to search for prices and whether or not to purchase now or wait for the next period. We will describe the consumer problem first, taking the price distribution F_t as given, and then return to the firm problem.

2.1. Consumer Search Problem

In each period t, there is a unit measure of consumers with unit demand for the product and a common within-period reservation value $v_t > 0$. Each consumer is endowed with a time-invariant, per-firm search cost c which is an i.i.d. draw from a continuous population distribution H which has full support on the non-negative real line. Consumers who purchase exit the market permanently, so the distribution of search costs of active consumers in period t, denoted G_t, may vary over time.

The ex-ante identical consumers in our model are heterogeneous in the sense that obtaining individual prices may be more or less costly. With only data on prices but not on individual consumers, a necessary limitation of our model, and others in the literature on which we build, is that consumer heterogeneity is one-dimensional. Importantly, as Hong and Shum (2006), Moraga-González and Wildenbeest (2008), and Moraga-González et al. (2013) have shown, even with only the price distribution at hand, identifying features of the search cost distribution is still feasible. This heterogeneity leads consumers to search different numbers of firms to obtain prices, which in turn supports the mixed strategy pricing equilibrium among firms.

Consumers observe the equilibrium price distribution F_t and search simultaneously in period t, receiving a chosen number of price draws from F_t. Let K denote the maximum number of firms a consumer can search and let $k_t \leq K$ denote the optimal choice of the number of firms to search in period t (i.e., the number of firms in a consumer's consideration set).

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 exist meaningful 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. Importantly, Hong and Shum (2006) show that the simultaneous search model can be estimated without either observing search costs or making parametric assumptions about the search cost distribution. Additionally, De los Santos et al. (2012) and Honka and Chintagunta (2017) find that observed online search behavior is consistent with simultaneous search. Our dynamic model builds on these insights: we do not need to impose a parametric assumption and the search process resembles the optimal search rule of Morgan and Manning (1985) in the sense that consumers search simultaneously within a period while retaining the option to continue searching next period (i.e., sequential search across periods and simultaneous search within periods).

Again following the literature, we assume consumers receive one free price quote from a random firm.[6] We assume that there is no recall across periods, so prices this period are not valid next period and may not be representative of next period's price distribution. For each additional firm searched beyond the first, consumers incur their individual search cost c. Hence, the total cost for becoming fully informed about all K prices is $c(K-1)$. Moraga-González et al. (2017, Proposition 3) show that the optimal consumer search response to any atomless price distribution F_t leads to a unique grouping of consumers based on how many prices each consumer will optimally obtain. The ex-post information asymmetry among consumers is an equilibrium outcome of optimal behavior.

At the start of each period, a consumer decides on the number of firms to search, prior to observing prices. The first price is obtained for free and a cost of c is incurred for each additional firm sampled. Consumers are attempting to minimize the expected total cost of purchase this period (search costs incurred and the expected minimum price from the sampled firms). A consumer decides to purchase or not, denoted $D_t \in \{0, 1\}$, by evaluating whether the value of continuing to search is less than the best terminal utility currently available. A consumer who purchases ($D_t = 1$) does so from the firm with the lowest price of the k_t firms sampled. Since the product is durable and consumers have unit demand, a consumer who purchases in the current period exits the market permanently. At the end of each period, consumers who purchase are replaced by consumers with i.i.d.

[6] Moraga-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.

search costs drawn anew from the population distribution H.[7] Consumers who do not purchase remain in the market and retain their idiosyncratic search costs.

Before actually searching, a rational consumer calculates the optimal number of firms to search by minimizing her expected total expenditure:

$$k_t(c) = \arg\min_{k \le K} \left[c(k-1) + \int_{\underline{p}_t}^{\overline{p}_t} pk\,(1-F_t(p))^{k-1} f_t(p)\,dp \right]. \tag{1}$$

The first term in the minimand is the total cost of searching and the second term is the expected minimum price paid after sampling k firms conditional on purchasing this period. For notational convenience, we rewrite the expected value of the minimum of k sampled prices in period t (the first order statistic) as $\mathrm{E}[p_{t,k}^{(1)}]$. The tension in the consumer's intertemporal 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, given a search cost c, there is a unique solution for k.

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 equals the marginal benefit of searching the $(k+1)$-th firm;

$$c_{t,k} = \mathrm{E}\left[p_{t,k}^{(1)}\right] - \mathrm{E}\left[p_{t,k+1}^{(1)}\right]. \tag{2}$$

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; these are the consumers who possess search costs in the interval $(c_{t,k+1}, c_{t,k}]$. Combining the adding-up restriction $\sum_1^K \mu_{t,k} = 1$ 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 CDF:

$$\mu_{t,k} = \begin{cases} 1 - G_t(c_{t,1}) & k = 1, \\ G_t(c_{t,k-1}) - G_t(c_{t,k}) & 2 \le k \le K-1, \\ G_t(c_{t,K-1}) & k = K. \end{cases} \tag{3}$$

Notice that this condition involves only a finite number of values of G_t and so price data from a single market will be insufficient to identify the entire distribution. It is particularly difficult to identify the upper quantiles of the search cost distribution. For very high costs, we can only identify the smallest cost that makes a consumer want to only

[7] The measure of consumers is therefore fixed over time. However, one could also consider expansion and contraction of the market size over time if an observable measure of market size is available. This would change the relative proportion of consumers entering the market from the population distribution H and would dampen or amplify the dynamic selection bias effects as the case may be.

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.

2.2. Consumer Purchase Decision: A Model-Free Approach

We consider the possibility that forward-looking consumers may choose to continue searching beyond the current period if they expect searching will strictly improve their expected utility. Let $\sigma_t(c) \equiv \Pr(D_t = 1 \mid C = c, \mathcal{I}_t)$ denote the purchase probability (i.e., the policy function) of a consumer with search cost c in period t, where D_t is a purchase indicator and $\mathcal{I}_t$ represents the public information available in period t. In other words, $\sigma_t(c)$ is the conditional probability of purchasing and exiting the market in period t given search cost c and information $\mathcal{I}_t$. Similarly, $1 - \sigma_t(c)$ is the conditional probability of remaining in the market (not purchasing). While static models are based on the assumption that all consumers search for only one-period, uniformly, our model permits heterogeneity in search durations allowing σ_t to depend on c and other factors.

In determining this period's purchase policy function σ_t, consumers may incorporate expectations about next period's equilibrium price distribution, F_{t+1}. Prices next period are determined by next period's realized search cost distribution G_{t+1}, which is in turn determined by this period's search cost distribution G_t, the time invariant population search cost distribution H, and the policy function σ_t of interest. There are a myriad of rational ways in which consumers could solve this problem. For example, one approach would be to assume that the equilibrium policy σ_t is determined in a rational expectations equilibrium. Another option would be to use an adaptive expectations equilibrium.[8] Specifying such explicit models quickly becomes analytically intractable, even with specific functional form assumptions.[9] Instead, we take a largely model-free approach which does

[8] Chow (2011) demonstrates that adaptive expectations would be better suited for consumers in our setting, since past period prices become increasingly less relevant over time in forming expectations about future prices.

[9] To illustrate the difficulty in fully specifying, solving, and estimating a dynamic representative consumer model here, let $W_t(p \mid c)$ denote the value function of a consumer with search cost c given an initial observed price p. The value that a consumer receives from searching and purchasing in period t is $v - \mathrm{E}[p^{(1)}_{t,k(c)}] - [k(c) - 1]c$, which is the expected value of buying after searching $k(c)$ stores. On the other hand, the option value of waiting and not purchasing is $\beta\,\mathrm{E}[W_{t+1}(p' \mid c) \mid p, c]$. The Bellman equation describing this situation is

$$W_t(p \mid c) = \max_{D_t \in \{0,1\}} \left\{ \left(v - \mathrm{E}(p^{(1)}_{t,k(c)}) - (k(c) - 1)c\right) D_t + \beta\,\mathrm{E}[W_{t+1}(p' \mid c) \mid p](1 - D_t) \right\},$$

where D represents the purchase decision. To complete the model, one specify consumer beliefs about future price and search cost distributions, which are in turn determined in part by the purchase policy function, which might be determined in an equilibrium. Therefore, this is closer to a dynamic game than single agent model, and so straightforward "full solution" techniques are likely impractical here. However, we have been not require us to specify exactly how the policies σ_t are determined. However, since σ_t can depend on beliefs about F_{t+1}, our procedure follows the spirit of the optimal search rule of Morgan and Manning (1985).

We consider a general model in which we assume only that consumers follow a sequence of unspecified policy functions $\sigma_t : [0, \infty) \to [0, 1]$ for $t = 1, 2, \ldots, T$. These policies may depend on the search cost c in an arbitrary way. Any particular dynamic model one might specify for consumers will yield such a sequence of policy rules, but as we show below our approach yields easily interpretable results while being robust to misspecification of σ_t. Since σ_t is indexed by t, the purchase probabilities may depend on the search cost and price distributions G_t and P_t in the current, previous, or future periods as well as the population distribution H. This allows for the possibility of non-stationary, time- or duration-dependent purchase behavior.[10] We also note that the static case where consumers always purchase is nested in our framework since one possible policy sequence is $\sigma_t(c) = 1$ for all c and t.

Although economic theory may not guide us in selecting a particular parametric family for σ_t, it may inform us about its shape. For example, since consumers with higher search costs search fewer stores, their ex-ante probability of finding a sufficiently low price may be lower and they may be more likely to delay purchase until next period. In other words, σ_t may be monotonic. We show how to incorporate such a theoretical restriction in the course of discussing our identification and estimation results below.

Consider a single consumer making a decision about whether to purchase this period or remain in the market until the next period. If the consumer does not purchase ($D = 0$), then she remains in the market and retains her search cost next period ($C' = c$). 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 ($C' \sim H(\cdot)$). Therefore, the conditional search cost distribution next period is

$$\Pr(C' \leq c' \mid C = c, D = d) = \begin{cases} 1\{c \leq c'\} & \text{if } d = 0, \\ H(c') & \text{if } d = 1. \end{cases} \tag{4}$$

If the consumer does not purchase, then her search cost stays the same and so there is a mass point at c. When the consumer does purchase, the search cost of the consumer who replaces her is drawn anew from H and is therefore independent of the previous cost c.

able to find equilibria numerically in a simple two-period instance of this model where policy functions are assumed to have a threshold or indicator function form. These results are available from the authors upon request.

[10] We note that the procedure is robust to seasonality in prices (e.g., changes in the mean of the distribution over time), since the within-period decision in (2) depends only on differences.

The object of interest in this model is the unconditional population search cost distribution H. To find this distribution we can first integrate (4) with respect to c and d to obtain the unconditional distribution of search costs in period $t+1$:

$$G_{t+1}(c') = \int_0^{c'} (1-\sigma_t(c))\, g_t(c)\, dc + H(c') \int_0^{\infty} \sigma_t(c) g_t(c)\, dc. \tag{5}$$

Hence, next period's search cost distribution is a mixture of the current search cost distribution and the population distribution where the weights are determined by the purchase policy function. Consider the two extreme cases. If all the consumers purchase, regardless of their search costs, then next period's distribution is simply the time invariant 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, G_t.

2.3. Firm Pricing

Following the previous literature, we assume that prices are determined by a collection of J oligopolistic, static profit maximizing firms that each produce a homogeneous, durable good at a constant within-period marginal cost $r_t > 0$. Firms do not take into account that their consumers are forward-looking. This is a limitation of our model, but it yields a computationally tractable model and allows us to leverage existing results on identification and estimation of within-period search cost distributions by Hong and Shum (2006), Moraga-González and Wildenbeest (2008), Moraga-González et al. (2013), and others. For now we assume these firms are identical, but we will show below that we can also allow for vertical differentiation in the sense of Wildenbeest (2011).

