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

Mateusz Myśliwski
University College London

Fabio Sanches
PUC-Rio

Daniel Silva Jr
City, University of London

Sorawoot Srisuma
University of Surrey

12 May 2017

Abstract

We propose a non-sequential search model with a continuum of consumers and a finite number of firms. Both consumers and firms are heterogeneous. Consumers differ in search costs. Firms have private marginal costs of production. We show that an equilibrium price dispersion can arise in this model as firms employ a Bayesian Nash pricing strategy. We provide conditions to identify the model using price and another supply side data (such as market share). Our identification strategy is constructive. We derive the uniform rate of convergence of our estimator.

JEL Classification Numbers: C14, C57, D44, D83

Keywords: Auctions, Nonparametric Methods, Search Cost

[*]We thank Guiherme Carmona, Hanming Fang, Alessandro Gavazza, Matt Gentry, Emmanuel Guerre, Tatiana Komarova, Jose Luis Moraga-González, Lars Nesheim, and Philip Reny for some very helpful comments and discussions. We also thank seminar participants at the University of Groningen and University of Surrey.

[†]*E-mails*: mateusz.mysliwski@ucl.ac.uk; fmiessi@gmail.com; danielsjunior@gmail.com; s.srisuma@surrey.ac.uk

1 Introduction

Many theoretical models have been developed to explain the price dispersion of homogeneous products relying on the notion that search is costly for consumers; see the survey of Baye, Morgan and Scholten (2005). Existing nonparametric identification results on an empirical model of search build on the *fixed sample search*[1] framework studied in Burdett and Judd (1983). The players in the games are consumers (buyers) and firms (sellers). Consumers differ by their search costs. Firms have identical costs of production. In this model firms compete by setting prices in a complete information environment. The price dispersion is generated by assuming that firms employ a mixed-strategy Nash pricing rule in equilibrium. Hong and Shum (2006) exploit the indifference condition that defines a mixed strategy and initiate a nonparametric approach to a structural search model by showing the consumer's search cost distribution can be identified from data on prices alone. When the dataset available is limited to a single market only finite points of the cost distribution can be identified. A sufficient condition for identification of the cost distribution over the whole support is possible, for instance, when we have more data of prices from different markets; see Moraga-González, Sándor and Wildenbeest (2013).

We consider a more general model where firms have heterogeneous costs of production. Our model build on the setting proposed in MacMinn (1980) where firms have independent private marginal costs. This approach leads to a game of incomplete information played between firms that resembles a procurement auction. The subsequent equilibrium solution concept is a Bayesian-Nash pure strategy. However, MacMinn only presents a partial equilibrium result as he only studies the best response of the firms' taking consumer search behaviour as given. We consider a full equilibrium model. The goal of our work is to provide a general framework for an empirical analysis of such search model.

The contribution of our paper is to provide a theoretical, both economic and econometrics, treatment for analyzing an empirical search model and give a corresponding estimation methodology. We characterize the equilibrium of our search model. We provide conditions to identify the consumers' search costs and firms' marginal costs. Finally, we propose nonparametric estimators for all of the identified objects in the model and provide some asymptotic properties of these estimators when appropriate data are available.

Our identification strategy differs from those employed to study models in the spirit of Burdett and Judd (1983). The insight of Hong and Shum (2006) uses the constancy condition imposed by a mixed strategy equilibrium to identify the distribution of consumer's search over all possible

[1] In a fixed sample search consumers decide before hand how many firms to search from. This is in contrast to a sequential search. Some recent studies have found support that nonsequential search models can better approximate consumers' search behavior (De los Santos et al. (2012), Honka and Chintagunta (2014)).

firms. We do not have such restriction to exploit with the pure strategy solution concept. Therefore, in additional to price, we require observations of another variable other than price to identify the proportions of consumer search. We then assume the regression of this variable on price to be related to the proportions of consumer search through a particular semiparametric index restriction. The index structure can be motivated from the model. These proportions are required as they appear in the firm's pricing problem. We subsequently use them to identify the distribution of the firm's marginal costs. Following our identification steps, we propose a companion two-step estimation procedure:

Step 1 The proportions of consumer search are estimated. When the index specification is linear our estimator can be computed in closed-form as an OLS estimator.

Step 2 The firms' marginal costs are estimated. These generated variables are then used to construct a nonparametric estimator for the probability density function of the marginal costs in a similar fashion to Guerre, Perrigne and Vuong (2000, hereafter GPV).

Despite its seemingly natural scope for applications[2], we are not aware of any theoretical work that considers our search model previously. We directly extend MacMinn's partial equilibrium analysis for a fixed sample search model to a full equilibrium one. We build on his insight that makes the connection between the search and procurement auction models. The pricing problem of each firm can be seen as a first price procurement auction problem with random participation; the number and identity of bidders are stochastic. We characterize an equilibrium that generates a continuous price distribution.

