Identification and Estimation of a Search Model: a Procurement Auction Approach*†

Mateusz Myśliwski — NHH
Fabio Sanches — Insper
Daniel Silva Jr — City, University of London

Sorawoot Srisuma — NUS and University of Surrey

1 March 2021

Abstract

We propose a model of nonsequential consumer search where consumers and firms differ in search and production costs respectively. We characterize the equilibrium of the game. We nonparametrically identify the model using market shares and prices. Our strategy to identify firm's cost distribution is similar to Guerre, Perrigne and Vuong (2000) as the price setting problem for firms resembles that of a procurement auction. Our estimator for the firm's cost pdf has the same convergence rate as the optimal rate in Guerre et al. uniformly over any fixed subset in the interior of the support, and it can converge arbitrarily close to that rate when the subset increases to the full support asymptotically. The difference in rates is due to the possible presence of a pole in the price pdf, which we show to be a common feature in equilibrium. Our simulation shows estimators that ignore the pole can have substantial bias. Our results apply to search models with homogenous or vertically differentiated products.

JEL Classification Numbers: C14, C57, D83

Keywords: Auctions, Kernel Smoothing, Nonparametric Identification, Search Costs

---

*We thank Jaap Abbring, Guiherme Carmona, Hanming Fang, Alessandro Gavazza, Matt Gentry, Emmanuel Guerre, Kohei Kawaguchi, Tatiana Komarova, Nianqing Liu, Jose Luis Moraga-González, Lars Nesheim, Joris Pinkse (discussant), Fabien Postel-Vinay, Philip Reny, Xiaoxia Shi, Mikkel Sølvsten, Pai Xu and seminar participants at Academia Sinica, Cardiff University, Durham University, Hitotsubashi University, LSE, SUFE, TSE, UCL, University of Groningen, University of Hong Kong, University of Wisconsin–Madison (virtual) and Conference on "Auctions, competition, regulation, and public policy" in Lancaster for comments and discussions.

†*E-mails*: mateusz.mysliwski@nhh.no; fmiessi@gmail.com; danielsjunior@gmail.com; s.srisuma@surrey.ac.uk

1 Introduction

It is a common occurrence that seemingly homogeneous goods and services are available for sale at different prices. If the market is competitive, the discrepancy from "the law of one price" indicates some frictions or other forms of inefficiencies may exist. One classic explanation of this phenomenon attributes the existence of price dispersion to consumer search costs Stigler (1961). Many theoretical search models have been proposed to show that price dispersion can arise in equilibrium; with the earlier theoretical literature focusing on models with minimal or even no ex-ante heterogeneity. The survey of Baye, Morgan, and Scholten (2006) gives an account as well as examples of other sources of price dispersion.

An influential paper by Burdett and Judd (1983) shows that a continuous pricing rule can be generated by a mixed strategy Nash equilibrium in a *fixed sample*[1] search model with complete information consisting of infinitely many identical firms and consumers. There, firms are identical in the sense that they are known to have the same (constant) marginal cost of production and all consumers have the same search cost. Hong and Shum (2006) develop an empirical model based on Burdett and Judd (1983) by assuming that consumers draw search costs from some continuous distribution. They show, using just data on prices, nonparametric identification of the firms' marginal costs and parts of the distribution of consumer search costs. Their strategy can also be used to identify an analogous model with finite number of firms (Moraga-González and Wildenbeest (2008)), i.e. an oligopolistic setup, which is commonly assumed in practice. Identification of the search distribution in these papers is only partial as parts of the support of search cost cannot be identified. Moraga-González, Sandor and Wildenbeest (2013) show identification on the full support is possible if additional price data from other equilibria are available.

In this paper we propose an empirical model of fixed sample search that allows for heterogeneity across firms as well as consumers. We assume there are a finite number of firms who draw marginal costs from some continuous distribution. Costs are private and firms compete in prices in an *incomplete information* environment. We analyze both the theoretical and empirical aspects of this model. We make three main contributions:

(i) Provide a system of equations that characterize non-degenerate pure strategy Bayesian-Nash equilibria (BNE) in the model via a fixed-point;

[1] In a fixed sample (or nonsequential) search consumers decide from the onset how many price quotes to search for. This stands in contrast to sequential search. The two models are not nested. Morgan and Manning (1985) show the fixed and sequential search models can be optimal in different circumstances. The fixed search model may be more suitable, for example, in applications where time is a factor so that buyers prefer to gather information quickly. Some recent empirical studies found that nonsequential search models provide a better approximation to consumers' search behavior observed in real life (De Los Santos et al. (2012), Honka and Chintagunta (2017)).

(ii) Show both the marginal cost distribution of firms and search cost distribution of consumers can be nonparametrically identified from data on price and market share;

(iii) Propose nonparametric estimators for the primitives that achieve optimal convergence rate uniformly on any fixed subset in the interior of the support and the rate can be made arbitrarily close to the optimal rate when the subset expands asymptotically to the whole support.

For the ease of notation and clarity of idea, the paper focuses on the symmetric model where products are homogeneous. We provide a generalization of this to a model of vertical differentiated products as an extension. More specifically, in the context of a complete information game, Wildenbeest (2011) incorporates vertical product differentiation components into the model in order to explain systematic price differences between firms. He shows identification is still possible from just prices by assuming the *valuation-cost markup*, which is defined as the difference between quality and production cost, is the same for all firms, so that firms pass on the costs associated with quality to the consumers in equilibrium. A natural extension to his idea for an incomplete information game is to assume the cost distributions for all firms have the same shape but vary in location. In this setting, we show our results developed for the symmetric model can then be readily generalized when we restrict the pricing strategies for firms to be affine transformations of each other.

There are at least two motivations for incorporating heterogeneity amongst firms. First, the perfect competition assumption on the supply side – i.e. a common knowledge of identical production costs – may not be suitable for applications with a small number of firms. Another reason is that seemingly identical products may in reality differ in an unobservable way to the consumers but those differences were observed by the seller when she chose the price.[2] In this case, a model with heterogeneous costs can be interpreted as taking an agnostic approach on product differentiation to approximate situations where differentiation may occur in less transparent or unobservable way.

Our model is inspired by the theoretical work by MacMinn (1980) and we generalize his framework.[3] MacMinn gives one of the earliest account of a search model where firms' best responses can generate price dispersion. He assumes firms differ by drawing marginal costs from a uniform distribution. His result is a partial equilibrium because consumer search behavior is exogenously assumed and every consumer searches the same number of times (cf. Pereira (2005)). In contrast, our BNE is defined by simultaneous best responses for both consumers and firms. We do not specify any distributional assumption on the marginal costs and allow different consumers to search for a

[2] For example, online sellers of second-hand books or music records (e.g. on Amazon Marketplace or Discogs.com) tend to possess private information about the actual condition of the item that goes beyond the description provided in the offer because they physically own the product.

[3] Similarly to us, Benabou (1993) considers a search model with bilateral heterogeneity. He does not, however, have any results on nonsequential search.

different number of prices. Therefore the decision problem for firms in our model resembles a first-price procurement auction where each bidder has to form an expectation on the number and identity of her competitors. The similarities between search and auction models have been well documented in the theoretical literature, e.g. see McAfee and McMillan (1998) in a mechanism design context. Other applications include some job search models from the labor literature.[4]

Our identification strategy is different to what is used to identify complete information models. In particular, Hong and Shum (2006) use price distribution to identify consumers' search distribution and the firms' common marginal cost by exploiting the constant profit condition imposed by the mixed strategy condition. We propose to identify an incomplete information model with market shares of firms in addition to prices. This idea is analogous to linking market shares to choice probabilities, which is the starting point for the identification argument used in the literature on on demand for differentiated products (see Berry and Haile (2014)). We show that market shares generally relate to the equilibrium proportions of consumer search linearly in expectation conditional on price, so that the distribution of consumer search costs can be recovered from solving a linear equation. We then use it to help identify the marginal cost distribution for firms. Our approach here is similar to how Guerre, Perrigne and Vuong (2000, hereafter GPV) identify the distribution of the bidder's latent valuation in a first-price auction model. In particular, we derive the inverse of the equilibrium pricing function explicitly and use it to recover firms' latent costs from observed prices.

Our identification results are constructive. We show the parameters of interest can be written explicitly in terms of the joint distribution of observed variables. Subsequently we propose to estimate them using sample counterparts that does not require any numerical optimization. On the demand (consumers') side the consumer optimal behavior is characterized by the distribution of frequency of search that is a solution to a linear least squares problem where the regressors are written in terms of the cumulative density function (cdf) of price. It can then be computed using an OLS closed-form expression. On the supply (firms') side, we estimate the firms' latent costs probability density function (pdf) in two steps. First we invert observed prices into costs. Then we use them as generated variables to estimate a nonparametric pdf estimator.

The equilibrium with price dispersion has an interesting feature. Our analysis reveals that the price pdf generally has a *pole* at the upper support, i.e. it asymptotes to infinity at that point.[5]

[4] Well-known labor applications include Postel-Vinay and Robin (2002) and Cahuc, Postel-Vinay and Robin (2006) who model on-the-job search as a sequential auction over the worker between the current and prospective employer. A job search model that is closer to ours is the work by Bontemps, Robin and van den Berg (1999) as they allow for heterogeneous opportunity costs of keeping jobs among workers and continuous productivity among firms.

[5] While the pdf of optimal bids does not have this feature in a standard first price auction, it is present when there is a binding reserved price (see Section 4 in GPV). Unbounded densities also exist in some job search models. E.g. Bontemps, Robin and van den Berg (2000) find it for the wage distribution around the minimum wage.

Intuitively this happens because there are consumers who search only one time and will pay whatever price the firm charges up to their valuation of the good. Correspondingly the firm has an incentive to charge close to that price. This feature has a theoretical implication because the optimal convergence rate for the cost pdf for firms is determined by the convergence rates of the price pdf, and the convergence rate for nonparametric estimators in the neighborhood of the pole is slower than at an interior point. We show that a kernel density estimator for the firm's cost pdf attains the same optimal convergence rate as the GPV estimator on any fixed subset in the interior of the support, and their convergence rate over a suitably expanding support that increases to the full support asymptotically can be made arbitrarily close to the same optimal rate as on the interior after a transformation is made. Our estimator that accounts for the pole relies on a transformation that is effective without specifying the rate of divergence at the pole. We show using a simulation that not accounting for the pole can lead to substantial bias in finite sample.

Recent empirical papers that employ models closest to ours can be found in Salz (2020) and Myśliwski and Rostom (2020). The former, which is an independent work that precedes ours chronologically, proposes a search model to study the trade waste market in New York City. The latter extends our model and use it to study the UK's mortgage market. The common motivation for both of these papers is to study the role of intermediaries in markets with heterogeneous firms and consumers. Specifically, in additional to searching, consumers in their papers can choose to purchase a broker. The broker acts as a clearinghouse where a procurement auction game is played. Our model, which is a pure search model with no intermediary, is a special case of theirs. Importantly, however, Salz assumes a broker always exists so that he can directly identify the firm's cost distribution from the brokers independently of the search mechanism.[6] Identification of the consumer search subsequently relies on this. Therefore, as an econometric problem, ours and Salz's are different and not nested. On the other hand, no physical auction actually takes place in Myśliwski and Rostom (2020). They assume firms (mortgage brokers) search on behalf of consumers at a fee as a theoretical construct; their identification and estimation strategies are based on the results in this paper. These two approaches complement each other as data on intermediaries or shares may not be available in some situations.

We organize the rest of the paper as follows. Section 2 presents the model and characterizes the equilibrium of the game. We give identification results in Sections 3 and show how they lead to estimators with desirable properties in Section 4. Section 5 extends our search model and results to accommodate product differentiation. Section 6 contains a Monte Carlo exercise to illustrate theoretical features of the model. Section 7 concludes. The proofs of all results not given in the main

[6] Salz assumes there are two types of firms (carters). A H(igh) and L(ow) cost types. Both types are present in both the broker and search markets, and they are allowed to bid differently.

text can be found in the Appendix.

2 Model

We consider a model in which there is a unit mass of consumers and a finite number of firms. Each consumer has an inelastic demand for a single unit of a good supplied by the firms. Consumers differ by search costs. They have a belief on the price distribution and employ a nonsequential search strategy to decide on the number of firms to visit and purchase at the lowest price. Firms differ by production costs. They form beliefs about consumer search behavior and competing firms' pricing strategies, and set their price to maximize expected profits.

The primitives of our search model are $\{G(\cdot), H(\cdot)\}$, which respectively represent the search cost cdf and production cost cdf. The number of firms, denoted by I, is finite and known. We model consumers in the same way as Moraga-González and Wildenbeest (2008), Moraga-González, Sandor and Wildenbeest (2013), and Sanches, Silva and Srisuma (2018), which only differs from Hong and Shum (2006) in that I is infinite in the latter. We describe the decision problem and the best response for the consumers in Section 2.1. The aforementioned papers assume firms have the same production cost and that is common knowledge. We assume costs differ across firms and they are private information. We describe the firms' decision problem and derive their best response in Section 2.2. We define the equilibrium of our game in Section 2.3.

2.1 Consumers

All consumers have the same valuation of the object at some finite and positive $\overline{P}$. Each consumer draws a search cost c, which is assumed to be a continuous random variable support on $\left[0, \overline{C}\right] \subset \mathbb{R}^{+}$ with cdf $G(\cdot)$. A consumer with search cost c faces the following decision problem:

$$\min_{1 \leq k \leq I} c(k-1) + \mathbb{E}_F\left[P_{(1:k)}\right].$$

The first search is free and a purchase is always made. We use $P_{(k:k')}$ to denote the $k-$th order statistic from k' i.i.d. random variables of prices with some arbitrary distribution; $P_{(1:k)}$ denotes the minimum of such k prices. The game is symmetric as all firms have equal probability of being found. We use $\mathbb{E}_F[\cdot]$ to denote an expectation where the random prices have distribution described by the cdf $F(\cdot)$.