Following Burdett and Judd (1983), to ensure that a market exists we assume that no firm will price below marginal cost, which implies a lower bound $\underline{p}_t \geq r_t$, and that no firm will price above the valuation, implying an upper bound $\overline{p}_t \leq v_t$. In equilibrium, firms set prices according to a symmetric mixed strategy represented by a cumulative distribution function F_t. This distribution is absolutely continuous and assigns positive density everywhere on $[\underline{p}_t, \overline{p}_t]$. Moraga-González et al. (2017) prove that under mild conditions an equilibrium exists for any number of firms. They also show that for a given consumer search behavior there is a unique symmetric equilibrium and the equilibrium price distribution must be atomless.[11] Furthermore, simulations by Moraga-González et al. (2017) suggest that uniqueness is a more general result.

Firms choose prices simultaneously at the beginning of each period according to F_t. Given the consumer search behavior derived above, the optimal pricing strategy is a

[11] Baye, Kovenock, and De Vries (1992) examine asymmetric equilibria in Varian's model and conclude that only the symmetric equilibrium survives meaningful equilibrium refinement.

symmetric mixed strategy represented by a CDF F_t on $[\underline{p}_t, \overline{p}_t]$. For firms to be willing to mix in any one period, all prices in the support must yield the same profit:

$$\Pi_t = (p - r_t) \left[\sum_{k=1}^{K} \frac{k\mu_{t,k}}{K} \left(1 - F_t(p)\right)^{k-1} \right] \text{ for all } p \in [\underline{p}_t, \overline{p}_t]. \tag{6}$$

Firms charging the highest price will only sell to consumers who search exactly once. This condition defines the constant profit level: $\Pi_t = \frac{\mu_{t,1}(\overline{p}_t - r_t)}{K}$. $F_t(p)$ is then implicitly defined by (6) and the density of the price distribution can be determined by solving the first order condition of the profit equation for $f_t(p)$:

$$f_t(p) = \frac{\sum_{k=1}^{K} k\mu_{t,k} \left(1 - F_t(p)\right)^{k-1}}{(p - r_t) \sum_{k=2}^{K} k(k-1)\mu_{t,k} \left(1 - F_t(p)\right)^{k-2}}. \tag{7}$$

Given these density values, one can consistently estimate $\mu_t = (\mu_{t,1}, \ldots, \mu_{t,K})$ in each period using maximum likelihood as in MGW. By evaluating (6) at $p = \underline{p}_t$, this condition can be rearranged and solved to obtain the per unit marginal cost:

$$r_t = \frac{\underline{p}_t \cdot \left(\sum_{k=1}^{K} k \cdot \mu_k\right) - \mu_1 \overline{p}_t}{\sum_{k=2}^{K} k \cdot \mu_k} \tag{8}$$

2.4. Pricing with Vertically Differentiated Firms

So far, we assumed firms were identical and equally likely to be sampled by a consumer. Here, we consider vertically differentiated firms following the approach of Wildenbeest (2011). Firms still sell a homogeneous product, but they can differentiate themselves by offering different levels of in-store services (quality). Suppose that all consumers have the same preference for quality through an additively separable utility function of the form $u_{jt} = v_{jt} - p_{jt}$, where v_{jt} and p_{jt} are the consumer's valuation and the price of the product at firm j in period t, respectively. Consumers know their firm-specific valuations, but the prices are unknown until the consumer searches. Consumers purchase the product from the firm providing them the highest utility level.

The consumer's firm-specific valuation is made up of two components: a common value that is derived from consumption of the homogeneous good and a firm-specific value coming from level of service quality. As in Armstrong (2008), firms compete directly in utility space by offering consumers a quality-price pair. We assume that firm quality is fixed in the short term. In addition, there are constant returns to firm quality and markets for quality input factors are perfectly competitive. Under this assumption, as shown by Wildenbeest (2011), the consumer valuation for firm j's product is additively separable as $v_{jt} = x_t + q_j$, where x_t is the common utility level from the actual consumption of good in period t and q_j is firm j's quality.

Each firm j will choose q_j to maximize the valuation-cost margin, $v_{jt} - r(q_j)$, where $r(q_j)$ denotes firm j's cost of production given the quality level q_j. Assuming perfect competition in input markets and constant returns to scale in production, Euler's theorem implies that $r(q_j) = q_j$. The valuation-cost markup is constant and independent of quality and therefore the profit margins of all the firms are independent of the quality choice. Although firms might offer different prices due to their quality choices, the quality-adjusted price is the same for everyone. Firms are symmetric along the utility dimension, allowing for the development of a symmetric mixing equilibrium over utilities.

2.5. Extensions for Cases with More Observable Information

If both prices and quantities are observed in every period, the sharpness of our bounds can be improved in a way similar to Hortaçsu and Syverson (2004). With market shares and prices in hand, the within period equilibrium market clearing, with some mild normalization conditions, nonparametrically identifies G_t—the search cost CDF—at the critical cutoff costs. If we also know marginal cost, then g_t–the search cost PDF—can be derived directly from the competition assumption. Otherwise, it can be estimated from the price and market share data. The advantage that this additional information provides is that we are deriving G_t, g_t, and c_k from the model, rather than estimating them. Therefore, in this setting our method becomes a one-step estimator, rather than a two-step estimator where the first stage quantities are used as inputs in the second stage. Thus, if one observes prices and quantities this can yield sharper and more precisely estimated bounds for H.

Additionally, if individual data on search behavior is observed our model can be extended to allow for firm prominence along the lines of Armstrong, Vickers, and Zhou (2009). For example, De los Santos (2018) observes individual consumer characteristics and search behavior and estimates a model of search for New York Times Best Sellers with firm prominence and consumer heterogeneity in search intensities. In our setting, suppose that prominent firms have a higher probability of being sampled than other firms.[12] This situation could be captured in our model with the addition of firm-specific consideration probabilities ψ_i where $i = 1$ for prominent firms and $i = 0$ for other firms. This implies that the expected period profit in (6) would be different for prominent and non-prominent firms. While this would increase the computational burden of estimating the model, the MLE problem is essentially solved in the same fashion as before, based on profit indifference conditions that implicitly define $F_{t,i}$ each period which in turn determine $f_{t,i}$. Additional information such as this would improve the quality of estimated bounds

[12]One technical caveat is that the prominent firm cannot be sampled first by *all* consumers. Armstrong et al. (2009) show that firms will use a pure strategy and the price disparity is more limited and arises directly from search order effects.

by controlling for variation in prices arising from search order effects.

3. Identification

Our identification strategy is constructive, and can be summarized briefly as follows. We consider identification of the population search cost distribution H given a sequence of observable per-period price distributions F_t for $t = 1, \ldots, T$. First, in each period t the price distribution F_t identifies the per-period distribution function G_t at a finite number of points $\{c_{t,k} : k = 1, \ldots, K\}$ by Proposition 1 of Moraga-González et al. (2013). Then, we compare the identified values of G_t and G_{t+1} in adjacent periods. Since G_{t+1} is a σ_t-weighted mixture of G_t and H, letting σ_t vary over the space of probability-valued policies yields upper and lower bounds on H at fixed points. These bounds can be improved by exploiting variation in the price distributions across many periods. Hence, the asymptotic framework we consider for estimation is one where there are N_t observations drawn from F_t in each period $t = 1, \ldots, T$ with $2 \leq T < \infty$. Therefore, with only price data and no additional assumptions on the structure of the model, we can identify informative bounds on H. Finally, in an extension we show that theoretical restrictions, such as monotonicity of purchase probabilities in search costs, could be used to further refine the bounds.

3.1. Bounds with Unrestricted Policies

Using the observed per-period price distributions and Proposition 1 of Moraga-González et al. (2013), we take it as given 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 Σ be the set of all functions mapping $[0, \infty)$ to $[0, 1]$. Using the weak restriction that $\sigma_t \in \Sigma$ for all t will provide information about H. For a given cost value c', we will bound $H(c')$ by finding two functions in Σ that maximize and minimize $G_{t+1}(c')$, as defined in (5). For a given cost c', let $\sigma_t^{\mathrm{L}}(\cdot; c')$ denote the function for which $G_{t+1}(c')$ reaches the lower bound and let $\sigma_t^{\mathrm{U}}(\cdot; c')$ denote the function which achieves the upper bound. The purchase probabilities that generate the lower and upper bounds on $G_{t+1}(c')$, and the implied bounds are given by the following proposition. Proofs of this result and others can be found in the appendix.

Proposition 1. *When σ_t is unrestricted in Σ, the conditional, purchase probabilities that generate the lower and upper and bounds on G_{t+1} for a given cost c' are*

$$\sigma_t^L(c; c') = \begin{cases} 1 & \text{if } c \leq c', \\ 0 & \text{if } c > c', \end{cases} \quad \text{and} \quad \sigma_t^U(c; c') = \begin{cases} 0 & \text{if } c \leq c', \\ 1 & \text{if } c > c'. \end{cases}$$

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

$$G_{t+1}^L(c') \equiv G_t(c')H(c') \leq G_{t+1}(c') \leq G_t(c') + [1 - G_t(c')]H(c') \equiv G_{t+1}^U(c').$$

Next, we use this information to learn about the population distribution H. For a cost c at 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 $H(c)$:

$$H^{\mathrm{L}}_{t+1}(c) \equiv \max\left\{\frac{G_{t+1}(c) - G_t(c)}{1 - G_t(c)}, 0\right\} \leq H(c) \leq \min\left\{\frac{G_{t+1}(c)}{G_t(c)}, 1\right\} \equiv H^{\mathrm{U}}_{t+1}(c) \tag{9}$$

These bounds only depend on the estimable period search cost distributions. The following proposition establishes that except in singular cases, the bounds are different from the trivial bounds $[0,1]$ and are therefore informative. It also shows that the bounding interval is in fact sharp, meaning that it is as small as possible with the given information.

Proposition 2. *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 $H^{L}_{t+1}(c) > 0$ or $H^{U}_{t+1}(c) < 1$. In other words, at least one of the bounds will be "informative" in the sense of being different from the trivial bounds $[0,1]$.

b. If $G_t(c) = G_{t+1}(c)$, then $H^{L}_{t+1}(c) = 0$ and $H^{U}_{t+1}(c) = 1$. In other words, both bounds are trivial.

c. Let $G_t(\cdot)$, $G_{t+1}(\cdot)$, and $h \in [H^{L}_{t+1}(c), H^{U}_{t+1}(c)]$ be given. Then there exists a policy σ_t that would rationalize $H(c) = h$. In other words, the bounding interval is sharp.

Recall that the time subscripts for the bounds above indicate the within-period distributions used to obtain the bounds: the bounds H^{L}_{t+1} and H^{U}_{t+1} are constructed using only G_t and G_{t+1}. However, thus far we have only used information in adjacent periods t and $t+1$. Since the H distribution is time invariant, we can also aggregate the information in these bounds across periods. The most informative bounds are the least upper bound and the greatest lower bound on $H(c)$ at a given cost value c for any given adjacent periods. Aggregating across periods yields what we will refer to as the *envelope bounds*:

$$H^{\mathrm{L}}(c) \equiv \max_{t \in \{1,\ldots,T-1\}} H^{\mathrm{L}}_{t+1}(c) \qquad \text{and} \qquad H^{\mathrm{U}}(c) \equiv \min_{t \in \{1,\ldots,T-1\}} H^{\mathrm{U}}_{t+1}(c).$$

These are the bounds that we estimate and report in the application described in Section 6.

We note that because we aggregate information across periods, the envelope bounds $\hat{H}^{\mathrm{L}}$ and $\hat{H}^{\mathrm{U}}$ 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 cross for some lower quantiles (where there is typically more information about search costs). As in generalized method of moments (GMM) estimation of finite-dimensional parameters (Hansen, 1982), this suggests that we should choose our estimate of H to minimize a loss function. For given bounds H^{L} and H^{U}, the average of the two bounds will minimize the directional L^2 loss function.[13] Therefore we also report estimates of the average of the bounds:

$$\overline{H} = \frac{H^{\mathrm{L}} + H^{\mathrm{U}}}{2}.$$

We conclude with a brief discussion of related results in the search literature. In a static model, Moraga-González et al. (2013, Proposition 2) show that if price distributions are observed across many markets with the same search cost distribution, but with variation in valuations or marginal costs, then the full search cost distribution is nonparametrically identified. However, in our dynamic model the period-market-specific search cost distributions will be different across markets, due to the stochastic nature of individual consumer purchase decisions. There is also a single online market for textbooks in our application, so in the remainder, we focus on the case where data are sampled from only one market but where, unfortunately, the identification problem is more difficult.