The model of search that is closest to the one we consider in this paper can be found in a recent empirical study by Salz (2017). A version of our model can in fact be seen as a special case of his[3]. However, as an econometric problem, our search problems are not nested. In particular his identification strategy is not applicable to our model.

Salz studies the trade-waste market in New York City. In his model buyers (consumers) can haggle (search) directly with carters (firms), or use a broker who has access to a group of carters. The haggle part is the same as our search problem. A broker acts as a clearinghouse where a standard procurement auction game with known number and identity of bidders is played. Salz assumes an equilibrium exist in his model. Importantly Salz's identification strategy relies on the assumption

[2] Our model generates price dispersion in a transparent manner through heterogenous marginal costs. A mixed strategy solution is harder to interpret. We are not aware of any purified justification for it. Even then a purification will impose some restrictions on the primitives of the game.

[3] The independent work of Salz precedes ours chronologically. We only became aware of his work during his presentation at the London School of Economics in November 2016 when his 2015 version was circulated

that a broker *always* exists; see his Assumption 1. He also assumes both carters that can be searched and those who participate with brokers have the same cost distribution[4]. Therefore he can identify the firm's cost distribution using the procurement auction data from the brokers independently of the search mechanism. The identification for the remaining components of his model subsequently relies on this.

Brokers do not exist in our model. We emphasize that we are not being critical of Salz's approach. His model captures well important features of many real world markets. Nevertheless brokers, or other *clearinghouse* facilities, are not available in many other markets. For these pure search models we show identification is possible with additional data. Our key identifying assumption involving the semiparametric index restriction is empirically motivated. It contains as a special case the assumption that some observable market outcomes are proportional to the probabilities of firms completing a sale as consumers search in expectation. Natural candidates for such variable could be market shares or sales figures. This idea is identical to linking market shares to the choice probabilities, which is the starting point for the identification argument used in the study of differentiated products markets from the IO literature (see Berry and Haile (2014)).

In terms of the econometrics, the estimation of the demand side is relatively straightforward. The estimators for the demand parameters are smooth functionals of the empirical process of observed prices and will converge at a parametric rate; cf. Sanches, Silva, Srisuma (2016). The estimation of the distribution of the firm's marginal cost is more challenging. We follow the tradition set by GPV for an nonparametric analysis of auction models and focus on density estimation, and study its uniform convergence rate. The firm's density function is the *hardest* object to estimate in our model.

We employ the same estimation strategy as GPV. We first use the observed prices to generate the latent, or pseudo-, marginal costs and then perform nonparametric estimation using the generated variables. To this end we establish some key relations between the density function of the observed and latent variables in our model. These findings are not just for theoretical interests but have important practical implications. The most crucial one is we show the density of the observed price generally asymptotes to infinity as the price approaches its upper support. Estimating a density function with a pole requires particular care as standard kernel estimation techniques are only suitable when the underlying density is assumed to be bounded on its support. For this we characterize the behavior of the price density at the upper boundary and suggest a transformation that eliminates the boundary issue (cf. Marron and Ruppert (1994)). However, a slower uniform convergence rate

[4] Salz assumes there are two types of carters. H(igh) and L(ow) cost types. Both types are present in both the broker and search markets. A carter that participates in both markets generally will bid differently during the auction and haggling process.

in the neighborhood of the pole than other part of the support is a necessary feature. We show our estimator has the same convergence rate as the GPV estimator on any compact inner subset of the support. Uniform convergence over appropriately expanding support will converge at a slower rate depending on the speed of the support expansion. We can make the convergence rate to be arbitrarily close to the optimal convergence rate derived in GPV's auction problem.

The rest of the paper proceeds as follows. Section 2 presents the model and characterizes the equilibrium of the game. Section 3 presents our constructive identification strategy. Section 4 contains the theoretical results. Section 5 discusses ideas for extensions. Section 6 presents a simulation study.

2 Model

We consider a model where there are a continuum of consumers and a finite number of firms. Each consumer has an inelastic demand for a single unit of good supplied by the firms. Consumers differ by search costs and employ a non-sequential search strategy and purchase from the firms that sell at the lowest price. We next formally introduce the elements of the game.