Consumer's Best Response

The marginal saving from searching one more firm after having searched k firms is:

$$\Delta_k(F) \equiv \mathbb{E}_F\left[P_{(1:k)}\right] - \mathbb{E}_F\left[P_{(1:k+1)}\right]. \tag{1}$$

$\Delta_k(F)$ is non-increasing in k because $\mathbb{E}_F\left[P_{(1:k)}\right]$ is non-increasing in k. We define $\Delta_I(F)$ to be 0. When price has a continuous distribution, $\mathbb{E}_F\left[P_{(1:k)}\right]$ is strictly increasing and

$$\Delta_k(F) = \int F(p)(1-F(p))^k\, dp. \tag{2}$$

The optimal behavior for a consumer that draws $c > \Delta_1(F)$ is to search once and search k times for $1 < k \leq I$ if $c \in [\Delta_k(F), \Delta_{k-1}(F))$. For the purpose of defining equilibrium (see below), we can state the best response for consumers in terms of proportions of consumer search. In what follows, we use $\mathcal{F}$ to denote a set of all price cdfs and $\mathbb{S}^{I-1}$ to denote a unit simplex in $\mathbb{R}^{I+}$.

LEMMA 1. *A consumer's best response is a map $\sigma_D : \mathcal{F} \to \mathbb{S}^{I-1}$ such that for any F in $\mathcal{F}$,*

$$\sigma_D(F) = \begin{cases} 1 - G(\Delta_k(F)) & \text{for } k = 1 \\ G(\Delta_{k-1}(F)) - G(\Delta_k(F)) & \text{for } 1 < k \leq I \end{cases}. \tag{3}$$

where $\{\Delta_k\}_{k=1}^{I-1}$ is defined in (1) and $\Delta_I := 0$.

2.2 Firms

Firm i draws a marginal cost of production R_i. R_i is assumed to be a continuous random variable supported on $\left[\underline{R}, \overline{R}\right] \subset \mathbb{R}^+$ with cdf $H(\cdot)$ where $\overline{R}$ is finite. Firm costs are private information that are independent from each other. Under symmetry, firm i then faces the following decision problem:

$$\max_p \Lambda(p, R_i; \mathbf{q}), \text{ where}$$

$$\Lambda(p, R_i; \mathbf{q}) \quad = \quad (p - R_i) \sum_{k=1}^{I} q_k \frac{k}{I} \mathbb{P}\left[P_{(1:k-1)} > p\right].$$

Here $\mathbf{q} = (q_1, \ldots, q_I)^\top$ denotes a vector in $\mathbb{S}^{I-1}$, where q_k denotes the proportion of consumers searching k firms. The term $\frac{k}{I}$ is the probability that firm i gets included when k firms are sampled.[7] The standard first-price procurement auction can be seen as a special case of the game firms in our model play when $q_I = 1$.

[7] Let $\mathcal{C}_k^I \equiv \frac{I!}{(I-k)!k!}$ be the combinatorial number from choosing k objects from a set of I. Then $\mathcal{C}_{k-1}^{I-1}/\mathcal{C}_k^I = \frac{k}{I}$.

Firm's Best Response

We consider a pricing strategy $\beta : \left[\underline{R}, \overline{R}\right] \to \left[\underline{P}, \overline{P}\right] \subset \mathbb{R}$ that is strictly increasing almost everywhere and satisfies $\beta\left(\overline{R}\right) = \overline{R}$. The latter is a free entry condition. We assume $\overline{R} = \overline{P}$, so that firms always produce and a purchase is always made. For any $\mathbf{q} \in \mathbb{S}^{I-1}$, we can define $\Lambda^*\left(\cdot;\mathbf{q}\right)$ to be the value function for a representative firm when all players are assumed to employ a strictly increasing optimal pricing strategy that we denote by $\beta\left(\cdot;\mathbf{q}\right)$. We denote its inverse, $\beta^{-1}\left(\cdot;\mathbf{q}\right)$ by $\xi\left(\cdot;\mathbf{q}\right)$.

$$\Lambda^*\left(r;\mathbf{q}\right) = \left(\beta\left(r;\mathbf{q}\right) - r\right)\sum_{k=1}^{I} q_k \frac{k}{I}\left(1 - H\left(\xi\left(\beta\left(r;\mathbf{q}\right);\mathbf{q}\right)\right)\right)^{k-1}.$$

Then by the envelope theorem (Milgrom and Segal (2002)),

$$\begin{aligned}
\left.\frac{d}{dr}\Lambda^*\left(r;\mathbf{q}\right)\right|_{r=R} &= -\sum_{k=1}^{I} q_k \frac{k}{I}\left(1 - H\left(R\right)\right)^{k-1}, \text{ and} \\
\Lambda^*\left(\overline{R};\mathbf{q}\right) - \Lambda^*\left(R;\mathbf{q}\right) &= -\sum_{k=1}^{I} q_k \frac{k}{I}\int_{s=R}^{\overline{R}}\left(1 - H\left(s\right)\right)^{k-1} ds.
\end{aligned}$$

Solving this gives the solution of the firm's maximization problem, where for all r:

$$\beta\left(r;\mathbf{q}\right) = r + \frac{\sum_{k=1}^{I} q_k k \int_{s=r}^{\overline{R}}\left(1 - H\left(s\right)\right)^{k-1} ds}{\sum_{k=1}^{I} q_k k\left(1 - H\left(r\right)\right)^{k-1}}. \tag{4}$$

Suppose $H\left(\cdot\right)$ is differentiable and let $h\left(\cdot\right)$ denote the pdf of R_i. Differentiating the expression above gives,

$$\beta'\left(r;\mathbf{q}\right) = \frac{h\left(r\right)\left(\sum_{k=2}^{I} q_k k\left(k-1\right)\left(1 - H\left(r\right)\right)^{k-2}\right)\left(\sum_{k=1}^{I} q_k k \int_{s=r}^{\overline{R}}\left(1 - H\left(s\right)\right)^{k-1} ds\right)}{\left(\sum_{k=1}^{I} q_k k\left(1 - H\left(r\right)\right)^{k-1}\right)^2}. \tag{5}$$

It is clear that $\beta\left(\cdot;\mathbf{q}\right)$ is continuous and non-decreasing on $\left[\underline{R}, \overline{R}\right]$ as well as satisfying $\beta\left(\overline{R}\right) = \overline{R}$. In particular, note that $\beta'\left(\cdot;\mathbf{q}\right)$ is non-negative on $\left[\underline{R}, \overline{R}\right]$ and is finite on $\left[\underline{R}, \overline{R}\right)$. Furthermore, if $q_1 = 1$ then $\beta\left(r;\mathbf{q}\right) = \overline{R}$ for all r, otherwise $\beta\left(\cdot;\mathbf{q}\right)$ will be strictly increasing on $\left[\underline{R}, \overline{R}\right]$ if $h\left(\cdot\right) > 0$.

We define the firm's best response to the consumers in terms of the distribution of $\beta\left(R_i;\mathbf{q}\right)$.

Lemma 2. *The firm's best response is a map $\sigma_S : \mathbb{S}^{I-1} \to \mathcal{F}$ such that for any $\mathbf{q}$ in $\mathbb{S}^{I-1}$, $\sigma_S\left(\mathbf{q}\right)$ is the cdf of $\beta\left(R_i;\mathbf{q}\right)$ where $\beta\left(\cdot;\mathbf{q}\right)$ is defined as in (4).*

2.3 Equilibrium

We define a symmetric equilibrium for our game by any pair of consumer search proportions and induced cdf for firm's pricing strategy that simultaneously satisfy the best responses on both the demand and supply side.

Definition 1. *A pair* $(\mathbf{q}, F) \in \mathbb{S}^{I-1} \times \mathcal{F}$ *is a symmetric equilibrium if* $\mathbf{q} = \sigma_D(F)$ *and* $F = \sigma_S(\mathbf{q})$, *where* $\sigma_S(\cdot)$ *and* $\sigma_S(\cdot)$ *are defined in Lemmas 1 and 2 respectively.*

An equilibrium always exists in our model. For example the monopoly pricing strategy when all consumers search just once constitutes an equilibrium with: $\beta_M(r; \mathbf{q}_M) = \overline{R}$ for all r, and $\mathbf{q}_M$ such that $q_{1M} = 1$ and $q_{kM} = 0$ for $k \neq 1$. However, $\beta_M(R_i; \mathbf{q}_M)$ is degenerate and does not generate any price dispersion (cf. Diamond (1971)). Such equilibrium will be immediately rejected by the data where prices are not all the same. From (5), if $h(\cdot) > 0$, it is clear that $\beta(\cdot; \mathbf{q})$ is strictly increasing if and only if $q_1 < 1$. We will focus on this case. Theorem 1 characterizes such equilibria by $\mathbf{q}$ that satisfies (3) and (4) simultaneously.

Theorem 1. *In a symmetric equilibrium* $(\mathbf{q}, F)$ *where the equilibrium pricing strategy is strictly increasing,* $\mathbf{q}$ *satisfies the following system of equations:*

$$q_k = \begin{cases} 1 - G\left(\int F(p)(1 - F(p))\, dp\right) & \text{for } k = 1 \\ G\left(\int F(p)(1 - F(p))^{k-1}\, dp\right) - G\left(\int F(p)(1 - F(p))^{k}\, dp\right) & \text{for } k > 1 \end{cases}, \tag{6}$$

where $F(p) = H(\xi(p; \mathbf{q}))$ *for all* $p \in \left[\underline{P}, \overline{P}\right]$.

The characterization above shows that an equilibrium can be summarized by a fixed-point of some map, say $\mathcal{T}$. It can be shown using the implicit function theorem that $\mathcal{T}$ is a continuous map. Therefore an existence of an equilibrium with a price dispersion follows from a fixed-point theorem, such as Brouwer's, by showing that $\mathcal{T}$ maps a certain subset of $\mathbb{S}^{I-1}$ onto itself. There may be multiple equilibria that support price dispersion. We are not aware of any relevant uniqueness result in this context.

In subsequent sections we consider the econometric problem of identifying and estimating the model primitives from data generated from a particular equilibrium. We will henceforth drop the indexing arguments of equilibrium objects that are made explicit in this section for the purpose of defining best response and equilibrium. E.g. $\beta(\cdot; \mathbf{q})$ becomes $\beta(\cdot)$, $\mathbb{E}_F[\cdot]$ becomes $\mathbb{E}[\cdot]$ etc.

3 Nonparametric Identification

We assume to observe $\{(Y_{im}, P_{im})\}_{i=1,m=1}^{I,M}$ where Y_{im} and P_{im} respectively represent market share and price of firm i in market m. We describe properties of market shares below. Here M is the total number of markets and we will use a large M asymptotics framework.

Assumption D. $\{(Y_{im}, P_{im})\}_{i=1,m=1}^{I,M}$ *is a sequence of random variables such that:*

(i) there exists $(\mathbf{q}, F) \in \mathbb{S}^{I-1} \times \mathcal{F}$ *with* $q_1 < 1$ *so that* $P_{im} = \beta(R_{im}) \equiv \beta(R_{im}; \mathbf{q})$ *where* $\beta(\cdot; \mathbf{q})$ *has been defined in (4) for all* i, m *where* $\{R_{im}\}_{i=1,m=1}^{I,M}$ *is a sequence of i.i.d. random variables across* i *and* m*, with positive and finite density almost everywhere on* $[\underline{R}, \overline{R}]$*;*

(ii) $\{(Y_{1m}, \ldots, Y_{Im})\}_{m=1}^{M}$ *is an i.i.d. sequence of random variables across* m*, such that the joint distribution of* (Y_{im}, P_{im}) *for each* i, m *satisfies,*

$$\mathbb{E}[Y_{im}|P_{im}] = \sum_{k=1}^{I} q_k \frac{k}{I} (1 - F(P_{im}))^{k-1}. \tag{7}$$

Assumption D(i) assumes observed prices are a random sample that correspond to the model equilibrium. $q_1 < 1$ ensures $\beta(\cdot; \mathbf{q})$ is strictly increasing and price has a continuous distribution. Assumption D(ii) relates Y_{im} to P_{im}. Specifically, the RHS of (7) is precisely the *ex-ante* probability that a firm wins a sale when competing with $I-1$ other firms by setting her price to be P_{im}. This allows Y_{im} to be a mismeasurement of the theoretical market share. In particular, the theoretical *ex-post* market share can be formally defined by,

$$\overline{Y}_{im} = \frac{q_1}{\mathcal{C}_1^I} + \sum_{k=2}^{I} q_k \frac{\sum_{\mathcal{A} \in \mathcal{I}_{ik}} \mathbf{1}[P_{im} < \min_{j \in \mathcal{A}} \{P_{jm}\}]}{\mathcal{C}_k^I},$$

where $\mathcal{I}_{ik}$ is defined as $\left\{ \mathcal{A} = \bigcup_{j \in \{1,\ldots,I\} \setminus \{i\}} \{j\} \middle| |\mathcal{A}| = k-1 \right\}$ for $k = 2, \ldots, I$. We can motivate (7) by assuming that $Y_{im} = \overline{Y}_{im} + \epsilon_{im}$ where ϵ_{im} is an exogenous term that captures deviations from the model. If ϵ_{im} is mean independent of P_{im} then $\mathbb{E}[Y_{im}|P_{im}] = \mathbb{E}[\overline{Y}_{im}|P_{im}]$. Assumption D(ii) also assumes market shares across markets are i.i.d. but allows correlations within each market; Y_{im} and Y_{jm} are expected to be correlated when $i \neq j$ as market shares are jointly determined by the prices of all firms.

In Section 3.1 we consider identification on the demand side. We first identify $\mathbf{q}$ using (7), based on $\{(Y_{im}, P_{im})\}_{i=1,m=1}^{I,M}$, which can then be used to identify $G(\cdot)$. We identify $H(\cdot)$ in Section 3.2. For the latter, it suffices to show how to recover firm costs, $\{R_{im}\}_{i=1,m=1}^{I,M}$. In both Sections 3.1 and 3.2 we take $F(\cdot)$ and the joint distribution of (Y_{im}, P_{im}) for any (i, m) to be known. Both of these objects are nonparametrically identified under Assumption D when $M \to \infty$.

3.1 Consumers

Let X_{im} be a vector in $\mathbb{R}^I$ such that $(X_{im})_k = \frac{k}{I}(1 - F(P_{im}))^{k-1}$. We can write (7) as

$$Y_{im} = X_{im}^\top \mathbf{q} + \varepsilon_{im}, \tag{8}$$

where ε_{im} satisfies $\mathbb{E}[\varepsilon_{im}|P_{im}] = 0$. We can then identify $\mathbf{q}$ as the solution of a least squares problem.