There is also a related literature on models of labor search, where generally the full wage distribution can be identified, even below the reservation wage where no offers are accepted, by assuming a functional form that is recoverable from the observed, truncated distribution or observing a random sample of wages offers spanning the full distribution.[14] Unlike our model, sequential search is the standard technology for labor models where workers receive one wage offer at a time. These techniques are not directly applicable to our model where consumers and firms are facing distinct price and search costs distributions, respectively, every period.

3.2. Improved Bounds with Monotonic Policies

Finally, we consider whether we can strengthen the bounds by imposing weak monotonicity of the policy functions σ_t. Note that without this assumption, as established in Proposition 1 above, the minimizing policy function in Σ is weakly increasing 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 set of weakly decreasing functions $\Sigma_{\mathrm{M}} \subset \Sigma$.

Proposition 3. *When $\sigma_t \in \Sigma_M$, the conditional purchase probabilities that generate the upper*

[13] Given the two inequality constraints $H^{\mathrm{L}}(c) \leq H(c) \leq H^{\mathrm{U}}(c)$, we consider the directional L^2 loss function $L(H; H^{\mathrm{L}}, H^{\mathrm{U}}) = \int_0^\infty \left[\left|H^{\mathrm{L}}(c) - H(c)\right|_+^2 + \left|H^{\mathrm{U}}(c) - H(c)\right|_-^2 \right] dc$, where for any $z \in \mathbb{R}$ we define $|z|_- = |z| 1\{z < 0\}$ and $|z|_+ = |z| 1\{z > 0\}$.

[14] See Mortensen (1986) and Eckstein and van den Berg (2007) for surveys of this literature.

bound on G_{t+1} for a given cost c' are:

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

The implied lower bound on H is

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

The upper bound given in (9) still applies.

4. Dynamic Selection Effects: Simulation Evidence and Theoretical Results

Here we investigate the effects of dynamic selection on the period search cost distribution via a small-scale simulation study. We also discuss some additional theoretical results which characterize the extent to which dynamic selection causes problems with static estimation. In particular, we focus on the characteristics of the consumers' purchase policy function that are related to the dynamic selection bias. We define this bias to be the difference between the per-period search cost distributions obtained by a static analysis, G_t, and the population search cost distribution, H.

For each simulation, we choose: 1) a starting period search cost distribution, G_0, 2) a population distribution, H, and 3) a time-invariant purchase probability policy function $\sigma(\cdot)$. Each simulation begins with 1,000 initial consumers with costs drawn randomly from G_0. We then apply the policy function σ to determine which 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 drawn from the population distribution H. We repeat this process for 10,000 periods.

To analyze the effects of the functional form of σ on the evolution of the search cost distribution, we use an assortment of functions that are monotonically increasing, monotonically decreasing, non-monotonic, or independent of search cost. We considered all combinations of four initial distributions, three population distributions, and 20 policy functions for a total of 240 specifications.[15]

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. In this scenario, as search cost increases, consumers become less likely

[15]See Appendix A where Table 7 displays the primitives of four representative specifications which we use to illustrate our findings, and for the full details on the collection of distributions and policy functions used.

to purchase relative to their lower search cost counterparts. Intuitively, this implies that over time the fraction of consumers with high search costs grows. At the same time, low search cost consumers are purchasing and exiting the market; thus, the per period distribution is becoming more concentrated with higher search cost consumers. 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 less intensely before purchasing.

The opposite outcome occurs in the situation where the purchase probability is monotonically increasing in search cost. Regardless of the starting or population search cost distributions used, selection effects result in a downward distributional bias. This result arises from the same forces as above: consumers with higher search costs are purchasing at relatively higher rates, so the market becomes more concentrated with lower cost consumers over time.

When the probability of purchase is non-monotonic, there can be both positive and negative bias at various quantiles. The direction of bias depends on the nature of the non-monotonicity of the purchase probability. The bias is primarily driven by the purchase decisions of consumers who possess more extreme search cost realizations. In this case, consumers in both tails of the distribution have very low purchase probabilities. The result is that the search cost distribution spreads out over time, resulting in both types of bias. The direction of the bias at specific quantiles is determined by which tail of the purchase probability function has relatively more weight. In this case, we eventually see positive bias at the median because it is the lower-cost consumers who tend to purchase more frequently.

Finally, if the purchase probability is a constant, say $\sigma(\cdot) = p$, then there is a closed-form solution for period t's search cost distribution:

$$G_t(c) = (1-p)G_{t-1}(c) + pH(c).$$

Extending 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:

$$G_t(c) = (1-p)^t G_0(c) + pH(c)\sum_{s=1}^{t}(1-p)^{s-1}.$$

Taking the limit reveals that the period search cost distributions converge towards the population distribution, so there is no persistent bias: $\lim_{t\to\infty} G_t(c) = H(c)$ for all c. The rate of convergence depends on the consumers' purchase probability, with faster convergence as p increases.

We conclude this section with Proposition 4, which characterizes the nature of the bias over time in terms of the unconditional purchase probability $R_t(\sigma_t) \equiv \int_0^\infty \sigma_t(c)g_t(c)\,dc$.

The first part establishes that if G_t is biased in the current period, then G_{t+1} will be biased in the next period unless the non-purchase probability is sufficiently large. The second part states that when the within-period search cost distribution has reached an unbiased steady state, then it must be the case that the purchase policy function is constant (and therefore trivially equal to the unconditional purchase probability).

Proposition 4. *a. If $G_t \neq H$, then $G_{t+1} \neq H$ for any policy function σ_t such that $1 - R_t(\sigma_t) < \inf_{c'} \frac{g_t(c')}{h(c')}$.*

b. If $G_t = G_{t+1} = H$, then σ_t must be constant.

5. Estimation

In this section, we propose a method to estimate the model described above using only pooled cross-sectional 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., $p_{t,1} \leq \ldots \leq p_{t,N_t}$ for all t). The estimation method consists of two or three stages: 1) nonparametric estimation of the within-period search cost distributions G_t using N_t prices in each period t, 2) nonparametric estimation of bounds on the entry distribution H using variation in G_t across periods t, 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 $p_{t,N_t} = \overline{p}_t$.[16] The maximum observed price, $\overline{p}_t$, is also a super-consistent estimate of the consumer valuation, v_t (Hong and Shum, 2006; Moraga-González and Wildenbeest, 2008).

Recall that $\mu_{t,k}$ is the measure of consumers who search exactly k firms in period t. The maximum likelihood estimation problem for $\mu_t = (\mu_{t,1}, \ldots, \mu_{t,K})$ in each period t is:

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

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

where $f_t(p; \mu_t)$ is defined in (7).

[16] Refer to Kiefer and Neumann (1993) and Donald and Paarsch (1993) regarding the use of order statistics to estimate bounds.

Each search cost cutoff $c_{t,k}$ for $k = 1, \ldots, K-1$ can be found by evaluating the following integral:[17]

$$c_{t,k} = \int_{\underline{p}}^{\overline{p}} F_t^{-1}(z) \left((k+1)z - 1\right) (1-z)^{k-1} \, dz.$$

Recall that the price distribution is strictly monotonically increasing, so the inverse of F_t exists. Given estimates of $\mu_{t,k}$ and $c_{t,k}$ for each k, we use (3) to estimate the values of the search cost CDF $G_t(c_{t,k})$ at each of cutoff search cost values. These are the estimates one would obtain when carrying out a static analysis. If we have N_t observations in period t, then as established by Moraga-González and Wildenbeest (2008), the maximum likelihood estimates $\{(\hat{c}_{t,k}, \hat{G}_t(c_{t,k})\}_{k=1}^{K}$ are consistent as $N_t \to \infty$.

Next, we estimate the nonparametric upper and lower bounds on the population distribution H using the estimated cutoffs $c_{t,k}$ and CDF values $G_t(c_{t,k})$ as described in Section 3. Specifically, we define the estimates of the bounds as follows:

$$\hat{H}_{t+1}^{\mathrm{L}}(c) \equiv \max\left\{\frac{\hat{G}_{t+1}(c) - \hat{G}_t(c)}{1 - \hat{G}_t(c)}, 0\right\},$$

$$\hat{H}_{t+1}^{\mathrm{U}}(c) \equiv \min\left\{\frac{\hat{G}_{t+1}(c)}{\hat{G}_t(c)}, 1\right\}.$$

Provided that the sample sizes N_t and N_{t+1} in adjacent periods tend to infinity, we can obtain consistent estimates of the CDF values $\hat{G}_t(c)$ and $\hat{G}_{t+1}(c)$. The following proposition establishes that the endpoint estimates are consistent and therefore we have a consistent interval estimate for the population identified interval $[H^{\mathrm{U}}(c), H^{\mathrm{L}}(c)] \ni H(c)$.

Proposition 5. *Let $\sigma_t \in \Sigma$ and $c \in \{c_{t,k}\}_{k=1}^{K} \cap \{c_{t+1,k}\}_{k=1}^{K}$. If $N_t \to \infty$ and $N_{t+1} \to \infty$, then $\hat{H}_{t+1}^{L}(c) \xrightarrow{\mathrm{p}} H_{t+1}^{L}(c)$, $\hat{H}_{t+1}^{U}(c) \xrightarrow{\mathrm{p}} H_{t+1}^{U}(c)$, and therefore $d_H\left([\hat{H}_{t+1}^{L}(c), \hat{H}_{t+1}^{U}(c)], [H_{t+1}^{L}(c), H_{t+1}^{U}(c)]\right) \xrightarrow{\mathrm{p}} 0$ where $d_H(A, B)$ denotes the Hausdorff distance between sets A and B.*

To allow for vertical differentiation as described in Section 2.4, we need to transform prices into utilities before estimating the model. As noted by Wildenbeest (2011), consumer valuations for each firm can be estimated via a fixed effects regression of prices on firm dummies: $p_{jt} = \alpha + \delta_1 + \ldots + \delta_N + \varepsilon_{jt}$. The intercept, α, is the common component across stores and the fixed effects δ_j represent firm-specific services and in-store experiences (qualities). Stated another way, we can simply demean the prices to obtain utility values. Pooling prices by firm and noting that $p_{jt} = v_j - u_{jt}$, we can estimate the firm-specific

[17] Moraga-González et al. (2017, 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, r_t, and $G_t(c)$, are guaranteed to be numerically solvable.

quality v_j using the mean price of firm j over time and the firm-time-specific utility values u_{jt} using the negative demeaned prices $v_j - p_{jt}$. Again, as in Wildenbeest (2011), estimation of the period-specific search cost distributions proceeds using the same likelihood function as above, but with the negative utility values used instead of the observed prices.

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

$$Q_T(\theta) = \sum_{t=1}^{T-1} \sum_{k=1}^{K} \left| \hat{H}_t^{\mathrm{L}}(c_{t,k}) - H(c_{t,k};\theta) \right|_+^2 + \left| \hat{H}_t^{\mathrm{U}}(c_{t,k}) - H(c_{t,k};\theta) \right|_-^2 \tag{10}$$

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. Application: Online Market for Econometrics Textbooks

In this section, we apply our dynamic estimation procedure to data collected from the online markets for two widely-used graduate econometrics textbooks. The online market for textbooks 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, 2018).[18] Furthermore, the 2013 publishing industry's annual review found that 44% of American 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.

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 as many as 40–60% of consumers visit only one firm prior to purchasing.[19] More recently, De los Santos (2018) 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. Using similar data, Johnson, Moe, Fader, Bellman, and Lohse (2004) also find that the

[18] Refer to Clay et al. (2001) for a review of the online book industry.