2.1 Supply Side

There are I firms. Let $\mathcal{I} \equiv \{1, \ldots, I\}$. Firm i draws a marginal cost of production R_i. R_i is assumed to be a continuous random variable supported on $\left[\underline{R}, \overline{R}\right] \subset \mathbb{R}$. We denote its cumulative distribution function (CDF) by $H(\cdot)$. The marginal costs of firms are independent from each other. Firm i then faces the following decision problem:

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

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

Here $\mathbf{q} = \left(q_1, \ldots, q_I\right)^\top$ denotes a vector containing $\left(q_k\right)_{k=1}^{I}$ where q_k denotes the proportion of consumers searching for k firms. For a given k, $\frac{k}{I}$ is the number of combinations that firm i gets included when k firms are sampled[5]. 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; $P_{(1:k-1)}$ denotes the minimum of such $k-1$ prices. Here we implicitly assume that all firms have equal probability of being found thus the game is symmetric. We discuss how this assumption can be relaxed in Section 5.2.

[5]Let $\mathcal{C}_k^I \equiv \frac{I!}{(I-k)!k!}$ denote 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 assume there exists a candidate for an optimal symmetric pricing strategy $\beta : [\underline{R}, \overline{R}] \to [\underline{P}, \overline{P}] \subset \mathbb{R}$ with the following properties: (i) β is strictly increasing; (ii) $\beta(\overline{R}) = \overline{R}$, which is the *free-entry* condition imposing that $\overline{P} = \overline{R}$.

Let $\mathbb{S}^{I-1}$ denote a unit simplex in $\mathbb{R}^{I+}$. For any $\mathbf{q} \in \mathbb{S}^{I-1}$, we can define $\Lambda^*(\cdot; \mathbf{q})$ 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 $\beta(\cdot; \mathbf{q})$. We denote $\beta^{-1}(\cdot; \mathbf{q})$ by $\xi(\cdot; \mathbf{q})$.

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

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

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

Thus for any r,

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

It is easy to verify that $\beta(\cdot; \mathbf{q})$ is non-decreasing. In particular $\beta(\cdot; \mathbf{q})$ is continuously differentiable with the following derivative

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

where $h(\cdot)$ denotes the probability density function (PDF) of R_i. The form of the derivative suggests that: if $q_1 = 1$ then $\beta'(r; \mathbf{q}) = 0$ for all r; otherwise $\beta(\cdot; \mathbf{q})$ will be strictly increasing almost everywhere. We shall focus on the latter case as $\beta(R_i; \mathbf{q})$ has a continuous distribution.

2.2 Demand Side

All consumers have the same valuation of the object but differ in search effort cost. Each draws a search cost c from a continuous distribution with CDF $G(\cdot)$. She decides how many firms to visit before conducting the search. Then a consumer with search cost c faces the following decision problem:

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

We use $\mathbb{E}_F[\cdot]$ to denote an expectation where the random prices have distribution described by the CDF $F(\cdot)$. As standard we assume the is no cost for the first search. We assume a purchase is always made and set the valuation of the object to be $\overline{R}$.

Consumer's Best Response

It is easy to verify that $\mathbb{E}_F\left[P_{(1:k)}\right]$ is non-increasing in k, and we have strict monotonicity when price has a non-degenerate distribution. The marginal saving from searching one more store after having searched k stores is:

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

$\Delta_k(F)$ is also non-increasing in k. When price has a continuous distribution it can be shown that

$$\Delta_k(F) = \int F(p)(1 - F(p))^k \, dp. \tag{3}$$

It then follows that the proportions of consumers searching optimally will satisfy this rule:

$$q_k(F) = \begin{cases} 1 - G(\Delta_1(F)) & \text{for } k = 1 \\ G(\Delta_{k-1}(F)) - G(\Delta_k(F)) & \text{for } k > 1 \end{cases}. \tag{4}$$

The consumer's search behavior on the demand side in our model is standard.

2.3 Equilibrium

For any $\mathbf{q} \in \mathbb{S}^{I-1}$, $\beta(\cdot;\mathbf{q})$ in (1) gives an expression for the firm's best response that induces a price distribution. Conversely, given any price CDF, $F(\cdot)$, (4) gives the consumer's best response $\mathbf{q}(F) = (q_k(F))_{k=1}^{I}$. Therefore we can define a symmetric equilibrium for our game as follows.

Definition (Symmetric Bayesian Nash equilibrium). *The pair $(\mathbf{q}, \beta(\cdot;\mathbf{q}))$ is a symmetric equilibrium if:*

(i) for every $\mathbf{q}$ when all firms apart from i use pricing strategy $\beta(\cdot;\mathbf{q})$, $\beta(\cdot;\mathbf{q})$ is a best response for firm i;

(ii) given the price distribution induced by $\beta(\cdot;\mathbf{q})$, $\mathbf{q}$ is a vector of proportions of consumers' optimal search.