Lemma 3. *Suppose Assumption D holds. If* $\mathbb{E}\left[X_{im}X_{im}^{\top}\right]$ *has full rank then* $\mathbf{q}$ *is identified.*

We now treat both $\mathbf{q}$ and $F(\cdot)$ as known and use them to identify $G(\cdot)$ at $\{\Delta_k\}_{k=1}^{I-1}$

Proposition 1. *Suppose Assumption D holds. Then* $G(\Delta_k)$ *is identified for* $k = 1, \ldots, I-1$.

Proof. From (4), we see that $G(\Delta_k) = 1 - \sum_{k'=1}^{k} q_{k'}$ for $k = 1, \ldots, I-1$. The proof follows since both $\{\Delta_k\}_{k=1}^{I-1}$ and $\mathbf{q}$ are identified. In particular, note that Δ_k is a functional of $F(\cdot)$ for all k (see (1) and (2)) and $\mathbf{q}$ is identified from Lemma 3. ■

The proof of Proposition 1 is similar to how Hong and Shum (2006) and Moraga-González and Wildenbeest (2008) identify the search cost distribution in models with complete information where all firms produce at the same cost. Ours differs from theirs only in how we identify $\mathbf{q}$. Like their results, this is a partial identification result as we can only identify $G(\cdot)$ at $\{\Delta_k\}_{k=1}^{I}$.

It is possible to identify $G(\cdot)$ almost everywhere on $\left[0, \overline{C}\right]$ if there is sufficient exogenous variation across markets. For example, suppose there are L market types where consumers draw search costs from the same distribution but firms production costs have different distribution across types and/or the number of firms may vary with L. If we are able to identify $\{\Delta_{kL}\}_{k=1}^{I_L-1}$ for market type L and have $L \to \infty$ in such a way that $\bigcup_{l=1}^{L} \{\Delta_{kl} : k = 1, \ldots, I_L - 1\}$ *grow dense*[8] in $\left[0, \overline{C}\right]$, then we can identify $G(\cdot)$. Moraga-González, Sandor and Wildenbeest (2013) propose this identification strategy for their search model with complete information. Their strategy is also directly applicable in the incomplete information environment that we consider.

3.2 Firms

Our identification strategy for $H(\cdot)$ takes similar steps to how GPV identifies the distribution of bidder's valuation in a first-price sealed-bid auction model from bids data. We derive the inverse of the equilibrium pricing strategy in terms of identifiable objects the price distribution in Lemma 4. We then use it to recover the latent marginal costs from observed prices.

Lemma 4. *Suppose Assumption D(i) holds. Then the inverse of the equilibrium pricing strategy,* $\xi : \left[\underline{P}, \overline{P}\right] \to \left[\underline{R}, \overline{R}\right]$, *exists and takes the following form*

$$\xi(p) = p - \frac{\sum_{k=1}^{I} q_k k (1 - F(p))^{k-1}}{f(p) \sum_{k=2}^{I} q_k k (k-1) (1 - F(p))^{k-2}}, \tag{9}$$

and $\xi\left(\overline{P}\right) = \overline{R}$.

[8] Suppose A be an interval on the real line and $\{A_L\}$ is an increasing sequence such that $A_L \subseteq A$. We say that $\{A_L\}$ *grows dense* in A if for any $x \in A$ and $\varepsilon > 0$, there exists $L(x, \varepsilon)$ and $x_0 \in A_{L(x,\varepsilon)}$ such that $|x - x_0| < \varepsilon$.

Since $\xi(\cdot)$ is identified, we can use it to invert production costs from prices to identify $H(\cdot)$.

PROPOSITION 2. *Suppose Assumption D holds. Then $H(\cdot)$ is identified.*

PROOF. Under Assumption D, $\xi(\cdot)$ is identified. Therefore we can recover R_{im} from $\xi(P_{im})$ for all i, m. ■

4 Estimation and Convergence Rates

We now focus on estimating $h(\cdot)$ with the best possible convergence rate. Along the way we will discuss estimators of other parameters in the model. Parameters on the demand side can be estimated at the parametric rate and our discussion on them will be brief.

We consider two separate cases. First, we consider the uniform convergence for $h(\cdot)$ over any fixed closed interval that lies in the interior of $\left[\underline{R}, \overline{R}\right]$. In this case we can provide an estimator for $h(\cdot)$ that achieves the same optimal convergence rate as the GPV estimator. Second, we consider uniform convergence over an expanding interval that approaches $\left[\underline{R}, \overline{R}\right]$ as the sample size increases. In this case we will provide another estimator for $h(\cdot)$ that can converge at an arbitrarily close rate to the one over fixed support. The reason for a slightly slower convergence rate for the latter case is to account for $\lim_{p \to \overline{P}} f(p) = \infty$, which is a feature of the equilibrium. Our estimator for $h(\cdot)$ in both cases will be based on kernel smoothing using estimated $\{R_{im}\}_{i=1,m=1}^{I,M}$, which is to be obtained through the estimated inverse of the pricing function (see (9)). In particular, the inverse of the pricing function depends on $(\mathbf{q}, F(\cdot), f(\cdot))$ that have to be estimated.

To study convergence rates, we have to specify the degree of smoothness of $H(\cdot)$.

ASSUMPTION R. *$H(\cdot)$ admits upto $\tau + 1$ continuous derivatives on $\left[\underline{R}, \overline{R}\right]$ for some $\tau \geq 1$.*

LEMMA 5. *Suppose Assumptions D and R hold. Then $f(\cdot)$ admits upto $\tau + 1$ continuous derivatives on $\left[\underline{P}, \overline{P}\right)$ for the same τ as in Assumption R.*

Lemma 5 says that $f(\cdot)$ has the same degree of smoothness as $H(\cdot)$ everywhere on the equilibrium price support other than at $\overline{P}$. As alluded above, $f(p)$ may diverge to infinity as $p \to \overline{P}$ and its derivative will also not be defined at $\overline{P}$. However, this possibility has no effect on the convergence rate we will derive over any closed inner subset of $\left[\underline{R}, \overline{R}\right]$.

We next define estimators for $(\mathbf{q}, f(\cdot), F(\cdot))$ and discuss their convergence rate under assumptions D and R.

An estimator for $F(\cdot)$

A natural estimator for $F(\cdot)$ is the empirical cdf, defined as

$$\widehat{F}(p) = \frac{1}{MI}\sum_{m=1}^{M}\sum_{i=1}^{I}\mathbf{1}\left[P_{im} \leq p\right] \quad \text{for all } p. \tag{10}$$

It is well-known from Donsker's theorem that $\sqrt{M}\left(\widehat{F}(\cdot) - F(\cdot)\right)$ converges weakly to a Gaussian process on $\left[\underline{P}, \overline{P}\right]$. Then by the continuous mapping theorem, $\sup_{p\in\left[\underline{P},\overline{P}\right]}\left|\widehat{F}(p) - F(p)\right| = O_p\left(1/\sqrt{M}\right)$.

An estimator for $\mathbf{q}$

We estimate $\mathbf{q}$ by least squares. Let $\mathbf{Y}_m = (Y_{1m}, \ldots, Y_{Im})^\top$, $\mathbf{e}_m = (\varepsilon_{1m}, \ldots, \varepsilon_{Im})^\top$ and $\mathbf{X}_m$ be an $I \times I$ matrix such that $(\mathbf{X}_m)_{ik} = \frac{k}{I}(1 - F(P_{im}))^{k-1}$. Vectorize $\mathbf{Y}_m$, $\mathbf{X}_m$ and $\mathbf{e}_m$ across m to form $\mathbf{Y} = \left[\mathbf{Y}_1^\top : \cdots : \mathbf{Y}_M^\top\right]^\top$, $\mathbf{X} = \left[\mathbf{X}_1^\top : \cdots : \mathbf{X}_M^\top\right]^\top$ and $\mathbf{e} = \left[\mathbf{e}_1^\top : \cdots : \mathbf{e}_M^\top\right]^\top$ respectively. Then a vector version of (8) is,

$$\mathbf{Y} = \mathbf{X}\mathbf{q} + \mathbf{e}.$$

$F(\cdot)$ is unknown and has to be estimated. Let $\widehat{\mathbf{X}}$ be the feasible counterpart of $\mathbf{X}$ where $F(\cdot)$ is replaced by $\widehat{F}(\cdot)$. Then,

$$\begin{aligned}
\widehat{\mathbf{q}} &= \left(\widehat{\mathbf{X}}^\top\widehat{\mathbf{X}}\right)^{-1}\widehat{\mathbf{X}}^\top\mathbf{Y}, \\
&= \mathbf{q} + \mathbf{a}_M + b_{f,M}, \text{ where} \\
\mathbf{a}_M &= \left(\mathbf{X}^\top\mathbf{X}\right)^{-1}\mathbf{X}^\top\mathbf{e}, \\
b_{f,M} &= \left(\left(\widehat{\mathbf{X}}^\top\widehat{\mathbf{X}}\right)^{-1}\widehat{\mathbf{X}}^\top - \left(\mathbf{X}^\top\mathbf{X}\right)^{-1}\mathbf{X}^\top\right)\mathbf{Y}.
\end{aligned} \tag{11}$$

Using asymptotic theory for clustered samples (e.g. see Hansen and Lee (2019)), $\|\mathbf{a}_M\| = O_p\left(1/\sqrt{M}\right)$ as $\frac{1}{M}\mathbf{X}^\top\mathbf{X} = \frac{1}{M}\sum_{m=1}^{M}\left(\sum_{i=1}^{I}X_{im}X_{im}^\top\right)$ and $\frac{1}{\sqrt{M}}\mathbf{X}^\top\mathbf{e} = \frac{1}{\sqrt{M}}\sum_{m=1}^{M}\left(\sum_{i=1}^{I}X_{im}\varepsilon_{im}\right)$ would satisfy a law of large numbers and central limit theorem respectively. Since $\left(\widehat{\mathbf{X}}^\top\widehat{\mathbf{X}}\right)^{-1}\widehat{\mathbf{X}}^\top$ is a smooth functional of $\widehat{F}(\cdot)$, it can also be verified by applications of the continuous mapping theorem that $\|b_{f,M}\| = O_p\left(1/\sqrt{M}\right)$. Thus, $\|\widehat{\mathbf{q}} - \mathbf{q}\| = O_p\left(1/\sqrt{M}\right)$.

Furthermore, since Δ_k is a functional of $F(\cdot)$, we can estimate $G(\Delta_k)$ using $\widehat{\mathbf{q}}$ and $\widehat{F}(\cdot)$ based on the constructive identification result in Proposition 1. Such estimator will be a smooth functional of $\widehat{F}(\cdot)$ and have a $\sqrt{M}-$convergence rate. Estimating $G(\cdot)$ as a curve is also possible when there are data from different equilibria that can identify more points on the support of the search cost. In this case, Sanches, Silva and Srisuma (2018) propose a series estimator that pooled data across equilibria based on using estimated Δ_k and $G(\Delta_k)$ as generated regressor and regressand respectively; they also derive the convergence rate of such estimator. The same type of estimator can also be constructed here. We refer the reader to Section 4 of Sanches, Silva and Srisuma (2018) for details.

An estimator for $f(\cdot)$

We use a kernel density estimator, defined as:

$$\widehat{f}(p) = \frac{1}{MIb_{f,M}} \sum_{m=1}^{M} \sum_{i=1}^{I} K\left(\frac{P_{im}-p}{b_{f,M}}\right) \text{ for all } p, \tag{12}$$

where $K(\cdot)$ is a $(\tau+1)-$th higher order kernel function and $b_{f,M}$ is a bandwidth that is proportional to the optimal bandwidth that converges to zero at the rate $\left(\frac{\log M}{M}\right)^{\frac{1}{2\tau+3}}$, see Härdle (1991). Let $\eta_M^* \equiv \left(\frac{\log M}{M}\right)^{\frac{\tau+1}{2\tau+3}}$ denote the optimal rate of convergence for density estimation with $\tau+1$ continuous derivatives (Stone (1982)). Then it is well-known that

$$\sup_{p\in[\underline{P}+\delta,\overline{P}-\delta]} \left|\widehat{f}(p) - f(p)\right| = O(\eta_M^*) \text{ a.s.},$$

for any positive δ.

We summarize the convergence rates of $\widehat{\mathbf{q}}$, $\widehat{F}(\cdot)$, and $\widehat{f}(\cdot)$ in a proposition.

Proposition 3. *Suppose Assumptions D and R hold. Then for the estimators* defined *in (10) to (12):*

(a) $\sup_{p\in[\underline{P},\overline{P}]} \left|\widehat{F}(p) - F(p)\right| = O_p\left(1/\sqrt{M}\right)$;

(b) $\|\widehat{\mathbf{q}} - \mathbf{q}\| = O_p\left(1/\sqrt{M}\right)$;

(c) For any positive δ, $\sup_{p\in[\underline{P}+\delta,\overline{P}-\delta]} \left|\widehat{f}(p) - f(p)\right| = O(\eta_M^*)$ *a.s.*

We next proceed to estimate $h(\cdot)$ using the estimators for $\mathbf{q}$, $f(\cdot)$ and $F(\cdot)$ described above.

An estimator for $h(\cdot)$

We start by obtaining an estimator for R_{im}, using

$$\widehat{R}_{im} = \begin{cases} P_{im} - \frac{\sum_{k=1}^{I} \widehat{q}_k k\left(1-\widehat{F}(P_{im})\right)^{k-1}}{\widehat{f}(P_{im}) \sum_{k=1}^{I} \widehat{q}_k k(k-1)\left(1-\widehat{F}(P_{im})\right)^{k-2}} & \text{for } P_{im} \in [\underline{P}+\delta, \overline{P}-\delta] \\ +\infty & \text{otherwise} \end{cases}. \tag{13}$$

When $\widehat{R}_{im} < \infty$, $\widehat{R}_{im}$ is the estimator of R_{im} based on on the feasible version of (9). In this case $\widehat{R}_{im}$ is a smooth function of $\widehat{\mathbf{q}}$, $\widehat{F}(P_{im})$ and $\widehat{f}(P_{im})$. Lemma 6 shows that its convergence rate to R_{im} is determined by $\sup_{p\in[\underline{P}+\delta,\overline{P}-\delta]} \left|\widehat{f}(p) - f(p)\right|$. We will effectively be throwing away $\widehat{R}_{im}$ that is not finite when it comes to estimating $h(\cdot)$. We will prove this has no effect on our asymptotic results because the probability that $\widehat{R}_{im}=\infty$ goes to zero asymptotically.