[19] Refer to Newman (1977) for a extensive review of the literature for search in non-electronic markets.

average online book shopper only searched 1.2 firms before purchasing. Furthermore, De los Santos et al. (2012) and Honka and Chintagunta (2017) both find strong evidence 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 over 236 days from nine online retail outlets[20] from March to November 2006 for new hardcover editions of *Advanced Econometrics* by Amemiya (1985) and *Microeconometrics* by Cameron and Trivedi (2005). We abstract away from differences between sites and the individual "marketplace" sellers present on some sites, although we do allow for observable seller heterogeneity by type as we describe below. By incorporating all posted prices on all sites, the implicit assumption we make is that these prices are representative draws from the overall distribution of prices among the relevant set of online retailers that consumers search. An alternative would be to sequentially model the decisions of consumers first among sites (e.g., Amazon vs. Barnes & Noble) and then among specific sellers on the site (e.g., Amazon.com or an individual marketplace seller), but such a model is beyond the scope of this paper.

For the purposes for this analysis, we classified sellers as "verified" retailers if we could match the seller ID of the listing to either a physical or online firm. The first column of Table 1 contains overall summary statistics on prices across all periods for both books. The remaining columns list the same statistics but by seller type—for non-verified, verified (but non-major), and major retailers. 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.[21] 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. 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 observed price dispersion.

For both books, a sizable proportion of the listings are from what we refer to as "unverified" sellers. These are sellers that appear to be individual, non-commercial entities who generally are simply selling a single book. These sellers may have little experience in selling textbooks online. In this environment we think that our model, in which firms are

[20]The online retailers are Abebooks, Alibris, Amazon, Barnes & Noble, Half.com, Overstock, Powell's, Super Book Deals, and Walmart.

[21]Powell's, Super Book Deals, Walmart, Barnes & Noble, Amazon, and Overstock are considered major retailers for this application.

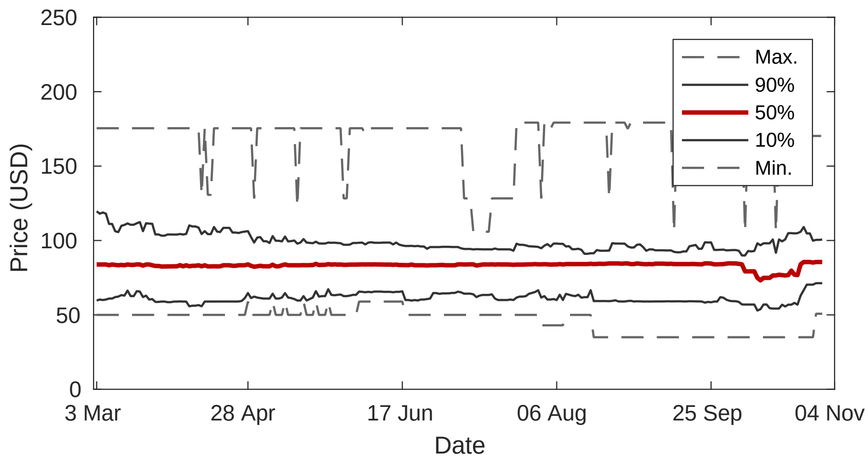
Table 1: Price Summary Statistics

	All Retailers	Non-Verified Retailers	Verified Retailers	Major Retailers
Amemiya (1985)				
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
Minimum	34.99	35.00	34.99	52.50
St. Dev.	19.96	30.41	13.59	6.60
Observations	11,457	2,874	8,583	1,018
Cameron and Trivedi (2005)				
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
Minimum	53.01	53.01	56.25	56.25
St. Dev.	22.72	30.91	14.93	4.46
Observations	15,791	5,951	9,841	1,163

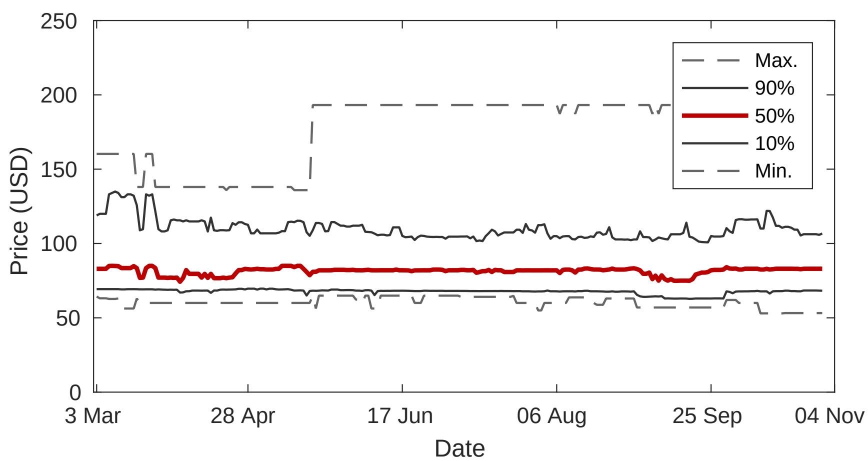
myopic, is a reasonable modeling compromise.

Quantiles and extrema of the daily price distributions for both books are plotted in Figure 1. The markets for both books are characterized by persistent and significant price dispersion, but also persistence in prices at the retailer level. For Amemiya, the minimum price is offered by a verified retailer in 95% of periods but the maximum price is posted by a verified retailer only once in the sample. For Cameron and Trivedi, the minimum per-period price is offered by a verified retailer in 69% of periods but the maximum price is only offered by a verified retailer on 7% of the days in our sample.

As can be seen in Table 1, the different types of sellers offer rather different prices, so the assumption of identical firms does not seem appropriate in this context. Therefore, we use the aforementioned approach of Wildenbeest (2011) to account for vertical differentiation among seller groups in this application. After transforming prices to obtain type-specific qualities and utilities, we carry out Kolmogorov-Smirnov (K-S) tests for distributional equality of the utility values in each of the 236 periods in our sample. There are three seller types, and therefore we carry out tests for all three pairwise comparisons for each period. To summarize the results, we report the fraction of periods where we fail to reject the null hypothesis of distributional equality for all pairings. At the 5% level, we fail to reject distributional equality in 84% of periods for Amemiya and 67% of periods for Cameron



(a) Amemiya (1985)



(b) Cameron and Trivedi (2005)

Figure 1: The Evolution of Price Distributions

and Trivedi. At the 1% level, we fail to reject distributional equality in all pairwise tests in 98% and 99% of periods, respectively.[22]

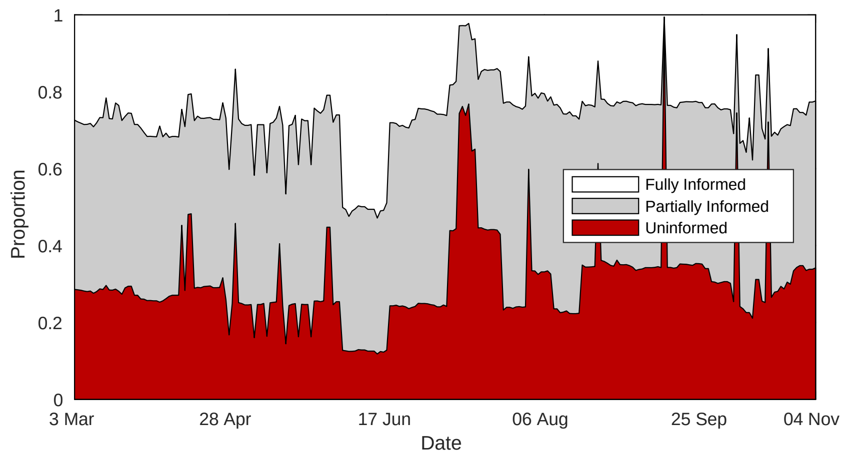
6.2. Results

The first stage of our procedure involves estimating the period search cost distributions G_t for each period 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 $K = 15$. To motivate this choice, we note that over our sample the online marketplaces average around 12 listings per day. Becoming a fully informed consumer with $K = 15$ (searching all K firms) then amounts to searching a marketplace site plus most of the major retailers. This choice of K is also consistent with Johnson et al. (2004) and De los Santos (2018), who find that the average online book shopper only searches, respectively, 1.2 or 1.29 online retailers. 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 $K = 15$.

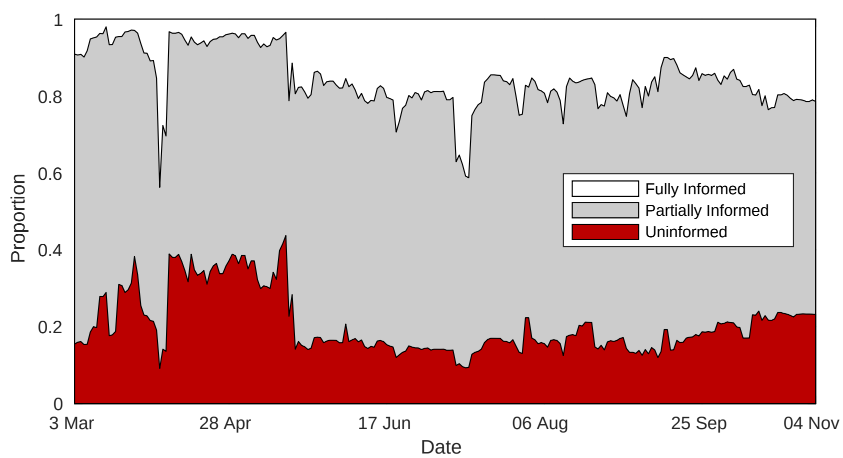
We allow marginal costs in the model to vary over time. Over all periods, the average estimated marginal cost for Amemiya was \$36.67, or about 44% of the overall average purchase price. For Cameron and Trivedi, the average estimated marginal cost was \$53.16, or about 62% of the average price. These results are broadly in line with typical marginal costs for retail bookstores, which are around 40–50% of the purchase price (Ashford, 2009; National Association of College Stores, 2009; Risk Management Association, 2016).

Traditionally, the theoretical literature on search distinguishes between two types of consumers: fully informed and fully uninformed consumers (see, e.g. Widle and Schwartz, 1979; Varian, 1980; Stahl, 1989). Empirically, Hong and Shum (2006) and Moraga-González and Wildenbeest (2008) find that a sizable proportion of consumers fall into the two polarized type of consumers. Additionally, Wildenbeest (2011) finds that basically no consumers are in the partially informed groups. In the notation above, the proportion of uninformed consumers in period t is $\mu_{t,1}$ and the proportion of fully informed consumers

[22] Additionally, pairwise Mann-Whitney rank-sum tests lead us to a similar conclusion. At the 5% level, we cannot reject the null hypothesis of distributional equality for 93% and 99% of periods for Amemiya and Cameron and Trivedi, respectively.



(a) Amemiya (1985)



(b) Cameron and Trivedi (2005)

Figure 2: Proportions of Fully Informed, Partially Informed, and Uninformed Consumers

is $\mu_{t,15}$. Hence, the proportion of partially informed consumers is $\sum_{k=2}^{14} \mu_{t,k}$.

Figure 2 shows the proportions of fully informed, partially informed, and fully uninformed consumers for both books. For Amemiya, around one quarter of consumers become fully informed in most periods. Similarly, in most periods the fraction of fully uninformed consumers is around 25–30%. For Cameron and Trivedi, in most periods about 15–20% of consumers are fully informed and 20% are fully uninformed. Periods that are characterized by a lower degree of price dispersion lead to a higher fraction of informed consumers. Still, the sizable fraction of uninformed consumers—those who do not search—implies that we cannot identify the upper tail of the search cost distribution. In other words, the estimated within-period search cost CDFs will not reach 1.

From the first stage, the full estimated per-period search cost distributions are also of interest. Table 2 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 \$10.99 on average. For Cameron and Trivedi, it is about \$8.58 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. The estimated bounds provide evidence that these selection effects may cause the period distribution to systematically shift away from the population distribution. This can also be seen in our simulation study in Section 4.