LEMMA 6. *Suppose Assumptions D and R hold. Then,*

$$\sup_{i,m\text{ s.t. }\widehat{R}_{im}<\infty}\left|\widehat{R}_{im}-R_{im}\right|=O\left(\eta_M^*\right)\ a.s.$$

We define our estimator for $h(\cdot)$ as follows:

$$\widehat{h}(r)=\frac{1}{MIb_{h,M}}\sum_{m=1}^{M}\sum_{i=1}^{I}K\left(\frac{\widehat{R}_{im}-r}{b_{h,M}}\right)\quad\text{for any } r. \tag{14}$$

Here $K(\cdot)$ is a kernel function with a bandwidth $b_{h,M}$. Under the conditions of Theorem 2, we can use the convergence rate of $\widetilde{R}_{im}$ to R_{im} to determine the uniform convergence rate of $\widehat{h}(\cdot)$ to $h(\cdot)$.

THEOREM 2. *Suppose Assumptions D and R hold. Assume the following properties for components in (14):*
(i) $K(\cdot)$ be a symmetric $\tau-$th order kernel with support $[-1,1]$;
(ii) $K(\cdot)$ is twice continuously differentiable on $[-1,1]$;
(iii) $b_{h,M}$ is proportional to $\left(\frac{\log M}{M}\right)^{\frac{1}{2\tau+3}}$.
Then for any $\varsigma>0$, there exists $\delta>0$ so that part (c) of Proposition 3 holds such that

$$\sup_{r\in\left[\underline{R}+\varsigma,\overline{R}-\varsigma\right]}\left|\widehat{h}(r)-h(r)\right|=O\left(\left(\frac{\log M}{M}\right)^{\frac{\tau}{2\tau+3}}\right)\quad a.s.$$

The rate $\left(\frac{\log M}{M}\right)^{\frac{\tau}{2\tau+3}}$ is equal to $\frac{\eta_M^*}{b_{h,M}}$, which is the optimal convergence rate GPV derived in their paper. This rate is achieved by choosing $b_{h,M}$ that oversmooths relative to the optimal bandwidth for a $\tau-$times continuously differentiable density function.

Next, we consider the uniform convergence for an estimator of $h(\cdot)$ over $\left[\underline{R}+\varsigma_M,\overline{R}-\varsigma_M\right]$ for some $\varsigma_M=o(1)$. We use a different estimator to $\widehat{h}(\cdot)$ due to the possibility of $f(\cdot)$ having a pole at $\overline{P}$. Specifically, one of the standard assumptions assumed when deriving uniform convergence of kernel density estimators is that the density is uniformly bounded (e.g. see Andrews (1995), Masry (1996), Fan and Yao (2003), Hansen (2008)). The concern regarding having an unbounded density is relevant for our model as the following lemma shows.

LEMMA 7. *Suppose Assumption D(i) holds. If $q_1>0$ and $H(\cdot)$ is continuously differentiable on $\left[\underline{R},\overline{R}\right]$ then $f(\cdot)$ is continuous on $\left[\underline{P},\overline{P}\right)$ and satisfies:*

(a) $\inf_{p\in[\underline{P},\overline{P}]} f(p) > 0$;

(b) $\lim_{p\to\overline{P}} f(p) = \infty$.

We expect $q_1 > 0$ to be the rule rather than an exception in equilibrium. For other cases, as we show in the appendix that $f(p) = \frac{h(\beta^{-1}(p))}{\beta'(\beta^{-1}(p))}$ (see equation (26)), the behavior of $f(\cdot)$ near $\overline{P}$ can be complicated as it generally depends on both $H(\cdot)$ and which components of $\mathbf{q}$ are zero in equilibrium.

We next propose an estimator for $h(\cdot)$ that can converge at an arbitrarily close rate to $\left(\frac{\log M}{M}\right)^{\frac{\tau}{2\tau+3}}$ uniformly over an expanding interval without having to impose further structure on $H(\cdot)$. Our strategy is to take a log-transformation of the price to suppress the pole. We show this suffices without having to specify the behavior of $f(\cdot)$ near $\overline{P}$ (c.f. Marron and Ruppert (1993)) since $f(\cdot)$ is integrable. Specifically, let $P_{im}^{\dagger} \equiv -\ln\left(\overline{P} - P_{im}\right)$. Denote the pdf of $P_{im}^{\dagger}$ by $f^{\dagger}(\cdot)$. By a change of variable, we have $f(p) = \frac{f^{\dagger}(-\ln(\overline{P}-p))}{\overline{P}-p}$ for $p \in [\underline{P}, \overline{P}]$. And from the perspective of $P_{im}^{\dagger}$, whose support is $[-\ln\left(\overline{P} - \underline{P}\right), \infty)$, we have for any $p^{\dagger} \in [-\ln\left(\overline{P} - \underline{P}\right), \infty)$, $f^{\dagger}\left(p^{\dagger}\right) = \exp\left(-p^{\dagger}\right) f\left(\overline{P} - \exp\left(-p^{\dagger}\right)\right)$.

Lemma 8 shows $f^{\dagger}(\cdot)$ and its derivatives are uniformly bounded, and $f^{\dagger}(\cdot)$ has the same degree of smoothness as that of $f(\cdot)$. This result holds regardless whether $f(\cdot)$ has a pole or not.

LEMMA 8. *Suppose Assumptions D(i) and R hold.*

(a) $\sup_{p^{\dagger}\in[-\ln(\overline{P}-\underline{P}),\infty)} f^{\dagger}(p^{\dagger}) < \infty$;

(b) $f^{\dagger}(\cdot)$ *admits upto* $\tau+1$ *continuous and uniformly bounded derivatives on* $[-\ln\left(\overline{P}-\underline{P}\right), \infty)$ *for the same* τ as in Assumption R.

We estimate $f(\cdot)$ using the following estimator,

$$\begin{aligned} \widetilde{f}(p) &= \frac{\widehat{f^{\dagger}}\left(-\ln\left(\overline{P}-p\right)\right)}{\overline{P}-p}, \text{ where} \\ \widehat{f^{\dagger}}\left(p^{\dagger}\right) &= \frac{1}{MIb_{f^{\dagger},M}} \sum_{m=1}^{M} \sum_{i=1}^{I} K\left(\frac{P_{im}^{\dagger} - p^{\dagger}}{b_{f^{\dagger},M}}\right) \quad \text{for all } p^{\dagger}, \end{aligned} \tag{15}$$

where $K(\cdot)$ is a $(\tau+1)$ −th higher order kernel function with a bandwidth $b_{f^{\dagger},M}$ that is proportional to $\left(\frac{\log M}{M}\right)^{\frac{1}{2\tau+3}}$. Then, based on Lemma 8, $\left|\widehat{f^{\dagger}}\left(p^{\dagger}\right) - f^{\dagger}\left(p^{\dagger}\right)\right| = O\left(\eta_M^*\right)$ a.s. uniformly over any fixed inner proper subset of $[-\ln\left(\overline{P}-\underline{P}\right), \infty)$. Since

$$\widetilde{f}(p) - f(p) = \frac{\widehat{f^{\dagger}}\left(-\ln\left(\overline{P}-p\right)\right) - f^{\dagger}\left(-\ln\left(\overline{P}-p\right)\right)}{\overline{P}-p},$$

it follows that $\left|\widetilde{f}(p) - f(p)\right| = O\left(\eta_M^*\right)$ a.s. uniformly over any fixed inner proper subset of $\left[\underline{P}, \overline{P}\right]$. However, when we consider uniform convergence over an expanding support the rates for this estimator can be worse than η_M^*. First, the bias from estimating $f^{\dagger}(p)$ when $p^{\dagger}$ lies within a $b_{f^{\dagger},M}$-neighborhood from $-\ln\left(\overline{P}-\underline{P}\right)$ is of larger order of magnitude than an interior point. Second, as p

approaches $\overline{P}$, $\left(\overline{P}-p\right)^{-1}$ diverges to infinity and this makes the η_M^* convergence rate unachievable uniformly over any expanding subset of the support that grows to arbitrarily close to $\overline{P}$.

We can avoid the bias at the boundary by choosing an expanding interval away from the boundary. For this, let's δ_M' be any positive number that decreases to zero such that $\overline{P}-\exp\left(\ln\left(\overline{P}-\underline{P}\right)-b_{f^\dagger,M}\right)=o(\delta_M')$. The divergence rate from $\left(\overline{P}-p\right)^{-1}$ can be controlled by how fast p is approaching $\overline{P}$. Then, for any δ_M'' that decreases to 0 we have

$$\sup_{p\in[\underline{P}+\delta_M',\overline{P}-\delta_M'']}\left|\widetilde{f}(p)-f(p)\right|=O\left(\frac{\eta_M^*}{\delta_M''}\right)\text{ a.s.}$$

We can therefore always find an interval expanding to $[\underline{P},\overline{P}]$ that $\widetilde{f}(\cdot)-f(\cdot)$ converges uniformly at a rate as close to η_M^* as we like.

PROPOSITION 4. *Suppose Assumptions D(i) and R hold. Then for any sequence of positive reals* $\{\eta_m\}_{m=1}^M$ *that decreases to* 0 *such that* $\eta_M^*=o(\eta_M)$, *there exists some sequence* $\{\delta_m\}_{m=1}^M$ *that decreases to* 0 *such that* $\sup_{p\in\left[\underline{P}+\delta_M,\overline{P}-\delta_M\right]}\left|\widetilde{f}(p)-f(p)\right|=O(\eta_M)$ *a.s.*

We can then estimate R_{im} on using $\widetilde{f}(\cdot)$, and use it to estimate $h(\cdot)$ as done previously in (13) and (14) respectively. Specifically, for any $\delta_M>0$ let

$$\widetilde{R}_{im}=\begin{cases}P_{im}-\frac{\sum_{k=1}^I\widehat{q}_k k\left(1-\widehat{F}(P_{im})\right)^{k-1}}{\widetilde{f}(P_{im})\sum_{k=1}^I\widehat{q}_k k(k-1)\left(1-\widehat{F}(P_{im})\right)^{k-2}} & \text{for } P_{im}\in[\underline{P}+\delta_M,\overline{P}-\delta_M]\\ +\infty & \text{otherwise}\end{cases},\tag{16}$$

$$\widetilde{h}(r)=\frac{1}{MIb_{h,M}}\sum_{m=1}^M\sum_{i=1}^I K\left(\frac{\widetilde{R}_{im}-r}{b_{h,M}}\right)\quad\text{for any } r,\tag{17}$$

where $K(\cdot)$ is a kernel function with a bandwidth $b_{h,M}$. The following results are similar to Lemma 6 and Theorem. They differ in that the rates below do not reach the optimal rate but can be arbitrarily close to it, and the results are valid over a sequence of expanding intervals instead of a fixed interval.

LEMMA 9. *Suppose Assumptions D and R hold. Then for any sequence of positive reals* $\{\eta_m\}_{m=1}^M$ *that decreases to* 0 *such that* $\eta_M^*=o(\eta_M)$, *there exists some sequence* $\{\delta_m\}_{m=1}^M$ *as described in Proposition 4 such that*

$$\sup_{i,m\text{ s.t. }\widetilde{R}_{im}<\infty}\left|\widetilde{R}_{im}-R_{im}\right|=O(\eta_M)\ \ a.s.$$

THEOREM 3. *Suppose Assumptions D and R hold. Assume the following properties for components in (17):*

(i) $K(\cdot)$ *be a symmetric* $\tau-$*th order kernel with support* $[-1,1]$*;*

(ii) $K(\cdot)$ *is twice continuously differentiable on* $[-1,1]$*;*

(iii) $b_{h,M}$ *is proportional to* $\left(\frac{\log M}{M}\right)^{\frac{1}{2\tau+3}}$.

Then for any η_M *that satisfies* $\eta_M^* = o(\eta_M)$ *and* $\eta_M = O\left(b_{h,M}^2\right)$*, and for* ς_M *that decreases to zero such that* $b_{h,M} = o(\varsigma_M)$*,*

$$\sup_{r\in\left[\underline{R}+\varsigma_M,\overline{R}-\varsigma_M\right]}\left|\widetilde{h}(r)-h(r)\right| = O\left(\frac{\eta_M}{b_{h,M}}\right) \quad a.s.$$

The uniform convergence rate for $\widetilde{h}(\cdot)$ is derived over an expanding support that avoids the boundary effect. The additional condition Theorem 3 imposes to handle the possibility of a pole, which is not required for Theorem 2, is that $\eta_M = O\left(b_{h,M}^2\right)$. This is a mild condition. E.g., when $\tau \geq 2$ this condition is not restrictive. To see this, suppose $\eta_M = \eta_M^*\phi_M$ for some ϕ_M with $\lim_{M\to\infty}\phi_M = \infty$, then $\eta_M = O\left(b_{h,M}^2\right)$ is equivalent to $\phi_M\left(\frac{\log M}{M}\right)^{\frac{\tau-1}{2\tau+3}} = O(1)$. Since we are only be interested in ϕ_M that diverges to infinity slowly for a tight upper bound we can choose it to diverge at an aribrarily slow rate.

In practice, however, even if one is not interested in the uniform convergence rate over an expanding support it is important to be aware of the presence of the pole. We will illustrate this in Section 6.

5 Extension

Thus far our sellers were assumed ex-ante identical. We now introduce a search model where firms offer vertically differentiated products. Our model can be seen as an incomplete information counterpart to the model proposed in Wildenbeest (2011), which generalizes the homogeneous product search model with complete information of Moraga-González and Wildenbeest (2008) to the case of differentiated products. Our discussion here will focus on identification. The estimation strategy and the convergence rates of developed in Section 4 can be readily extended to this setting.

5.1 Product Differentiation

Firm i's product is characterized by $\nu_i \in \mathbb{R}$, which is a measure of differentiated quality. Consumers and firms observe quality of all products, which these are not observable to the econometrician. The main modelling assumption employed by Wildenbeest (2011)is that the difference between quality and marginal cost is the same for all firms. A natural way to extend his idea to an incomplete information game is to put a common distribution around ν_i for all i. We will show a quasi-symmetric equilibrium, where optimal pricing strategies between firms differ only by the differences in their qualities, can then be characterized analogously to Theorem 1.

Consumer's Best Responses

Consumers now value products from different firms differently. The utility they derive from purchasing from seller i is U_i. We assume,

$$U_i := \nu_0 + \nu_i - P_i, \tag{18}$$

where ν_0 denotes the common value of the product, ν_i denotes the valuation of the differentiating component due to firm i, and P_i denotes its corresponding price. One can, for example, attribute ν_i to physical quality or other experience associated with purchasing from firm i. A consumer with search cost c faces the following decision problem:

$$\max_{1 \leq k \leq I} \mathbb{E}_L \left[U_{(k:k)}\right] - \ c\,(k-1)\,.$$

We again assume that the first search is free and a purchase is always made. Note that ν_0 does not enter our analysis, just as it does not in the model with a homogeneous product.[9] For the moment suppose firms set prices such that $\{U_i\}_{i=1}^{I}$ is a random sample. Then, for $k \geq 1$, let $U_{(k:k)}$ be the maximum of k i.i.d. random variables of utilities and $\mathbb{E}_L[\cdot]$ denotes an expectation where the random utilities have distribution described by the cdf $L(\cdot)$.

Consumers' search strategy will again be determined by the marginal gain they expect to get from searching. We denote the expected marginal utility gain from a purchase when a consumer searches one more firm when she has already searched $k-1$ firms by

$$\Upsilon_k(L) := \mathbb{E}_L\left[U_{(k:k)}\right] - \mathbb{E}_L\left[U_{(k-1:k-1)}\right]. \tag{19}$$

We normalize the outside option such that $U_{(0:0)} = 0$. The consumer's best response is to search once if and only if $c > \Upsilon_1(L)$, and search $k > 1$ times if and only if $\Upsilon_{k-1}(L) < c < \Upsilon_k(L)$. Analogous to the discussions in Section 2.1, $\Upsilon_k(L)$ is positive and strictly decreasing when the distribution of U_i is non-degenerate.

Firm's Best Responses

We assume firm i's production cost consists of a sum of deterministic (determined by quality) and random (variable) components:

$$R_i = \nu_i + R_{0i},$$

where R_{0i} has cdf $H_0(\cdot)$ supported on $\mathcal{R}_0 \in \left[\underline{R}_0, \overline{R}_0\right]$ for some $\overline{R}_0 > \underline{R}_0 > 0$. We denote the support of R_i by $\mathcal{R}_i := \left[\nu_i + \underline{R}_0, \nu_i + \overline{R}_0\right]$. We assume firm costs are independent draws to preserve the independent value environment. Subsequently $\{R_{0i}\}_{i=1}^{I}$ is an i.i.d. sequence of random variables.

[9] One can think of ν_0 to be sufficiently high so that consumers always purchase. But the firms cannot extract further profit from consumers due to an upper bound on the price they can set.

We restrict our attention to quasi-symmetric pricing strategies where firms' strategies are affine translations from one another. Denote firm i's pricing strategy by $\beta_i(\cdot;\mathbf{q}) : \mathcal{R}_i \to \mathcal{P}_i$, where $\mathcal{P}_i = \left[\nu_i + \underline{P}_0, \nu_i + \overline{P}_0\right]$ and $\beta_i(\cdot;\mathbf{q}) = \nu_i + \beta_0(\cdot;\mathbf{q})$ and $\beta_0(\cdot;\mathbf{q}) : \mathcal{R}_0 \to \mathcal{P}_0 = \left[\underline{P}_0, \overline{P}_0\right]$. We denote the valuation-cost markup by $X_i := \nu_i - R_i$. By construction $X_i = -R_{0i}$ and $\{X_i\}_{i=1}^I$ is i.i.d. across firms. Since $U_i = \nu_i - P_i$, we can equivalently study the firm i's profit maximization problem where the firm sets the level of utility consumers would get from buying its product instead of setting prices. I.e., for any $x_i \in \left[-\overline{R}_0, -\underline{R}_0\right]$, consider

$$\max_u \Gamma(u, x_i;\mathbf{q}), \text{ where}$$
$$\Gamma(u, x_i;\mathbf{q}) = (x_i - u)\sum_{k=1}^I q_k \frac{k}{I}\mathbb{P}\left[U_{(k-1:k-1)} \leq u\right].$$

Suppose a solution to the maximization problem above exists and let $\mu(x_i;\mathbf{q}) := \arg\max_u \Gamma(u, x_i;\mathbf{q})$ for any $(x_i, \mathbf{q})$. We assume that $\mu(x_i;\mathbf{q})$ to be increasing in x_i and satisfies the boundary condition that $\mu\left(-\overline{R}_0;\mathbf{q}\right) = \overline{R}_0$. Under this premise, we can apply the same type of arguments used to obtain (4) to show that: for any $r_{0i} \in \mathcal{R}_0$ and $x(r_{0i}) := -r_{0i}$,

$$\mu(x(r_{0i});\mathbf{q}) = x(r_{0i}) - \frac{\sum_{k=1}^I q_k k \int_{s=r_{0i}}^{\overline{R}_0} (1 - H_0(s))^{k-1} ds}{\sum_{k=1}^I q_k k (1 - H_0(r_{0i}))^{k-1}}. \tag{20}$$

Therefore $\{\mu(x(R_{0i});\mathbf{q})\}_{i=1}^I$ is an i.i.d. sequence of random utilities that firms offer to the consumers upon drawing $\{R_{0i}\}_{i=1}^I$ as a best response given $\mathbf{q}$.

For any $r_i = \nu_i + r_{0i}$, since $\mu_i(x_i(r_i);\mathbf{q}) = \nu_i - \beta_i(r_i;\mathbf{q})$, it follows that $\beta_i(r_i;\mathbf{q}) = \nu_i + \beta_0(r_{0i};\mathbf{q})$ where,

$$\beta_0(r_{0i};\mathbf{q}) = r_{0i} + \frac{\sum_{k=1}^I q_k k \int_{s=r_{0i}}^{\overline{R}_0} (1 - H_0(s))^{k-1} ds}{\sum_{k=1}^I q_k k (1 - H_0(r_{0i}))^{k-1}}. \tag{21}$$

$\beta_0(\cdot;\mathbf{q})$ has an identical structure to $\beta(\cdot;\mathbf{q})$ as defined in (4). Therefore the properties of each firm's pricing strategy derived here are the same as that of the homogeneous product case other than being shifted by a constant ν_i. In particular $\beta_0(\cdot;\mathbf{q})$ is strictly increasing when $q_1 < 1$, and its inverse takes the same form as (9) in Lemma 5.

Equilibrium

We can now define a quasi-symmetric equilibrium, where players using pricing strategies that are affine translation from each other, and characterize the equilibrium of the game using an analogous expression to Theorem 1. We state the characterization of the equilibrium in Theorem 3. We provide its proof and the definition of a quasi-symmetric equilibrium in the Appendix.

THEOREM 4. *In a quasi-symmetric equilibrium* $(\mathbf{q}, F_0)$ *where the equilibrium pricing strategies are strictly increasing,* $\mathbf{q}$ *satisfies the following system of equations:*

$$q_k = \begin{cases} 1 - G\left(\int F_0(p)(1 - F_0(p))\, dp\right) & \text{for } k = 1 \\ G\left(\int F_0(p)(1 - F_0(p))^{k-1}\, dp\right) - G\left(\int F_0(p)(1 - F_0(p))^{k}\, dp\right) & \text{for } k > 1 \end{cases},$$

where $F_0(p) = H_0(\xi_0(p; \mathbf{q}))$ *for all* $p \in [\underline{P}, \overline{P}]$ *and* $\xi_0(\cdot; \mathbf{q})$ *is the inverse of* $\beta_0(\cdot; \mathbf{q})$.

5.2 Identification

For the rest of this section, suppose we have data generated from a non-degenerate equilibrium described in the previous subsection as described by the following assumption.

ASSUMPTION D'. $\{(Y_{im}, P_{im})\}_{i=1,m=1}^{I,M}$ *is a sequence of random variables such that:*

(i) there exists $(\mathbf{q}, F_0) \in \mathbb{S}^{I-1} \times \mathcal{F}$ *with* $q_1 \in (0,1)$ *so that* $P_{im} = \nu_i + \beta_0(R_{0im}; \mathbf{q})$ *where* $\beta_0(\cdot; \mathbf{q})$ *has been defined in (21) for all* i, m *where* $\{R_{0im}\}_{i=1,m=1}^{I,M}$ *is a sequence of i.i.d. continuous random variables with almost everywhere positive density on* $[\underline{R}_0, \overline{R}_0]$;

(ii) $\{(Y_{1m}, \ldots, Y_{Im})\}_{m=1}^{M}$ *is an i.i.d. sequence of random vectors such that the joint distribution of* (Y_{im}, P_{0im}) *for each* i, m *satisfies,*

$$\mathbb{E}[Y_{im}|P_{0im}] = \sum_{k=1}^{I} q_k \frac{k}{I}(1 - F_0(P_{0im}))^{k-1}. \tag{22}$$

If we observe $\{\nu_i\}_{i=1}^{I}$, we can construct $\{P_{0im}\}_{i=1,m=1}^{I,M}$. Then identification immediately follows the same steps described in Section 3. In particular, in this order, (i) use $\{P_{0im}\}_{i=1,m=1}^{I,M}$ to identify $f_0(\cdot)$ and $F_0(\cdot)$; (ii) identify $\mathbf{q}$ from (22) (cf. Lemma 3), combine it with $\{\Upsilon_k\}_{k=1}^{I-1}$, we can identify $\{G(\Upsilon_k)\}_{k=1}^{I-1}$ (cf. Proposition 1); (iii) recover $\{R_{0im}\}_{i=1,m=1}^{I,M}$ from

$$R_{0im} = P_{0im} - \frac{\sum_{k=1}^{I} q_k k (1 - F_0(P_{0im}))^{k-1}}{f_0(P_{0im}) \sum_{k=2}^{I} q_k k (k-1)(1 - F_0(P_{0im}))^{k-2}}, \tag{23}$$

cf. (9), which in turn identifies $H_0(\cdot)$ (cf. Proposition 2).

In practice, however, we do not know $\{\nu_i\}_{i=1}^{I}$. The key insight to proceed is that optimal search behavior is determined by the shape of the equilibrium price distributions, which is the same for all firms, and not their locations that may differ. Subsequently, relative utilities are identified by relative demeaned prices. To see this, recall that $U_{im} = \nu_i - P_{im}$, we have for all i and j,

$$\begin{aligned} U_{im} - U_{jm} &= P_{0jm} - P_{0im} && (24) \\ &= \omega_{jm} - \omega_{im}, && (25) \end{aligned}$$

where ω_{im} denotes $P_{im} - \mathbb{E}[P_{im}]$, and we use the fact that $\mathbb{E}[P_{0im}] = \mathbb{E}[P_{0jm}]$ to go from (24) to (25).

Our identification results rely on the distribution of ω_{im}, which identified. We denote the pdf and cdf of ω_{im} by $w(\cdot)$ and $W(\cdot)$ respectively. Note that $F_0(\cdot)$ and $W(\cdot)$ are parallel to each other by construction. A useful relation that immediately follows from inspecting (24) and (25) is that the cdfs of P_{0im} and ω_{im} coincide when evaluated at their respective points of realizations. We state this as a lemma.

LEMMA 10. *Suppose Assumption D' holds. Then* $F_0(P_{0im}) = W(\omega_{im})$ *for all* i *and* m.

This enables us to identify the consumer search distribution.

PROPOSITION 5. *Suppose Assumption D' holds. Then* $G(\Upsilon_k)$ *is identified for* $k = 1, \ldots, I-1$.
PROOF. By Lemma 10, any $\mathbf{q}$ that satisfies (22) also satisfies

$$\mathbb{E}[Y_{im}|\omega_{im}] = \sum_{k=1}^{I} q_k \frac{k}{I} (1 - W(\omega_{im}))^{k-1}.$$

We can then identify $\mathbf{q}$ in closed-form as done in Lemma 3. From (19), we can also identify Υ_k in the same way we identify Δ_k in Section 3.1 by replacing the raw prices with the demeaned prices. We can then apply the argument used to prove Proposition 1 to identify $\{G(\Upsilon_k)\}_{k=1}^{I-1}$ from $\mathbf{q}$ and $\{\Upsilon_k\}_{k=1}^{I-1}$. ■

On the supply side, we can identify the shape of the distribution of R_{0im} but not its location. This is clear from (23) because we can only identify the shape of the distribution of P_{0im}. More precisely, what we can identify is the distribution of $\rho_{im} := R_{0im} - \mathbb{E}[P_{0im}]$.

PROPOSITION 6. *Suppose Assumption D' holds. Then the distribution of* ρ_{im} *is identified.*
PROOF. Construct ρ_{im} by replacing $(P_{0im}, f_0(P_{0im}), F_0(P_{0im}))$ in the RHS of (23) by $(\omega_{im}, w(\omega_{im}), W(\omega_{im}))$ and apply Lemma 10. ■

Our Propositions 5 and 6 show that we can use $\{\omega_{im}\}_{i=1,m=1}^{I,M}$, which is a random sample, in place of the observed prices that are heterogeneous due to $\{\nu_i\}_{i=1}^{I}$ to identify the demand and supply side parameters in the same way as done in Sections 3.1 and 3.2 respectively. The proposed estimators and convergence rates discussed in Section 4 are therefore immediately applicable to our search model with differentiated products as well.

Not knowing $\{\nu_i\}_{i=1}^{I}$ does not limit the scope of counterfactual studies relative to the symmetric model. This is because consumers in our model bear the cost of quality differences and get compensated in equal amount with utility. Thus we can study changes in search behavior and the price distribution associated with quality adjusted production costs. We can identify these effects by comparing the difference of price distributions generated from the old and new equilibria where firms are treated symmetrically such that every firm draws cost from the same distribution as ρ_{im}.

6 Numerical Studies

The purpose of this section is to numerically illustrate theoretical features of our model and discuss issues that may be relevant for applications. We consider a simple design for a model of search with $I = 3$. Consumers draw search costs from a distribution with cdf $G(c) = \sqrt{c}$ for $c \in [0, 1]$. Firms draw marginal costs from a uniform distribution on $[0, 1]$. We use the system of equations in (6) to iteratively solve for the equilibrium of the game. We use numerous random initial values to iterate from. We have found only one equilibrium that generates price dispersion, where $\mathbf{q} = (0.7852, 0.0455, 0.1693)$. We generate data from this equilibrium for 333 markets, so $IM = 999$, by drawing prices prices from (4) and market shares from (7).