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 3 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 $\hat{H}^{\mathrm{L}}$ and $\hat{H}^{\mathrm{U}}$ for the case where σ_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, in the lower panels we also plot and report 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 $\hat{H}^{\mathrm{L}}$, $\hat{H}^{\mathrm{U}}$, 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 in the \$8.50–11.00 range. The first two rows of Table 3 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, using both the nonparametric average and the parametric approaching, are smaller than for the per-period distributions for both books. In the lower panel of Table 3, we report estimates of $\hat{H}^{\mathrm{L}}$, $\hat{H}^{\mathrm{U}}$, and $\hat{H}$ which exploit monotonicity of σ_t. After imposing monotonicity, the estimated population CDF H places more weight on lower search cost values.

We also compare the quantiles obtained by estimating two parametric models of the form $H(\cdot) = H(\cdot;\theta)$, where θ is estimated using the NLS approach outlined in Section 5. We consider models where H is the CDF of the exponential distribution, mean θ and variance θ^2 and where H is the CDF of a two-exponential mixture with mean parameters θ_1 and θ_2 respectively. The estimates are reported in Table 4 along with bootstrap standard errors in parentheses, based on the same 1,000 replication samples used before.

Table 2: Estimated Search Cost Quantiles, Static Utility Model

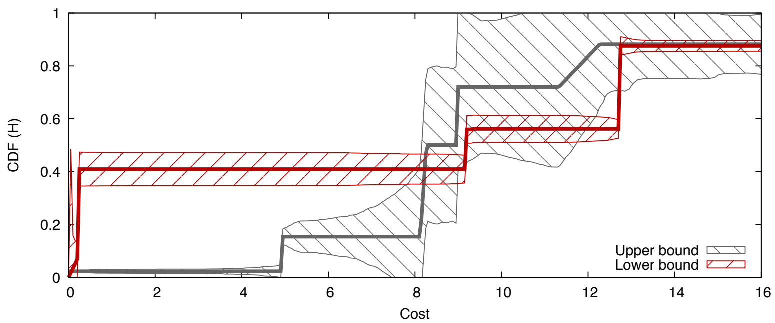
	Amemiya			Cameron and Trivedi		
Specification	0.25	0.50	0.75	0.25	0.50	0.75
Mean across periods	3.82	11.10	13.56	4.55	8.67	12.47
	(0.23)	(0.27)	(0.48)	(0.14)	(0.13)	(0.17)
Median across periods	0.48	10.99	14.44	5.43	8.58	12.11
	(1.73)	(0.19)	(0.28)	(0.11)	(0.12)	(0.15)

Note: Bootstrap standard errors based on 1,000 replications are reported in parentheses.

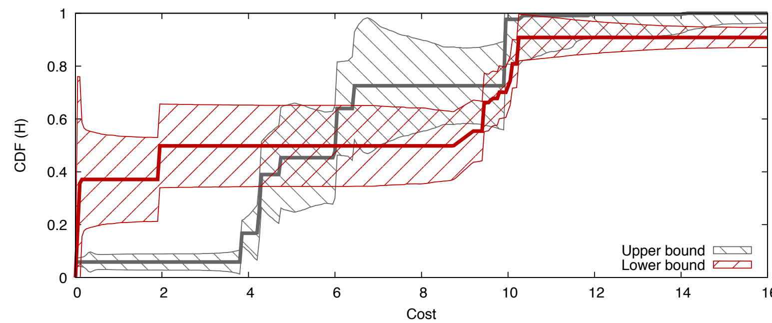
Table 3: Estimated Search Cost Quantiles

	Amemiya			Cameron and Trivedi		
Specification	0.25	0.50	0.75	0.25	0.50	0.75
Unrestricted σ						
Nonparametric H, Lower	0.23	5.62	11.36	0.07	2.16	10.01
	(0.01)	(0.02)	(0.28)	(0.03)	(0.03)	(1.83)
Nonparametric H, Upper	5.85	8.26	9.61	4.20	5.05	6.78
	(0.55)	(0.55)	(0.93)	(0.37)	(0.49)	(1.06)
Nonparametric H, Average	4.26	8.57	12.35	1.10	5.06	9.83
	(0.02)	(1.45)	(0.88)	(0.03)	(1.19)	(0.87)
Parametric H, Exponential	1.84	4.44	8.87	1.81	4.37	8.75
	(0.17)	(0.40)	(0.81)	(0.11)	(0.26)	(0.52)
Parametric H, Mixture	0.85	6.42	15.94	1.05	4.48	10.35
	(0.08)	(0.72)	(1.17)	(0.16)	(0.29)	(0.38)
Monotonic σ						
Nonparametric H, Lower	0.12	0.23	11.29	0.07	1.34	9.92
	(0.01)	(0.02)	(0.02)	(0.03)	(0.03)	(1.83)
Nonparametric H, Upper	5.85	8.26	9.61	4.20	5.05	6.78
	(0.55)	(0.55)	(0.93)	(0.37)	(0.49)	(1.06)
Nonparametric H, Average	0.22	8.20	12.30	1.11	4.78	9.83
	(0.02)	(1.45)	(0.88)	(0.03)	(1.19)	(0.87)
Parametric H, Exponential	0.70	1.68	3.36	0.75	1.80	3.61
	(0.12)	(0.29)	(0.58)	(0.08)	(0.19)	(0.39)
Parametric H, Mixture	0.00	4.46	15.38	0.12	3.74	9.93
	(0.04)	(0.68)	(1.03)	(0.14)	(0.27)	(0.32)

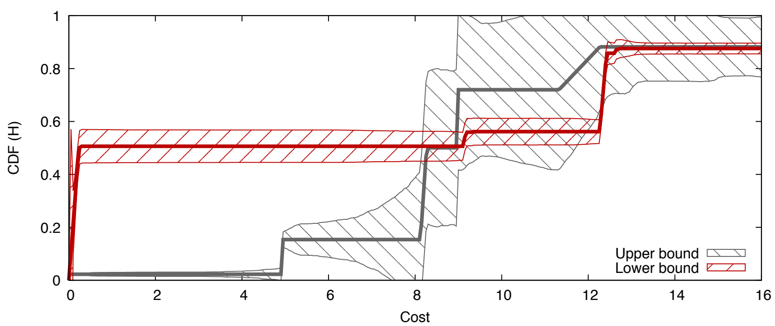
Note: Bootstrap standard errors are reported in parentheses using 1000 replications.



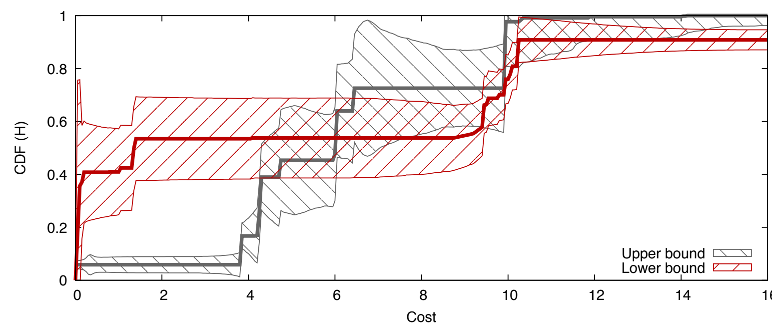
(a) Amemiya (1985), Unrestricted σ



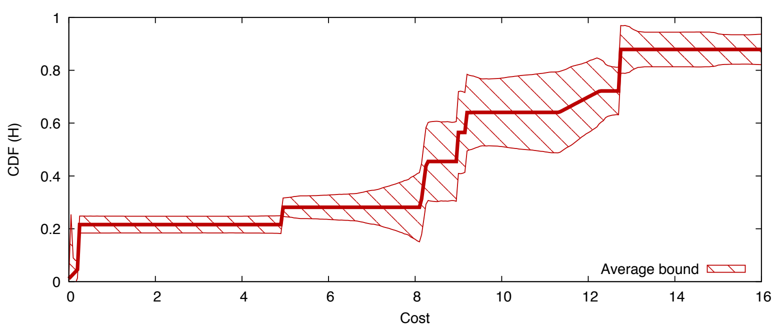
(b) Cameron and Trivedi (2005), Unrestricted σ



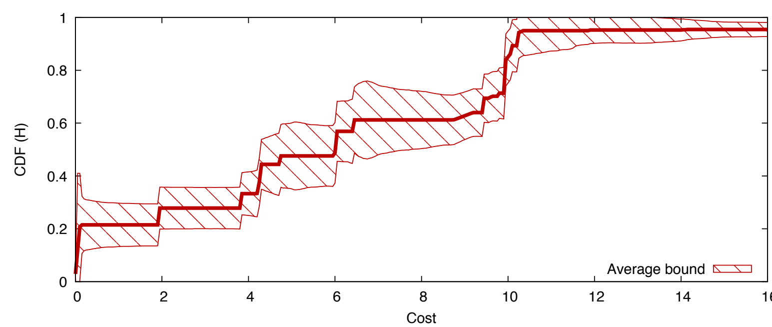
(c) Amemiya (1985), Monotonic σ



(d) Cameron and Trivedi (2005), Monotonic σ



(e) Amemiya (1985), Unrestricted σ



(f) Cameron and Trivedi (2005), Unrestricted σ

Figure 3: Estimated Bounds for H with 95% Confidence Bands

Figure 4 plots the estimated CDFs for both books and for both parametric models, with and without imposing monotonicity. For the exponential model with σ_t being unrestricted, we estimate the mean search cost for Amemiya to be \$6.40 (0.58), with a median of \$4.44 (0.40). For Cameron and Trivedi, we estimate that the mean search cost is \$6.31 (0.37) and median is \$4.37 (0.26). When imposing monotonicity of σ_t, the estimated means are much lower, at \$2.43 (0.42) and \$2.60 (0.28), respectively. The same is true for the medians, at \$1.68 (0.29) and \$1.80 (0.29).

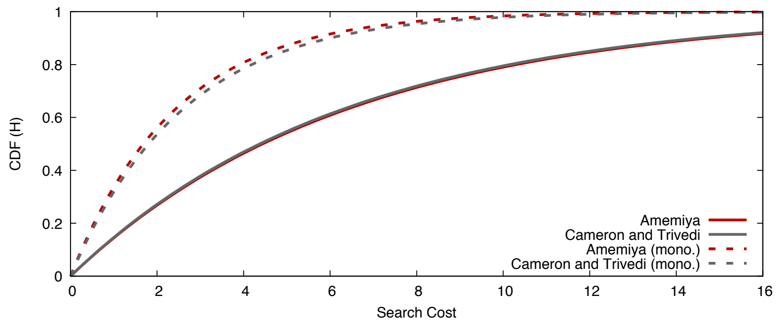
Given the apparent bimodality observed in the nonparametric estimates, we also estimated a model based on a mixture of two exponential distributions. Let θ_1 and θ_2 denote the two mean parameters and let α and $1-\alpha$ be the relative mixing weights. With a monotonic σ_t, we estimate the two means to be \$0.003 (0.001) and \$15.75 (1.65) for Amemiya, with $\hat{\alpha} = 0.34(0.03)$. For Cameron and Trivedi, the means are \$0.003 (0.004) and \$8.93 (0.50), with $\hat{\alpha} = 0.24(0.03)$. The results are qualitatively similar for the unrestricted σ_t case, as reported in the lower panel of Table 4.