We focus on the nonparametric estimators of $f(\cdot)$ and $h(\cdot)$. Estimation of other components is straightforward and the corresponding estimators are well-behaved with the parametric convergence rate. Specifically, we estimate $F(\cdot)$ and $\mathbf{q}$ using the estimators described in Section 4 that satisfy Assumption H. For $f(\cdot)$ and $h(\cdot)$, while the estimators mentioned in Section 4 are sufficient in delivering the desired convergence rate uniformly over an expanding interval in practice we may want to make use of the data outside of the interval that are closer to the boundaries if they can be well-estimated. In particular, it has been documented that the trimming procedure GPV proposed to avoid the boundary bias in estimating pdfs in a first-price auction model does not perform well in practice because trimming and the general bias from estimating the bids pdf in the first stage have direct effects on the estimated (object) valuations that are used as a generated regressor in the second stage of estimating the valuation pdf. In this context, Hickman and Hubbard (2015) suggest one uses boundary corrected kernel to avoid trimming, i.e. use all observations. Their choice for the boundary correction is based on the estimator of Karunamuni and Zhang (2008, henceforth KZ), and they show it works well in small samples (also see Li and Liu (2015) in another auction application). Our estimation problem is more challenging than the GPV setup because, other than the boundary issue, (i) for estimating $f(\cdot)$ we have a pole at the upper boundary and (ii) for estimating $h(\cdot)$ we have additional sampling errors from estimating $\mathbf{q}$ and $F(\cdot)$.

We consider three estimators for $f(\cdot)$. First, $\widehat{f}_1(\cdot)$, is a standard kernel estimator that does not account for the boundary effect or the presence of the pole. Second, $\widehat{f}_2(\cdot)$, following KZ, accounts for both the lower and upper boundaries based on KZ but not acknowledge the presence of the pole. Third, $\widehat{f}_3(\cdot)$ uses the transformation described in equation (15) to accommodate the pole and applies boundary correction at the lower boundary. We note that the boundary correction estimator of KZ is a sum of a standard kernel density estimator and term from an endpoint kernel. The latter is zero outside of a bandwidth-neighborhood of the boundaries so KZ's estimator is just a regular density estimator when evaluated at interior points. We use the Epanechnikov kernel for all of our estimators. Boundary correction uses the *optimal endpoint* kernel and associated plug-in constants and bandwidths suggested in KZ. Figures 1 to 3 plot the mean and the 5th and 95th percentiles for each $\left(\widehat{f}_1(\cdot), \widehat{f}_2(\cdot), \widehat{f}_3(\cdot)\right)$ against the true price pdf.

We see that $\widehat{f}_1(\cdot)$ performs poorly near both lower and upper support points. Boundary correction removes the bias near the lower boundary but not near the pole, we also see boundary correction makes $\widehat{f}_2(\cdot)$ more variable in the sense that its distribution is concentrated around the mean relative to $\widehat{f}_1(\cdot)$'s near the boundaries. The transformation method we suggest leads to a very good estimator near pole. The distribution of $\widehat{f}_3(\cdot)$ is also highly concentrated around its mean, but the transformation with bias correction has larger bias near the lower boundary than the non-transformed counterpart.

We next compare three estimators for $h(\cdot)$ in Figures 4 to 6. There, we plot the mean and the 5th and 95th percentiles of KZ boundary corrected estimators, $\left(\widehat{h}_1(\cdot), \widehat{h}_2(\cdot), \widehat{h}_3(\cdot)\right)$, for the marginal cost density that correspond respectively to $\left(\widehat{f}_1(\cdot), \widehat{f}_2(\cdot), \widehat{f}_3(\cdot)\right)$. These figures also include the mean and the 5th and 95th percentiles of an infeasible KZ boundary corrected estimator constructed from the estimated costs when the true $f(\cdot)$ is used; $\mathbf{q}$ and $F(\cdot)$ are still estimated. Naturally the infeasible estimator is generally a superior estimator than the feasible ones. Noting that even the infeasible estimator is also more variable closer to the boundaries and it still suffers from the boundary effect. For the feasible estimators, we see that $\widehat{h}_1(\cdot)$ performs poorly generally. In comparison, $\widehat{h}_2(\cdot)$ has lower bias over its lower half of the support that includes the lower boundary but its distribution in that region is more variable than $\widehat{h}_1(\cdot)$, and its bias at the upper half of the support is large. $\widehat{h}_3(\cdot)$ performs extremely well closer to the upper support and its distribution is concentrated around the mean over the whole support, however its bias increases as it approaches the lower boundary.

The simulation study illustrates the performance of $\widehat{h}_j(\cdot)$ inherits characteristics of $\widehat{f}_j(\cdot)$. Therefore it is clear one should account for the boundary effect at the lower support as well as the pole at the upper support. The transformation approach we propose estimates regions near the upper of

Figure 1: $\widehat{f}_1(\cdot)$ - No boundary correction or transformation

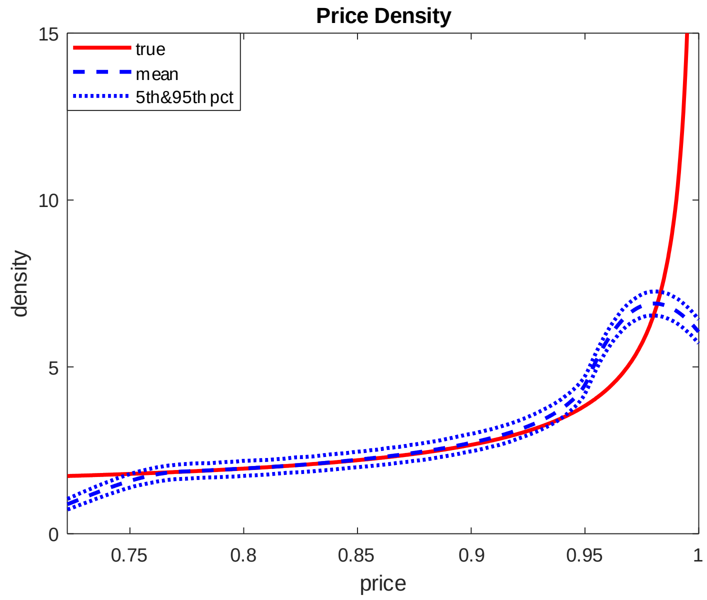


Figure 2: $\widehat{f}_2(\cdot)$ - KZ boundary correction

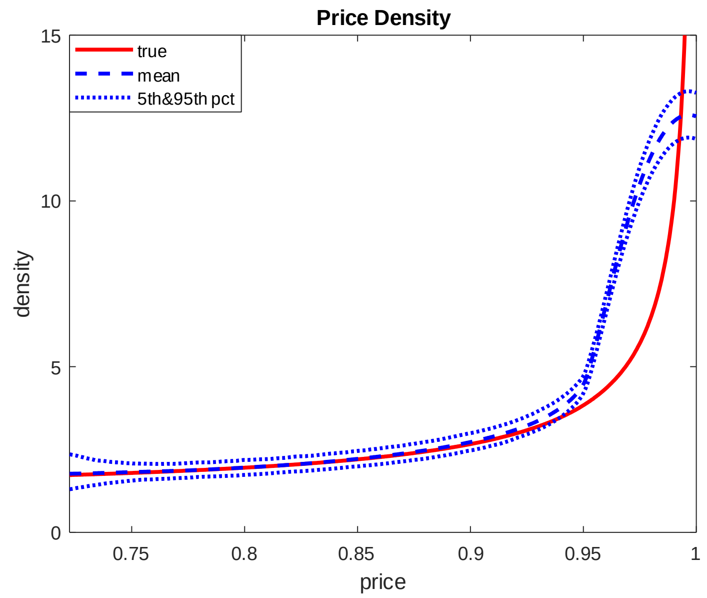


Figure 3: $\widehat{f}_3(\cdot)$ - Transformation and KZ boundary correction

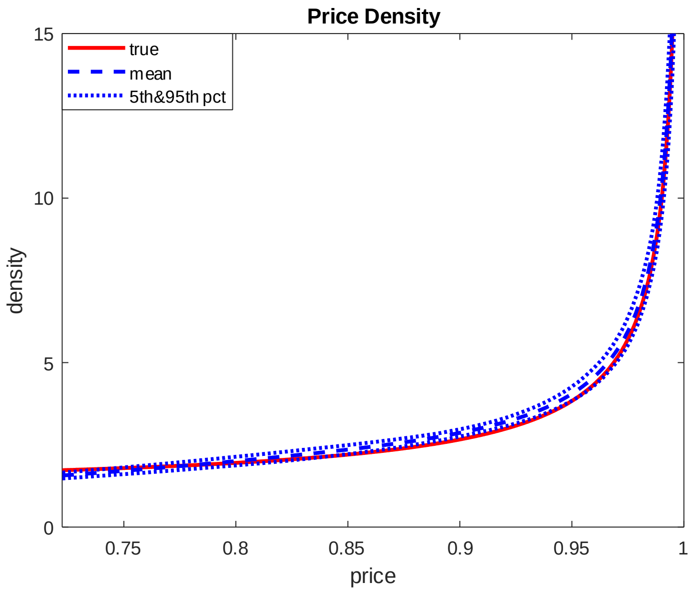


Figure 4: $\widehat{h}_1(\cdot)$ - Based on $\widehat{f}_1(\cdot)$

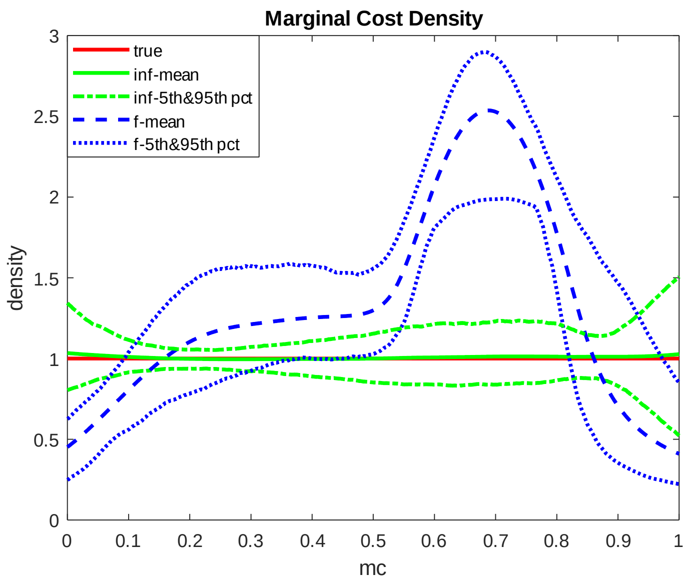


Figure 5: $\widehat{h}_2(\cdot)$ - Based on $\widehat{f}_2(\cdot)$

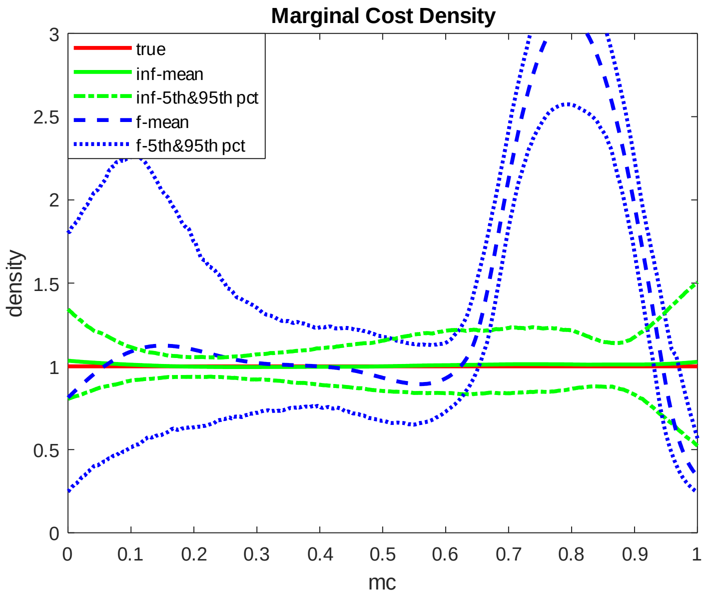


Figure 6: $\widehat{h}_3(\cdot)$ - Based on $\widehat{f}_3(\cdot)$

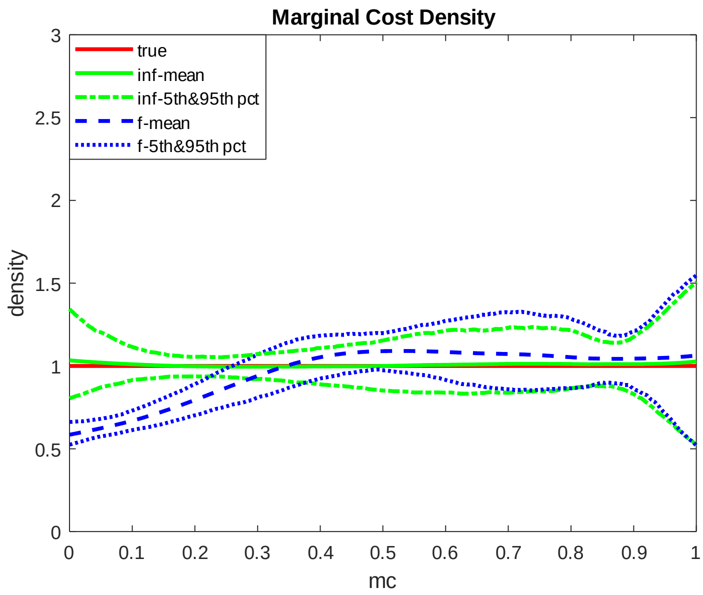
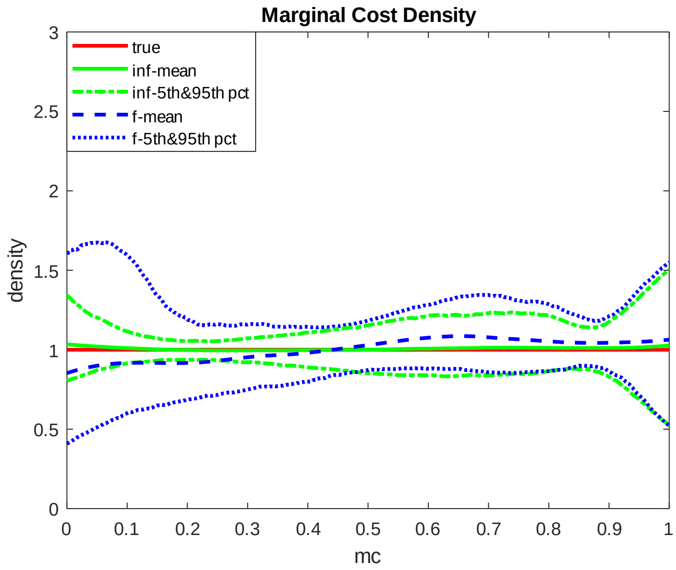


Figure 7: $\widehat{h}_4(\cdot)$ - Based on $\widehat{f}_2(\cdot)$ and $\widehat{f}_3(\cdot)$



marginal cost almost as well as the infeasible estimator because $\widehat{f}_3(\cdot)$ is a good estimator for $f(\cdot)$ near the pole, where the sampling error of its reciprocal in that region has a low order of magnitude.[10] The difficulty in estimating a nonparametric object is unavoidable for other parts of the distribution, especially in regions where price has a low density. A potentially relevant observation is that transforming the price data to estimate the pdf with boundary correction and transforming it back to get the price pdf may give a good estimator for $f(\cdot)$ near the pole but the bias correction near to lower support point can become less effective. One way to safeguard against this is to use both $\widehat{f}_2(\cdot)$ and $\widehat{f}_3(\cdot)$ to estimate parts of $h(\cdot)$. For instance, we can average the estimated costs from using $\widehat{f}_2(\cdot)$ and $\widehat{f}_3(\cdot)$ in computing (12) for observations that correspond to the lower half of the prices and use $\widehat{f}_3(\cdot)$ to estimate costs evaluated at the upper half of the prices. This leads to $\widehat{h}_4(\cdot)$. Figure 7 plots the mean and the 5th and 95th percentiles of this estimator.

7 Concluding Remarks

Hong and Shum (2006) and a series of papers by Moraga-González and Wildenbeest consider empirical models of consumer search in a complete information environment where firms have the same production cost and show how they can be identified with just observed prices alone. We propose a consumer search model with incomplete information where firms have different private costs. We characterize the equilibrium in such model and provide conditions to identify the model from price

[10] We thank Nianqing Liu for pointing this out.

and market share data. Our identification strategy is constructive and leads to natural estimators that can be computed without any optimization. Parameters on the consumer's side can be estimated at the same rate of convergence as in the complete information model. On the firm's side, the density function of the firms' production costs can be nonparametrically estimated with the same rate as the optimal rate derived in related auction models of Guerre, Perrigne and Vuong (2000) over any fixed subset in the interior of the support, and the rate can be made arbitrarily close to that if the subset is allowed to expand to the full support generally. The reason for slower convergence rate when an expanding subset is due to the potential pole at the upper support of the equilibrium price density. Irrespective whether one is interested in obtaining a convergence rate over a fixed or expanding intervals, our simulation shows the presence of the pole to have practical relevance and should be accounted for. We also present a search model with vertically differentiated products, following Wildenbeest (2011), as a way to account for systematic price differences that may exist amongst firms where the above results can be generalized to. The results from the homogeneous product model readily extends to the differentiated product case. Another way to model heterogeneity is for firms to have different probabilities of being found by consumers. Our results can also be readily extended to this case if we assume that an equilibrium exist where the optimal pricing strategies of firms are strictly increasing and share the same support. We are optimistic that such equilibrium exists based on some positive results from the literature on asymmetric first-price auctions[11]. The difficulty in establishing these features in such asymmetric search model is due to the fact that the (quasi-)inverse of the optimal pricing strategies are solutions to a system of nonlinear differential equations.[12] We leave this task for future research.

Our paper focuses on convergence rates of a search model involving products without any observed characteristics. In applications product characteristics can be readily incorporated into the model if they are available. As in the auction model of Guerre et al. (2000), however, the nonparametric rate of convergence for the conditional distributions will be slower than that of the unconditional ones. We expect it is possible for a recent approach to mitigate the dimensionality issue in the auction literature, particularly the quantile regression approach of Gimenes and Guerre (2020), can be applied to a search model like ours. Parametric assumptions can also be readily incorporated, for these we refer the readers to Myśliwski and Rostom (2020) and Salz (2020) for substantive examples of empirical modelling for closely related search models.

[11] Some existence results do exist, e.g. see Lebrun(1999) and Maskin and Riley (2000). Furthermore, a common support for the optimal bids in the first-price context is also known to hold (e.g. see Athey and Haile (2007).

[12] It is not trivial even to show existence of such equilibrium numerically. For instance, in a related problem, numerical studies of the equilibrium in asymmetric auctions is a current topic of research - e.g. see the discussion in Fibich and Gavish (2011).

Finally, we do not deal with inference in this paper. Inference on the demand side parameters is relatively straightforward, e.g. see Sanches, Silva and Srisuma (2018). Establishing the asymptotic distribution and validity for the bootstrap of $\widehat{h}(\cdot)$ and $\widetilde{h}(\cdot)$ is more challenging. We conjecture this can be obtained by suitably adapting the arguments in a recent article by Ma, Marmer and Schneyerov (2019), where they derive the asymptotic variance for the GPV's estimator as well as showing inference using the bootstrap is valid.

Appendix

This Appendix provides the proofs of Lemmas and Theorems. We omit the proofs of Lemmas 1, 2 and 10, Theorem 1, and the Propositions because these are either immediate consequences of what have discussed or proven in the main text. We also omit the proofs of Lemma 6 and Theorem 2 because they are very similar to the proofs of Lemma 9 and Theorem 3 respectively.

Proof of Lemma 3. From (8), we have $\mathbb{E}[Y_{im}|X_{im}] = X_{im}^{\top}\mathbf{q}$. Multiply both sides by X_{im} and take expectation yields $\mathbb{E}[X_{im}Y_{im}] = \mathbb{E}\left[X_{im}X_{im}^{\top}\right]\mathbf{q}$. Since $\mathbb{E}\left[X_{im}X_{im}^{\top}\right]$ has full rank, the proof follows from solving for $\mathbf{q}$. ■

Proof of Lemma 4. Under D(i), by inspecting (4) and (5), $\beta(\cdot)$ is strictly increasing and continuously differentiable on $\left[\underline{R},\overline{R}\right]$ with $\beta\left(\overline{R}\right)=\overline{R}$. Therefore $\xi(\cdot)$ exists, it is also strictly increasing on $\left[\underline{P},\overline{P}\right]$ with $\xi\left(\overline{P}\right)=\overline{R}$.

To obtain the desired expression, for any r, we know that $\beta(r)$ is the maximizer of the following function,

$$\Lambda(p,r) = (p-r)\sum_{k=1}^{I} q_k \frac{k}{I}\left(1-H(\xi(p))\right)^{k-1}.$$

$\beta(r)$ is also the zero to $\frac{\partial}{\partial p}\Lambda(p,r)$, where

$$\begin{aligned}\frac{\partial}{\partial p}\Lambda(p,r) &= \sum_{k=1}^{I} q_k \frac{k}{I}\left(1-H(\xi(p))\right)^{k-1} \\ &\quad + (p-r)\,\xi'(p)\,h(\xi(p))\sum_{k=2}^{I} q_k \frac{k(k-1)}{I}\left(1-H(\xi(p))\right)^{k-2}.\end{aligned}$$

Noting that the cdf and pdf of P_{im} and R_{im} are related through,

$$F(p) = H(\xi(p)) \quad \text{and} \quad f(p) = \xi'(p)\,h(\xi(p)). \tag{26}$$

Substitute these in and impose the first-order condition leads to

$$\sum_{k=1}^{I} q_k k\left(1-F(p)\right)^{k-1} = (p-\xi(p))\,f(p)\sum_{k=2}^{I} q_k k(k-1)\left(1-F(p)\right)^{k-2}.$$

Rearranging the relation above to make $\xi(p)$ the subject of the equation above gives (9). ■

PROOF OF LEMMA 5. Using (5), we can write $f(p) = \psi\left(\beta^{-1}(p)\right)$, where $\psi(\cdot)$ is a real-value function defined on $\left[\underline{R}, \overline{R}\right]$ such that

$$\psi(r) = \frac{\left(\sum_{k=1}^{I} q_k k (1 - H(r))^{k-1}\right)^2}{\left(\sum_{k=2}^{I} q_k k (k-1)(1 - H(r))^{k-2}\right)\left(\sum_{k=1}^{I} q_k k \int_{s=r}^{\overline{R}} (1 - H(s))^{k-1} ds\right)}. \tag{27}$$

From (5), we see that $\beta^{-1}(\cdot)$ is $\tau + 1$ times continuously differentiable on $[\underline{P}, \overline{P})$ as $\beta'(\cdot) > 0$ on $[\underline{R}, \overline{R})$. The result then follows from the fact that $\psi(r)$, see (27), is a smooth functional of $H(\cdot)$ for all $r \in [\underline{R}, \overline{R})$. ■

PROOF OF LEMMA 7. The proof can be seen from inspecting (27).

Part (a) follows from $\inf_{p \in [\underline{P}, \overline{P}]} f(p) = \inf_{r \in [\underline{R}, \overline{R}]} \psi(r) \geq \frac{q_1^2}{\left(\sum_{k=2}^{I} q_k k(k-1)\right)\left(\overline{R} \sum_{k=1}^{I} q_k k\right)} > 0.$

Part (b) follows from $\lim_{p \to \overline{P}} f(p) = \lim_{r \to \overline{R}} \psi(r) = \infty$. ■

PROOF OF LEMMA 8. Given that $f^{\dagger}\left(p^{\dagger}\right) = \exp\left(-p^{\dagger}\right) f\left(\overline{P} - \exp\left(-p^{\dagger}\right)\right)$ for all $p^{\dagger} \in \left[-\ln\left(\overline{P} - \underline{P}\right), \infty\right)$, it suffices to show $\lim_{p \to 0} p f\left(\overline{P} - p\right) = 0$, which is a fact that we will now show. For any $\underline{P} \leq p_n < p_n' \leq \overline{P}$, by the mean value theorem for definite integrals, there exists $c_n \in (p_n, p_n')$ such that

$$\int_{p=p_n}^{p_n'} f(p)\, dp = (p_n' - p_n) f(c_n) \geq \frac{(p_n' - p_n)}{\left(\overline{P} - p_n\right)}\left(\overline{P} - c_n\right) f(c_n).$$

Suppose $p_n = \overline{P} - 2^{-n+1}$ and $p_n' = \overline{P} - 2^{-n}$. Then for all n,

$$\int_{p=p_n}^{p_n'} f(p)\, dp \geq \frac{1}{2}\left(\overline{P} - c_n\right) f(c_n).$$

Since $f(\cdot)$ is a proper density and $\lim_{n \to \infty} p_n = \overline{P}$, the proof follows from $\lim_{n \to \infty} \int_{p=p_n}^{p_n'} f(p)\, dp = 0$.■

PROOF OF LEMMA 9. From (16) when $\widetilde{R}_{im} < \infty$ we can write,

$$\begin{aligned}
\widetilde{R}_{im} - R_{im} &= I_1(P_{im}) + I_2(P_{im}), \text{ where} \\
I_1(P_{im}) &= \Psi\left(\widehat{\mathbf{q}}, \widetilde{f}(P_{im}), \widehat{F}(P_{im})\right) - \Psi\left(\mathbf{q}, \widetilde{f}(P_{im}), F(P_{im})\right), \\
I_2(P_{im}) &= \Psi\left(\mathbf{q}, \widetilde{f}(P_{im}), F(P_{im})\right) - \Psi\left(\mathbf{q}, f(P_{im}), F(P_{im})\right),
\end{aligned}$$

where $\Psi(\mathbf{q}, f(P_{im}), F(P_{im})) = \frac{\sum_{k=1}^{I} q_k k (1 - F(P_{im}))^{k-1}}{f(P_{im}) \sum_{k=2}^{I} q_k k(k-1)(1 - F(P_{im}))^{k-2}}$ so that $\Psi\left(\widehat{\mathbf{q}}, \widetilde{f}(P_{im}), \widehat{F}(P_{im})\right)$ and $\Psi\left(\mathbf{q}, \widetilde{f}(P_{im}), F(P_{im})\right)$ are estimated counterparts of $\Psi(\mathbf{q}, f(P_{im}), F(P_{im}))$ where some or all components of $(\mathbf{q}, f(P_{im}), F(P_{im}))$ are replaced by $\left(\widehat{\mathbf{q}}, \widetilde{f}(P_{im}), \widehat{F}(P_{im})\right)$ accordingly. By Lemma 7(a) we know $\inf_{p\in\left[\underline{P}+\delta_M,\overline{P}-\delta_M\right]}\widetilde{f}(p)>c_0$ for some $c_0>0$ with probability approaching one as $M\to\infty$. Given the convergence rates in Propositions 3 and 4, it is straightforward to verify that the partial derivatives of $\Psi\left(\mathbf{q},\widetilde{f}(P_{im}),F(P_{im})\right)$ with respect to its first and third arguments are also almost surely uniformly bounded. Therefore by the mean value theorem it follows that

$$|I_1(P_{im})|=O_p\left(\|\widehat{\mathbf{q}}-\mathbf{q}\|+\sup_{p\in\left[\underline{P}+\delta_M,\overline{P}-\delta_M\right]}\left|\widehat{F}(p)-F(p)\right|\right)$$

So that $|I_1(P_{im})|=o(\eta_M)$ almost surely. For I_2, we can write

$$I_2(P_{im})=-\left(\frac{\widetilde{f}(P_{im})-f(P_{im})}{\widetilde{f}(P_{im})f(P_{im})}\right)\frac{\sum_{k=1}^{I}q_k k(1-F(P_{im}))^{k-1}}{\sum_{k=2}^{I}q_k k(k-1)(1-F(P_{im}))^{k-2}},$$

so that

$$|I_2(P_{im})|=O\left(\sup_{p\in\left[\underline{P}+\delta_M,\overline{P}-\delta_M\right]}\left|\widetilde{f}(p)-f(p)\right|\right)\ a.s.$$

The upper bounds for $|I_1(P_{im})|$ and $|I_2(P_{im})|$ are independent of P_{im}. The proof then follows from applying the convergence rates of the quantities in $|I_1(P_{im})|$ and $|I_2(P_{im})|$ as stated in Proposition 4.■