Table 4: Parametric Estimates

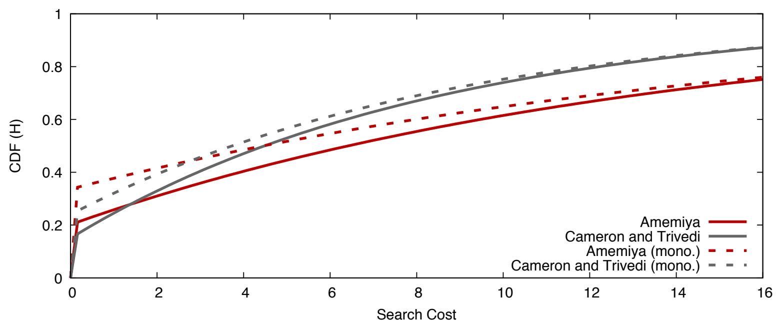
		Amemiya		Cameron and Trivedi	
Specification	Param.	Estimate	S.E.	Estimate	S.E.
Unrestricted σ					
Exponential	θ	6.402	(0.583)	6.309	(0.372)
Mixture	θ_1	0.010	(0.001)	0.001	(0.004)
	θ_2	13.721	(1.811)	8.460	(0.569)
	α	0.202	(0.034)	0.151	(0.027)
Monotonic σ					
Exponential	θ	2.427	(0.417)	2.603	(0.279)
Mixture	θ_1	0.003	(0.001)	0.003	(0.004)
	θ_2	15.750	(1.648)	8.927	(0.502)
	α	0.336	(0.032)	0.240	(0.026)

As shown in the final rows of each panel in Table 3, these estimated parameters yield median population search costs of between \$4.44–\$6.42 for Amemiya \$1.80–\$4.48 for Cameron and Trivedi, depending on which specification and assumptions are considered. In both cases, these are significantly lower than the typical within-period median search cost, which are \$11.10 and \$8.58 for the two books, 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 within-period distribution of search costs of the current consumers over time and the population distribution of search costs. Otherwise, the magnitude of search costs can be dramatically overstated. The distributions of search costs of consumers who are active in the market do indeed appear to differ substantively from the population distribution of search costs of newly entering consumers.



(a) Exponential



(b) Exponential Mixture

Figure 4: Parametric Estimates of Population CDF (H)

Based on our simulation results in Section 4, we can infer some of the characteristics of the consumers' purchase policy function, σ_t. The purchase probabilities are likely not monotonically increasing in cost as this would cause the period medians to be biased downwards (as opposed to the observed upward bias) relative to the population median. It also does not appear that the purchase policy is constant (independent of c), since in this case we would expect to see the period quantiles converging to the quantiles of the population distribution. Hence, there is evidence that purchase probabilities in these markets are larger for smaller search costs, which leads consumers with high search costs to stay in the market relatively longer. This leads to dynamic selection which in turn leads to the observed persistent upward bias in the within-period search cost distributions relative to the population distribution. As a result, the monotonicity restriction seems largely valid and the qualitative findings are similar to the unrestricted results for both books.

7. Counterfactuals

This section reports the results of a series of counterfactual simulations. Using the results from our empirical application in Section 6 as a benchmark, we designed three counterfactual scenarios to study different components of the market. We first examine the influence of the search cost distribution G_t on the price distribution F_t within a single period. Next, we consider the interplay between market structure and the population search cost distribution H. Finally, we look at the long-run effects of changing the population search cost distribution H on the sequence of within-period distributions G_t.

7.1. Within-Period Search Costs (G_t) and Prices (F_t)

Firms optimize prices in each period in response to the within-period search cost distribution, G_t, that they face in that particular period. Therefore, the current within-period search cost distribution, G_t, will influence the within-period price distribution, F_t. To illustrate the effect that changes in G_t have on F_t, we designed a counterfactual where the search cost distribution of a randomly selected period of our data ($t = 36$) is uniformly increased and decreased. Specifically, we scale search costs up or down by 10%.

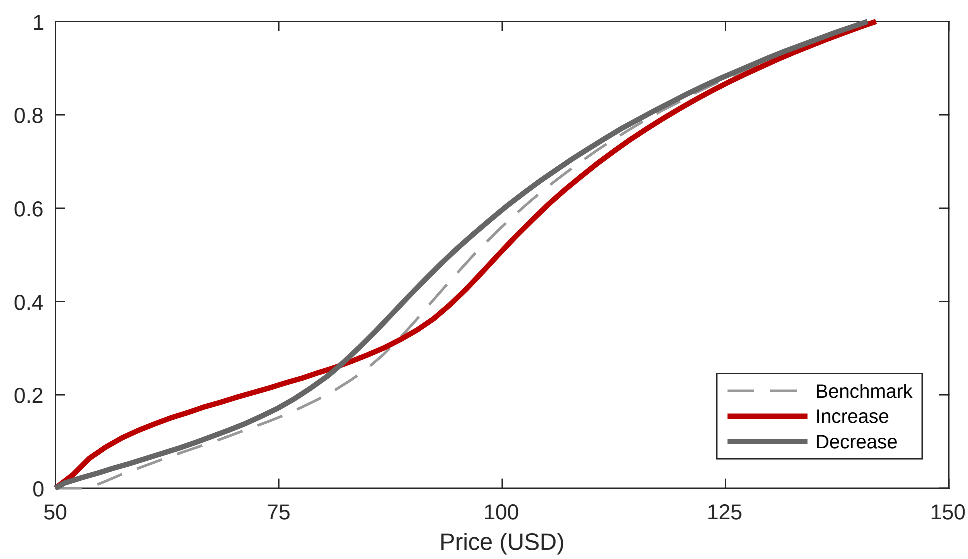


Figure 5: Effects of Within-period Search Costs on Prices

As shown by Moraga-González et al. (2017), the price effects of search cost changes are nonlinear and depend on the shape of the distribution. Figure 5 displays the observed price distribution (Benchmark) along with the resulting price distributions following an increase or decrease. When search costs decrease by 10%, they become relatively more concentrated around lower values, meaning that consumers will search more intensely and firms will lower prices. In this simulation, the price distribution arising from reduced search costs is first-order stochastically dominated by the benchmark distribution. The median and mean prices fall by \$2.62 and \$1.93, respectively.

Interestingly, the price effects are asymmetric when search costs are increased by 10%. In this case, the search cost distribution becomes more dispersed as probability mass moves out to the tail. This causes low prices to occur with higher probability, but also leads to the average and median prices increasing by \$1.11 and \$2.74, respectively.

The observed nonlinearity arises because of an interplay between the Diamond paradox and the Varian effect. As the heterogeneity of search costs in the right tail decreases, more consumers stay uninformed and therefore the average price increases, as this leads firms to gather around the monopoly price (Diamond paradox). However, as competition for fully informed consumers lessens, it becomes more likely that a low priced firm will capture all the informed consumers (Varian effect). This incentive leads to the increase in the probability of firms drawing a low price from the equilibrium price distribution.

7.2. Market Structure the Population Search Cost Distribution H

As demonstrated in the previous section, the within-period search cost distribution determines how firms price each period; however, the within-period distributions are strongly influenced by the underlying population search cost distribution. Therefore, in this counterfactual we investigate what the population distribution needs to be to be able to support different the price distributions over time. To change the price distribution, we alter the composition of firms (Major, Verified, Non-verified) in a systematic way, and then rerun our procedure to estimate H. These changes will have nonlinear effects and depend on the structure of search costs due to the tension between the Diamond paradox effect and the Varian effect.

We consider six scenarios in which we exogenously introduce additional firms of a certain type with a certain pricing rule. In each scenario we add either major retailers (High quality) or non-verified firms (Low quality) who set prices similar to either the highest (Max), average (Avg), or lowest (Min) price firms in their respective group.[23] We refer to these scenarios by their quality-pricing parameters, for example "High-Max" or "Low-Avg".

When adding a firm, it is important how the mean and variance of the price distribution changes as a result, since different search cost structures are required to support different price distributions. The six scenarios we consider cover all four combinations of an

[23] The prices set by the new firms are equal to quality-specific moments of the original price distribution in each period with the addition of a random $U(-2,2)$ perturbation.

increase/decrease in the mean price interacted with an increase/decrease in the variance of prices. Table 5 displays the percentage change in the mean, median, and variance of H relative to the distribution estimated in the application in Section 6, with the exponential mixture specification.

By changing the composition of firms, the price dispersion and average price change. When price dispersion decreases, we find that population search costs would need to become more homogeneous to support that. Additionally, increases in average price generally would require an increase in search cost dispersion. However, variance seems to be the dominant force in determining the within-period price distribution. This can be seen by examining the results for our six scenarios grouped by the implied mean/variance changes, which are shown in columns three and four of Table 5.

The results are unambiguous when the average and variance of price move in the same direction. When more high-price, low-quality firms are added to the market (Low-Max), the mean price and variance both increase. To support this larger and more dispersed price distribution, the market needs to have not only more consumers with larger search costs, but also must have more dispersion in search costs across consumers. Therefore, the mean and variance of H needs to be larger. However, when more average-price, high-quality firms or low-price, low-quality firms enter the market (High-Avg or Low-Min), the mean and variance of prices both decrease, since the price distribution becomes more concentrated around lower prices. To support this smaller and more compressed price distribution, the market needs less dispersion in search costs with consumers concentrated around smaller search costs. Therefore, the mean and variance of H need to be smaller. The addition of more low-price, low-quality firms causes a larger decrease in the average and variance of within-period prices, so the reduction in the distribution of H for the scenario by necessity is larger.

On the other hand, when the mean and variance of prices move in opposite directions, the component effects of each factor can be examined. When more average-price, low-quality firms enter the market (Low-Avg), the variance of the price distribution decreases while the mean increases. To support this, the market needs a large increase in the number of consumers with both low and high search costs; however, the distribution will be positively skewed with more mass concentrated around smaller search costs values. Conversely, when more high- or low-price, high-quality firms are added to the marketplace, the variance of the price distribution increases while the mean decreases. In other words, there will be more mass around lower prices, but the spread will increase. In this case, the market needs more consumers with larger search costs because of the increased price dispersion, but the difference in prices is smaller so the distribution needs less spread in the left tail to support the increased dispersion at low prices relative to when the average

Table 5: Percent Change in the Parameters of H

Scenario		Prices		Amemiya H			Cameron and Trivedi H		
Type	Price	Mean	Var	Mean	Median	Var	Mean	Median	Var
High	Max	-	+	0.51	4.98	1.06	0.92	4.46	1.56
High	Avg	-	-	-3.97	-2.49	-7.76	-8.43	-0.89	-2.95
High	Min	-	+	7.03	5.76	14.59	11.70	2.01	7.24
Low	Max	+	+	20.86	17.76	46.10	24.14	16.07	4.92
Low	Avg	+	-	-3.93	-1.25	-7.70	-4.87	-0.45	-17.15
Low	Min	-	-	-40.87	-80.06	-66.28	-10.71	-56.47	-20.29

price also increases.

7.3. The Influence of H on the Evolution of G_t

Our model implies that the current, within-period search cost distribution is a weighted mixture of last period's distribution and the population distribution. Therefore, we can measure how changes in the population distribution H influence the evolution of the sequence of period search cost distributions $\{G_t\}$. To investigate this, we consider a counterfactual where we adjust the population distribution H and measure how the within period distributions G_t change over time. We assume that H and G_t are exponential distributions using the parameter values estimated in Section 6.[24] Due to analytical and computational considerations, the purchase policy function σ_t here is assumed to be constant, $\sigma_t = \sigma$ for all t, and have an exponential functional form. In this case, we can easily determine the parameter of σ by solving our structural model using two adjacent, randomly selected periods from our data.

To analyze the effect of changes in the population distribution H, we perturb the estimated parameter and simulate the model for $T = 10,000$ periods. The results are shown in Table 6. The first column lists the perturbation amount δ, in dollars. Recall that the parameter θ is the mean of the distribution and θ^2 is the variance. The remaining columns show the percentage changes in the long-term period search cost distribution G under the perturbed population distribution H (with parameter $\hat{\theta} \pm \delta$), relative to H. The first row gives the baseline case, with no perturbation (i.e., $\delta = 0$). For example, the benchmark mean of the long run G distribution for Amemiya is 1.68% higher than the mean of H. After decreasing θ by one dollar (a 15.2% decrease), the mean of G falls to 14.20% below the mean of H. Given the results above regarding the effects of G_t on F_t,

[24] Gamma and Log Normal distributions were also tested and found to yield qualitatively similar results.

we can infer that, decreasing population search costs will lead to lower prices, even after accounting for the upward bias from dynamic selection effects.

Table 6: Percentage Change in Long-Term G Relative to H

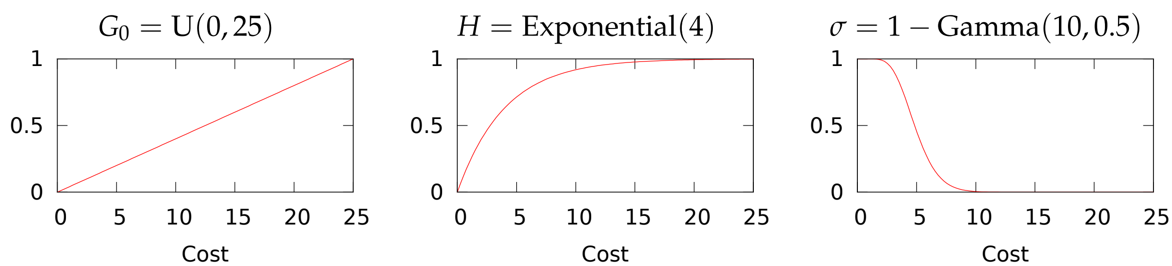
	Amemiya			
	$\theta + \delta$		$\theta - \delta$	
δ	Mean	Median	Mean	Median
0.00	1.68	3.30	1.68	3.30
0.50	9.62	11.37	-6.26	-4.76
1.00	17.56	19.44	-14.20	-12.83
2.00	33.45	35.58	-30.08	-28.97
3.00	49.33	51.71	-45.97	-45.10
	Cameron and Trivedi			
	$\theta + \delta$		$\theta - \delta$	
δ	Mean	Median	Mean	Median
0.00	1.81	3.43	1.81	3.43
0.50	10.89	11.84	-7.18	-7.47
1.00	17.06	20.31	-14.69	-15.34
2.00	34.65	33.40	-31.22	-31.30
3.00	50.28	51.23	-46.92	-44.78

8. Conclusion

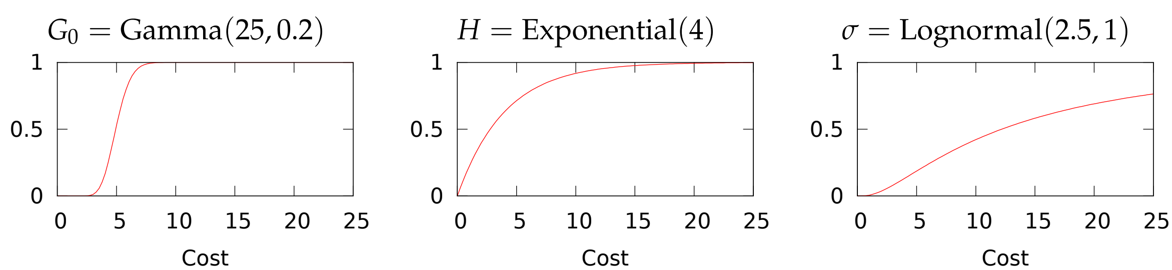
Thus far, the previous literature on structural estimation of consumer search costs has relied primarily on static models, particularly in models which are estimated using only price data. However, when consumers shop for a durable good over multiple periods and search is costly, selection effects may cause the distribution estimated using a static model to be biased relative to the actual population search cost distribution. This paper extends existing methods to estimate search cost distributions in a setting with forward-looking consumers, taking dynamic selection effects into account. As our application to online textbook markets and our simulation studies demonstrate, it is important to distinguish between the within-period and population distributions of search costs. Both may be of interest, and the methods we describe yield estimates of both without the need to fully specify the exact dynamic decision making process of consumers.

Table 7: Representative Simulation Specifications

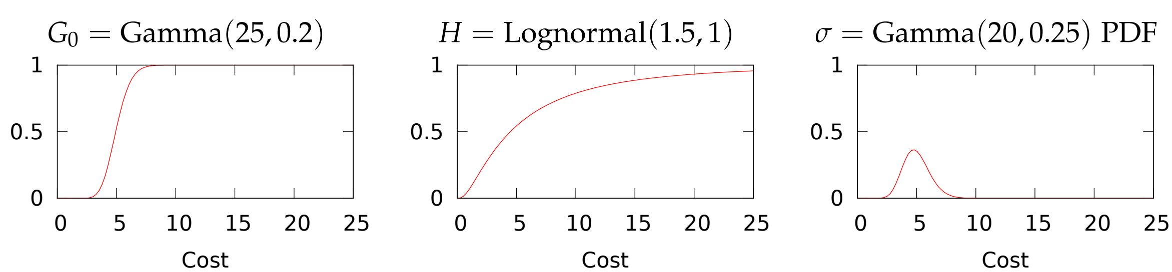
Initial Distribution (G_0) | Population Distribution (H) | Purchase Prob. (σ)



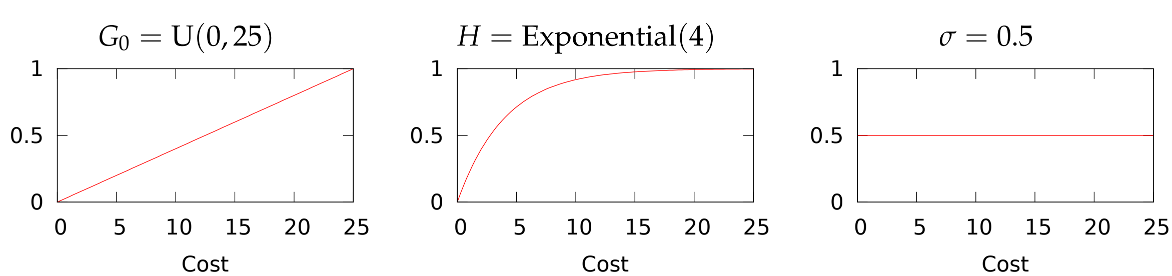
Specification 1: Positive Bias



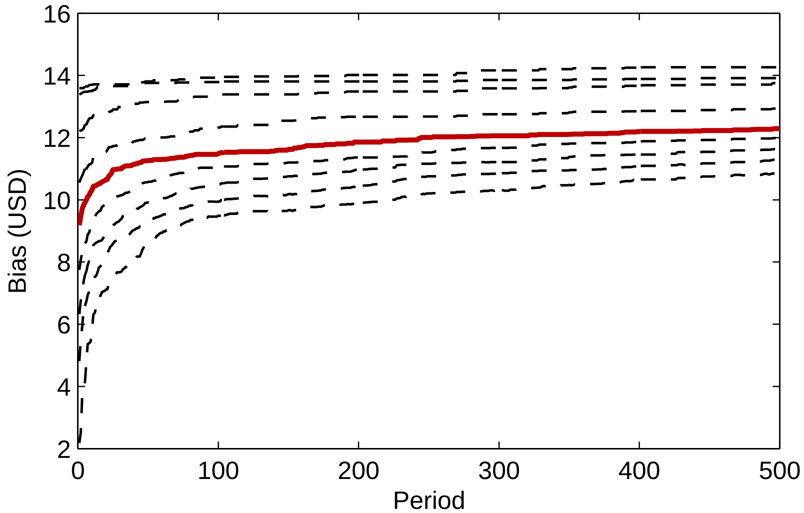
Specification 2: Negative Bias



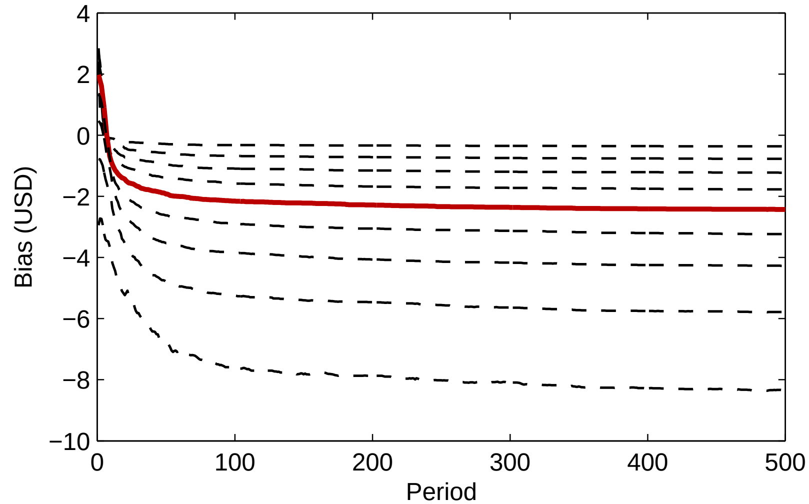
Specification 3: Both Positive and Negative Bias



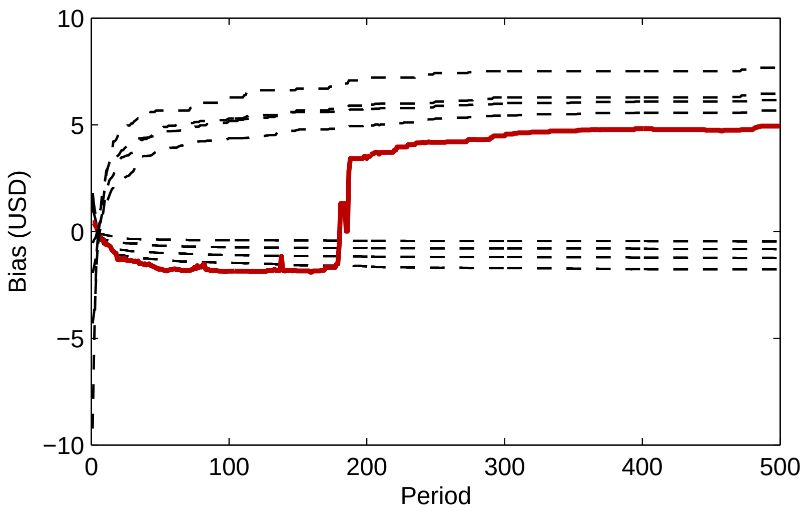
Specification 4: No Bias



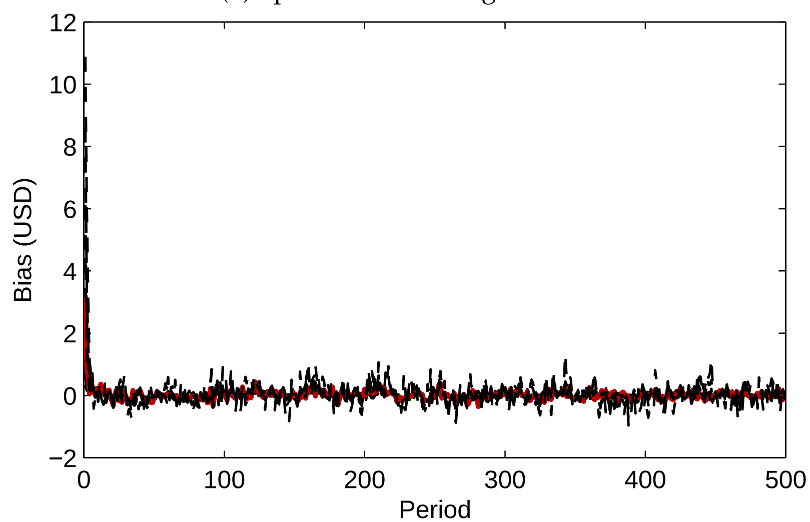
(a) Specification 1: Positive Bias



(b) Specification 2: Negative Bias



(c) Specification 3: Positive and Negative Bias



(d) Specification 4: No Bias

Figure 6: Decile Bias of Period Search Cost Distributions (G_t)

A. Simulation Details and Figures

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 purchase policy functions (σ) used were $\sigma(c) \in \{0.25, 0.5, 0.75\}$, 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 $\Gamma(10, 0.5)$, one minus the absolute value of $\sin(c)$, one minus the absolute value of $\cos(c)$, an increasing step function $\sigma(c) = 0, 0.5, 1$, respectively, when $c \leq 5, 5 < c < 20, 20 \leq c$, a decreasing step function $\sigma(c) = 1, 0.5, 0$ when $c \leq 5, 5 < c < 20, 20 \leq c$, one minus the PDF of $\Gamma(20, 0.25)$, one minus the triangular distribution, where $\sigma(c) = 1 - c/\beta$ for $c \leq \beta$ and $\sigma(c) = c/\beta - 1$ for $c > \beta$, and one minus the double triangular distribution, where $\sigma(c) = c/\beta$ for $c \leq \beta$ and $\sigma(c) = 2 - c/\beta$ for $c > \beta$, 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 (σ) from above were used to maintain internal comparability.