PROOF OF THEOREM 3. From (17),

$$\begin{aligned}
\widetilde{h}(r)-h(r) &= J_1(r)+J_2(r)+J_3(r),\ \text{where}\\
J_1(r) &= \frac{1}{MIb_{h,M}}\sum_{m=1}^{M}\sum_{i=1}^{I}\left(K\left(\frac{\widetilde{R}_{im}-r}{b_{h,M}}\right)-K\left(\frac{R_{im}-r}{b_{h,M}}\right)\right)\mathbf{1}\left[\widetilde{R}_{im}<\infty\right],\\
J_2(r) &= -\frac{1}{MIb_{h,M}}\sum_{m=1}^{M}\sum_{i=1}^{I}K\left(\frac{R_{im}-r}{b_{h,M}}\right)\mathbf{1}\left[\widetilde{R}_{im}=\infty\right],\\
J_3(r) &= \frac{1}{MIb_{h,M}}\sum_{m=1}^{M}\sum_{i=1}^{I}K\left(\frac{R_{im}-r}{b_{h,M}}\right)-h(r).
\end{aligned}$$

For J_1:

$$J_1(r)=\frac{1}{MIb_{h,M}}\sum_{m=1}^{M}\sum_{i=1}^{I}\left(K'\left(\frac{R_{im}-r}{b_{h,M}}\right)\left(\frac{\widetilde{R}_{im}-R_{im}}{b_{h,M}}\right)+\frac{1}{2}K''\left(\frac{\overline{R}_{im}-r}{b_{h,M}}\right)\left(\frac{\widetilde{R}_{im}-R_{im}}{b_{h,M}}\right)^2\right)\mathbf{1}\left[\widetilde{R}_{im}<\infty\right]$$

where $\overline{R}_{im}$ is some mid-point between $\widetilde{R}_{im}$ and R_{im}. Then we have

$$\begin{aligned}
|J_1(r)| \leq\ & \frac{\sup_{i,m\text{ s.t. }\widetilde{R}_{im}<\infty}\left|\widetilde{R}_{im}-R_{im}\right|}{b_{h,M}}\frac{1}{MIb_{h,M}}\sum_{m=1}^{M}\sum_{i=1}^{I}\left|K'\left(\frac{R_{im}-r}{b_{h,M}}\right)\right|\\
&+\frac{\left(\sup_{i,m\text{ s.t. }\widetilde{R}_{im}<\infty}\left|\widetilde{R}_{im}-R_{im}\right|\right)^2}{b_M^3}\frac{1}{2MI}\sum_{m=1}^{M}\sum_{i=1}^{I}\mathbf{1}\left[\widetilde{R}_{im}<\infty\right]\sup_{v\in\mathbb{R}}K''(v).
\end{aligned}$$

It can be shown using standard methods for kernel estimators that

$$\sup_{r\in\left[\underline{R}+\varsigma_M,\overline{R}-\varsigma_M\right]}\left|\frac{1}{MIb_{h,M}}\sum_{m=1}^{M}\sum_{i=1}^{I}\left|K'\left(\frac{R_{im}-r}{b_{h,M}}\right)\right|-h\left(r\right)\int\left|K'\left(v\right)\right|dv\right|=o\left(1\right),$$

where $\sup_{r\in\left[\underline{R},\overline{R}\right]}h\left(r\right)\int\left|K'\left(v\right)\right|dv$ is finite. Since $\left[\widetilde{R}_{im}<\infty\right]$ is an almost sure set asymptotically, $\frac{1}{MI}\sum_{m=1}^{M}\sum_{i=1}^{I}\mathbf{1}\left[\widetilde{R}_{im}<\infty\right]$ converges to 1 almost surely and,

$$\frac{1}{2MI}\sum_{m=1}^{M}\sum_{i=1}^{I}\mathbf{1}\left[\widetilde{R}_{im}<\infty\right]\sup_{v\in\mathbb{R}}K''\left(v\right)=\frac{1}{2}\sup_{v\in\mathbb{R}}K''\left(v\right)+o\left(1\right).$$

It follows that

$$\sup_{r\in\left[\underline{R}+\varsigma_M,\overline{R}-\varsigma_M\right]}\left|J_1\left(r\right)\right|\leq O\left(\frac{\eta_M}{b_{h,M}}+\frac{\eta_M^2}{b_M^3}\right).$$

When $\eta_M=O\left(b_M^2\right)$ it follows that,

$$\sup_{r\in\left[\underline{R}+\varsigma_M,\overline{R}-\varsigma_M\right]}\left|J_1\left(r\right)\right|\leq O\left(\frac{\eta_M}{b_{h,M}}\right)\ a.s.$$

For J_2, since $\left[\widetilde{R}_{im}=\infty\right]$ is a null set asymptotically and $\mathbf{1}\left[\widetilde{R}_{im}=\infty\right]=o\left(\upsilon_M\right)$, by choosing $\upsilon_M=o\left(\frac{\eta_M}{b_{h,M}}\right)$,

$$\sup_{r\in\left[\underline{R}+\varsigma_M,\overline{R}-\varsigma_M\right]}\left|J_2\left(r\right)\right|\leq o\left(\upsilon_M\right)\ a.s.$$

For J_3, it is a standard result in kernel estimation that

$$\sup_{r\in\left[\underline{R}+\varsigma_M,\overline{R}-\varsigma_M\right]}\left|J_3\left(r\right)\right|=O\left(b_{h,M}^{\tau}+\eta_M^*\right)\ a.s.$$

The bias component in J_3 is of the same order as $\frac{\eta_M^*}{b_{h,M}}=o\left(\frac{\eta_M}{b_{h,M}}\right)$ and the stochastic part is also $o\left(\frac{\eta_M}{b_{h,M}}\right)$. ■

PROOF OF THEOREM 4. It suffices to provide the best responses of consumers and firms analogous to those in Lemmas 1 and 2 respectively, and give the definition of a quasi-symmetric equilibrium. These results are stated in Lemmas 11 and 12 below. In particular, Lemma 11 replaces $\{\Delta_k\}_{k=1}^{I-1}$ in equation (3) by $\{\Upsilon_k\}_{k=1}^{I-1}$ and Lemma 12 use the distribution of random utilities based on (20) instead of prices.

LEMMA 11. *Suppose Assumption D' holds. Then the consumer's best response is a map $\sigma_D:\mathcal{L}\rightarrow\mathbb{S}^{I-1}$ such that for any L in $\mathcal{L}$,*

$$\sigma_D\left(L\right)=\begin{cases}1-G\left(\Upsilon_k\left(L\right)\right) & \text{for } k=1\\ G\left(\Upsilon_{k-1}\left(L\right)\right)-G\left(\Upsilon_k\left(L\right)\right) & \text{for } k>1\end{cases}.\qquad(28)$$

Lemma 12. *Suppose Assumption D' holds. Then the firm's best response is a map* $\sigma_S : \mathbb{S}^{I-1} \to \mathcal{L}$ *such that for any* $\mathbf{q}$ *in* $\mathbb{S}^{I-1}$, $\sigma_S(\mathbf{q})$ is the cdf of $\mu(x(R_{0i});\mathbf{q})$ *where* $\mu(x(\cdot);\mathbf{q})$ *is defined as in (20).*

We can now define a quasi-symmetric equilibrium as follows.

Definition 2 . *A pair* $(\mathbf{q}, L) \in \mathbb{S}^{I-1} \times \mathcal{L}$ *is a quasi-symmetric equilibrium if* $\mathbf{q} = \sigma_D(L)$ *and* $L = \sigma_S(\mathbf{q})$, *where* $\sigma_S(\cdot)$ *and* $\sigma_S(\cdot)$ *are defined in Lemmas 10 and 11 respectively.*

The proof of the proposition follows immediately from here. ■

References

[1] Andrews, D.W.K. (1995): "Nonparametric Kernel Estimation for Semiparametric Models," *Econometric Theory*, **11**, 560–596.

[2] Athey, S. and P. Haile (2007): "Nonparametric Approaches to Auctions," in Handbook of Econometrics, ed. by J. J. Heckman and E. Leamer, North-Holland.

[3] Baye, M. R., J. Morgan, and P. Scholten (2006): "Information, Search, and Price Dispersion," in Handbook of Economics and Information Systems, ed. by T. Hendershott, Elsevier Press, Amsterdam.

[4] Bénabou, R. (1993): "Search Market Equilibrium, Bilateral Heterogeneity, and Repeat Purchases," *Journal of Economic Theory*, **60**, 140-158.

[5] Berry, S., and P. Haile (2014): "Identification in Differentiated Prodicts Markets," *Econometrica*, **82**, 1749-1797.

[6] Bontemps, C., J.-M. Robin, and G. J. van den Berg (1999): "An Empirical Equilibrium Search Model with Continuously Distributed Heterogeneity of Workers' Opportunity Costs of Employment and Firms' Productivities, and Search on the Job," *International Economic Review*, **40**, 1039-1074.

[7] Bontemps, C., J.-M. Robin, and G. J. van den Berg (2000): "Equilibrium Search with Continuous Productivity Dispersion: Theory and Nonparametric Estimation," *International Economic Review*, **40**, 1039-1074.

[8] Burdett, K., and K. Judd (1983): "Equilibrium Price Dispersion," *Econometrica*, **51**, 955–969.

[9] Diamond, P.A. (1971): “A Model of Price Adjustment,” *Journal of Economic Theory*, **3**, 156–168.

[10] De los Santos, B., A. Hortaçsu, and M. R. Wildenbeest (2012): “Testing Models of Consumer Search Using Data on Web Browsing and Purchasing behavior,” *American Economic Review*, **102**, 2455-2480.

[11] Fan, J. and Q. Yao (2003): *Nonlinear Time Series: Nonparametric and Parametric Methods*, Springer-Verlag.

[12] Fibich, G. and N. Gavish (2011): “Numerical Simulations of Asymmetric First-Price Auctions,” *Games and Economic Behavior*, **73**, 479-495.

[13] Gimenes, N. and E. Guerre (2020): “Quantile Regression Methods for First-Price Auction: A Signal Approach,” *forthcoming in the Journal of Econometrics.*

[14] Guerre, E., I. Perrigne and Q. Vuong (2000): “Optimal Nonparametric Estimation of First-Price Auctions,” *Econometrica*, **68**, 525-574.

[15] Hansen, B. (2008): “Uniform Convergence Rates for Kernel Estimation with Dependent Data,” *Econometric Theory*, **24**, 726-748.

[16] Hansen, B. and S. Lee (2019): “Asymptotic Theory for Clustered Samples,” *Journal of Econometrics*, **210**, 268-290.

[17] Härdle, W. (1991): *Smoothing Techniques with Implementation in S.* New York: Springer Verlag.

[18] Hickman, B.R. and T.P. Hubbard (2015): “Replacing Sample Trimming with Boundary Correction in Nonparametric Estimation of First-Price Auctions,” *Journal of Applied Econometrics*, **30**, 739-762.

[19] Hong, H. and M. Shum (2006): “Using Price Distribution to Estimate Search Costs,” *RAND Journal of Economics*, **37**, 257-275.

[20] Honka, E. and P. Chintagunta (2017): “Simultaneous or Sequential? Search Strategies in the US Auto Insurance Industry,” *Marketing Science*, **36**, 21-42.

[21] Karunamuni R.J. and S. Zhang (2008): “Some Improvements on a Boundary Corrected Kernel Density Estimator,” *Statistics & Probability Letters*, **78**, 499-507

[22] Lebrun, B. (1999): “First Price Auctions in the Asymmetric N Bidder Case,” *International Economic Review*, **40**, 125-142.

[23] Li, H. and N. Liu (2015): "Nonparametric Identification and Estimation of Double Auctions with Bargaining,", *Working Paper, Shanghai University of Finance and Economics.*

[24] Ma, J., V. Marmer and A. Shneyerov (2019): "Inference for First-Price Auctions with Guerre, Perrigne, and Vuong's Estimator," *Journal of Econometrics*, **211**, 507-538.

[25] MacMinn, R.D. (1980): "Search and Market Equilibrium," *Journal of Political Economy*, **88**, 308-327.

[26] Marron, J.S. and D. Ruppert (1993): "Transformations to Reduce Boundary Bias in Kernel Density Estimation," *Journal of the Royal Statistical Society. Series B*, **56**, 653-671.

[27] Masry, E. (1996): "Multivariate Local Polynomial Regression for Time Series: Uniform Strong Consistency and Rates," *Journal of Time Series Analysis*, **17**, 571-599.

[28] McAfee, R.P. and J. McMillan (1988): "Search Mechanisms," *Journal of Economic Theory*, **44**, 99-123.

[29] Milgrom, P. and I. Segal (2002): "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, **70**, 583-601.

[30] Moraga-González, J.L. and M. Wildenbeest (2008): "Maximum Likelihood Estimation of Search Costs," *European Economic Review*, **52**, 820-48.

[31] Moraga-González, J.L., Z. Sándor and M. Wildenbeest (2012): "Consumer Search and Prices in the Automobile Market," *Working Paper, University of Indiana.*

[32] Moraga-González, J.L., Z. Sándor and M. Wildenbeest (2013): "Semi-nonparametric Estimation of Consumer Search Costs," *Journal of Applied Econometrics*, **28**, 1205-1223.

[33] Moraga-González, J.L., Z. Sándor and M. Wildenbeest (2017): "Nonsequential Search Equilibrium with Search Cost Heterogeneity," *International Journal of Industrial Organization*, **50**, 392-414.

[34] Myśliwski, M. and M. Rostom (2020): "Value of Information, Search and Competition in the UK Mortgage Market," *Working Paper, NHH.*

[35] Pereira, P. (2005): "Multiplicity of Equilibria In Search Markets with Free Entry and Exit," *International Journal of Industrial Organization*, **23**, 325-339.

[36] Postel-Vinay, F. and J.-M. Robin (2002): "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, **70**, 2295-2350.

[37] Salz, T. (2020): “Intermediation and Competition in Search Markets: An Empirical Case Study,” *Working Paper, MIT.*

[38] Sanches, F., D. Silva Junior and S. Srisuma (2018): “Minimum Distance Estimation of Search Costs using Price Distribution,” *Journal of Business and Economic Statistics*, **36**, 658-671.

[39] Stigler, G. (1961): “The Economics of Information,” *Journal of Political Economy*, **69**, 213-225.

[40] Wildenbeest, M. R. (2011): “An Empirical Model of Search with Vertically Dierentiated Products,” *RAND Journal of Economics*, **42**, 729-757.

[41] Stone, C.J. (1982): “Optimal Rate of Convergence for Nonparametric Regressions,” *Annals of Statistics*, **10**, 1040-1053.