B. Proofs

Proof of Proposition 1

First, note that (5) is an integral equation of the form $G_{t+1}(c') = \int_0^\infty L_t(c, \sigma_t(c), \sigma_t'(c); c')\, dc$ where

$$L_t(c, \sigma_t(c), \sigma_t'(c); c') = \left[(1 - \sigma_t(c))\, 1\{c \leq c'\} + \sigma_t(c) H(c')\right] g_t(c).$$

We can apply techniques from the calculus of variations to find policy functions σ_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

$$L_{t,2}(c, \sigma_t(c), \sigma_t'(c); c') - \frac{d}{dc} L_{t,3}(c, \sigma_t(c), \sigma_t'(c); c') = 0$$

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

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

For all $c' > 0$, we have $H(c') > 0 > H(c') - 1$ and so we can see immediately that of functions in the set Σ, $\sigma_t^{\mathrm{U}}(c;c') = 1\{c > c'\}$ will maximize $G_{t+1}(c')$ while $\sigma_t^{\mathrm{L}}(c;c') = 1\{c \leq c'\}$ will minimize $G_{t+1}(c')$.

Given these bounding policy functions, we can simplify (5) 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 $\sigma_t^{\mathrm{L}}(c;c') = 1$ when $c < c'$. The second term 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 $\sigma_t^{\mathrm{U}}(c;c') = 0$ when $c > c'$ effectively only limits the domain of integration for the second term.

Proof of Proposition 2

To establish the first two parts of the Proposition, we will say that the bounds are informative when either $H_{t+1}^{\mathrm{L}}(c) > 0$ or $H_{t+1}^{\mathrm{U}}(c) < 1$ (or both). 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 $H(c)$.

To establish sharpness of the bounds for the value of $H(\cdot)$ evaluated a given search cost c, we consider a parametric purchase policy $\sigma(\cdot;c,\alpha)$ with $\alpha \in [0,1]$ and

$$\sigma(\tilde{c};c,\alpha) = \begin{cases} 1-\alpha & \text{if } \tilde{c} \leq c, \\ \alpha & \text{if } \tilde{c} > c. \end{cases}$$

For a given value of α, and hence for a given policy rule $\sigma(\cdot;c,\alpha)$, the implied value of $H(c)$ is defined by (5). Solving for the value of $H(c)$, which we will denote $h(c,\alpha)$, we find

$$h(c,\alpha) = \frac{G_{t+1}(c) - \alpha G_t(c)}{G_t(c) - 2\alpha G_t(c) + \alpha}.$$

To restrict our consideration to proper probability values between 0 and 1, define $\overline{h}(c,\alpha) \equiv \min\{\max\{h(c,\alpha),0\},1\}$. Note that for extreme values of α, the upper and lower bounds on $H(c)$, respectively, from Proposition 1 are achieved:

$$\overline{h}(c,0) = \min\left\{\frac{G_{t+1}(c)}{G_t(c)},1\right\} \qquad \text{and} \qquad h(c,1) = \max\left\{\frac{G_{t+1}(c) - G_t(c)}{1 - G_t(c)},0\right\}.$$

We now have that $\overline{h}(c,\cdot)$ is a continuous mapping from a connected interval $[0,1]$ to a connected interval $\left[H_{t+1}^{\mathrm{L}}(c), H_{t+1}^{\mathrm{U}}(c)\right]$ and therefore the mapping is surjective. In other words, for all possible values $H(c) \in \left[H_{t+1}^{\mathrm{L}}(c), H_{t+1}^{\mathrm{U}}(c)\right]$ within the bounds, there is a policy rule $\sigma_t(\cdot;c,\alpha)$, for some $\alpha \in [0,1]$, which rationalizes $H(c)$. Since all values in the interval are possible for some σ_t, the bounds are therefore sharp.

Proof of Proposition 3

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

For maximizing $G_{t+1}(c')$, assigning non-zero probability on $c > c'$ (corresponding to the positive part of the functional derivative) must counter-balance the negative effects of necessarily assigning non-zero probability on $c \leq c'$ (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 $c \leq c'$, so any maximizing function must be constant on $c \leq c'$ and equal to the value at c'. Let $\alpha \geq 0$ denote the constant value on $[0, c']$. 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 α on (c', ∞). Hence, the maximizing policy σ_t must be constant on $[0, \infty)$. Finally, the constant value of the maximizing function α 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 $\sigma_t = 0$, $G_{t+1}(c') = G_t(c')$ and for $\sigma_t = 1$, $G_{t+1}(c') = H(c')$. Hence, the maximizing weakly decreasing policy function is

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

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

$$G_{t+1}(c') \leq G_{t+1}^{\mathrm{U}}(c') \equiv \max\{G_t(c'), H(c')\}.$$

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 for $c' = c$: if $G_t(c) < G_{t+1}(c)$ then $G_t(c) < H(c)$.

Proof of Proposition 4

Suppose to the contrary that there exists a policy σ_t such that $1 - R_t(\sigma_t) > \inf_{c'} \frac{g_t(c')}{h(c')}$ for which there is no bias, $G_{t+1} = H$. From (5), for all c'

$$H(c') \int_0^\infty (1 - \sigma_t(c))\, g_t(c)\, dc - \int_0^{c'} (1 - \sigma_t(c))\, g_t(c)\, dc = 0.$$

Differentiating with respect to c' and rearranging yields a contradiction:

$$1 - R_t(\sigma_t) = (1 - \sigma_t(c')) \frac{g_t(c')}{h(c')} \leq \frac{g_t(c')}{h(c')}. \tag{11}$$

The second result follows by noting that if $G_t = G_{t+1} = H$, then $g_t = g_{t+1} = h$ and by (11) it must be the case that σ_t is constant.

Proof of Proposition 5

Let k_t and k_{t+1} be the period-specific indices such that $c_{t,k_t} = c_{t,k_{t+1}} = c$. By Moraga-González and Wildenbeest (2008), the maximum likelihood estimates of the within-period search cost cutoffs and CDF values $\{(\hat{c}_{t,k}, \hat{G}_t(c_{t,k})\}_{k=1}^{K}$ are consistent for both periods t and $t+1$, provided that N_t and N_{t+1} both tend to infinity.

We consider the upper bound first:

$$\hat{H}^{\mathrm{U}}_{t+1}(c) \equiv \min\left\{\frac{\hat{G}_{t+1}(c)}{\hat{G}_t(c)}, 1\right\}.$$

For the CDF values of interest, by consistency of the per-period estimates we have $\hat{G}_t(c_{t,k_t}) \xrightarrow{\mathrm{p}} G_t(c)$ and $\hat{G}_{t+1}(c_{t+1,k_{t+1}}) \xrightarrow{\mathrm{p}} G_{t+1}(c)$. By Slutsky's theorem, it follows that $\hat{G}_{t+1}(c_{t+1,k_{t+1}})/\hat{G}_t(c_{t,k_t}) \xrightarrow{\mathrm{p}} G_{t+1}(c)/G_t(c)$. Then since $\min\{x/y, 1\}$ is continuous in x and y, the continuous mapping theorem yields consistency of the upper bound for $H(c)$: $\hat{H}^{\mathrm{U}}_{t+1}(c) \xrightarrow{\mathrm{p}} H^{\mathrm{U}}_{t+1}(c)$.

Consistency of the lower bound follows in a similar way, recalling that

$$\hat{H}^{\mathrm{L}}_{t+1}(c) \equiv \max\left\{\frac{\hat{G}_{t+1}(c) - \hat{G}_t(c)}{1 - \hat{G}_t(c)}, 0\right\}.$$

The numerator and denominator, respectively, converge in probability to $G_{t+1}(c) - G_t(c)$ and $1 - G_t(c)$. By continuity of $\max\{\cdot, 0\}$, it follows that $\hat{H}^{\mathrm{L}}_{t+1}(c) \xrightarrow{\mathrm{p}} H^{\mathrm{L}}_{t+1}(c)$.

Finally, convergence of the Hausdorff distance follows from consistency of the endpoints. In general, for two sets A and B, $d_H(A, B) = \max\left\{\sup_{\theta' \in B} \inf_{\theta \in A} d(\theta, \theta'), \sup_{\theta' \in A} \inf_{\theta \in B} d(\theta', \theta)\right\}.$

$$\begin{aligned} &d_{\mathrm{H}}([\hat{H}^{\mathrm{L}}_{t+1}(c), \hat{H}^{\mathrm{U}}_{t+1}(c)], [H^{\mathrm{L}}_{t+1}(c), H^{\mathrm{U}}_{t+1}(c)]) \\ &\qquad = \max\left\{\left|\hat{H}^{\mathrm{L}}_{t+1}(c) - H^{\mathrm{L}}_{t+1}(c)\right|, \left|\hat{H}^{\mathrm{U}}_{t+1}(c) - H^{\mathrm{U}}_{t+1}(c)\right|\right\} \xrightarrow{\mathrm{p}} 0. \end{aligned}$$

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.

Armstrong, M., J. Vickers, and J. Zhou (2009). Prominence and consumer search. *RAND Journal of Economics 40*, 209–233.

Ashford, K. (2009). For a $27.95 bestseller. *Money 38*.

Bailey, J. (1998). Electronic commerce: Prices and consumer issues fro 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.

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.

Chow, G. C. (2011). Usefulness of adaptive and rational expectations in economics. CEPS Working Paper 221, Princeton University.

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. (2018). Consumer search on the Internet. *International Journal of Industrial Organization 58*, 66–105.

De los Santos, B., A. Hortaçsu, 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. Hortaçsu, and M. Wildenbeest (2017). Search with learning for differentiated products: Evidence from e-commerce. *Journal of Business & Economic Statistics 35*, 626–641.

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.

Gautier, P., J. L. Moraga-Gonzalez, and R. Wolthoff (2016). Search costs and efficiency: Do unemployed workers search enough? *European Economic Review 84*, 123–139.

Gowrisankaran, G. and M. Rysman (2012). Dynamics of consumer demand for new durable goods. *Journal of Political Economy 120*, 1173–1219.

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.

Honka, E. (2014). Quantifying search and switching costs in the US auto insurance industry. *RAND Journal of Economics 45*(4), 847–884.

Honka, E. and P. Chintagunta (2017). Simultaneous or sequential? Search strategies in the US auto insurance industry. *Marketing Science 36*, 21–42.

Hortaçsu, A. and C. Syverson (2004). 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*(2), 403–456.

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), 299–308.

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

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.

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.

Melnikov, O. (2013). Demand for differentiated durable products: The case of the U.S. computer printer market. *Economic Inquiry 51*, 1277–1298.

Moraga-González, J. L. (2006). Estimation of search costs. Lecture, Economics Network for Competition and Regulation Workshop on "Consumers and Competition", University of Amsterdam.

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

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., Z. Sandor, and M. Wildenbeest (2017). Nonsequential search equilibrium with search cost heterogeneity. *International Journal of Industrial Organization 50*, 392–414.

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.

National Association of College Stores (2009). Industry statistics: 2015/2016 flash report.

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.

Risk Management Association (2016). Retail bookstore business insights.

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.

Sanches, F., D. Silva Junior, and S. Srisuma (2017). Minimum distance estimation of search costs using price distribution. *Journal of Business & Economic Statistics*. Forthcoming.

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.