For example the monopoly pricing strategy when all consumers search just once constitutes to an equilibrium with: $\beta^M\left(r;\mathbf{q}^M\right) = \overline{R}$ for all r, and $\mathbf{q}^M$ such that $q^M = 1$. However, $\left(\mathbf{q}^M, \beta^M\left(\cdot;\mathbf{q}^M\right)\right)$ does not generate any price dispersion. We will focus on an equilibrium where consumers search more than once with a positive measure. In an equilibrium where $\beta(\cdot;\mathbf{q})$, it can be characterized by $\mathbf{q}$ that satisfies (1) and (4) simultaneously. We state this as a proposition.

Proposition 1. *In an equilibrium with strictly increasing pricing strategy with an inverse function $\xi(\cdot;\mathbf{q})$, $\mathbf{q}$ satisfies the following system of equations:*

$$q_k = \begin{cases} 1 - G\left(\int H(\xi(p,\mathbf{q}))(1 - H(\xi(p,\mathbf{q})))\,dp\right) \quad \text{for } k = 1 \quad \text{otherwise,} \\ G\left(\int H(\xi(p,\mathbf{q}))(1 - H(\xi(p,\mathbf{q})))^k\,dp\right) - G\left(\int H(\xi(p,\mathbf{q}))(1 - H(\xi(p,\mathbf{q})))^{k+1}\,dp\right) \end{cases}. \tag{5}$$

The characterization above states 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 under some regularity conditions. It is clear that $\mathcal{T}$ maps $\mathbb{S}^{I-1}$ to some subset of $\mathbb{S}^{I-1}$. Therefore a general proof for an existence of an equilibrium with a price dispersion may be shown by using a fixed-point theorem, such as Brouwer's, by showing that $\mathcal{T}$ maps certain subset of $\mathbb{S}^{I-1}$ onto itself. However, it is difficult to show surjectivity in this general framework.

In subsequent sections we shall assume an existence of an equilibrium characterized by Proposition 1. We henceforth drop the indexing arguments of equilibrium objects that are made explicit in this Section for the purpose of discussions on best response; e.g. $\beta(\cdot;\mathbf{q})$ becomes $\beta(\cdot)$, $\mathbb{E}_F[\cdot]$ becomes $\mathbb{E}[\cdot]$ etc.

3 Identification

We identify the demand side first then proceed to the supply side. Our identification of the demand side focuses on $\mathbf{q}$. We assume another variable is that is related to price is available. Once we can identify $\mathbf{q}$, identification of the firm's cost distribution follows analogously to GPV.

3.1 Demand Side

Suppose we know the equilibrium price distribution of a search model. This is expected if we a random sample $\{P_{im},\}_{i=1,m=1}^{I,M}$ of prices for I firms from M markets, and we let $M \to \infty$. By assumption $P_{im} = \beta(R_{im})$. Here Y_{im} denotes an observable variable that is assumed to satisfy Assumption I below. The main identifying assumption we introduce in this paper links Y_{im} to the expected probability firm i winning the sale of the object conditioning on setting price to be P_{im}.

Assumption I. *There exists a finite and positive λ such that*

$$\mathbb{E}[Y_{im}|P_{im}] = \lambda \sum_{k=1}^{I} q_k \frac{k}{I}(1 - F(P_{im}))^{k-1}. \tag{6}$$

The expression above says: *Y_{im} is proportional to the probability firm i wins with price P_{im}.* Assumption I is analogous to the well-known assumption in the demand estimation literature in IO that equates the observed market share with the choice probabilities; e.g. as used in Berry, Levinsohn and Pakes (1995). In our case, depending on the context, candidates for Y_{im} could be market share or sales volume. The unknown λ does not prevent identification since we have the restriction that $\sum_{k=1}^{I} q_k$ must be 1. It is important to note that unlike in a discrete choice model, where the choice probabilities sums to 1, the ex-post probability $\sum_{k=1}^{I} q_k \frac{k}{I} (1 - F(P_{im}))^{k-1}$ will almost surely not sum to one across i. The role of λ in equation (6) ensures $\mathbf{q}$ can be interpreted independently from this scale. For simplicity we assume λ to be the same for all m but this is not necessary.

Let $Y_m = (Y_{1m}, \ldots, Y_{Im})^\top$ and $\mathbf{X}_m$ be a $I \times I$ matrix such that $(\mathbf{X}_m)_{ik} = \frac{k}{I} (1 - F(P_{im}))^{k-1}$. We vectorize Y_m and $\mathbf{X}_m$ across m to form: $\mathbf{Y} = \left[\mathbf{Y}_1^\top : \cdots : \mathbf{Y}_M^\top\right]^\top$ and $\mathbf{X} = \left[\mathbf{X}_1^\top : \cdots : \mathbf{X}_M^\top\right]^\top$. Then under Assumption I, we have

$$\mathbf{q} = \frac{\mathbb{E}\left[\mathbf{X}^\top \mathbf{X}\right]^{-1} \mathbb{E}\left[\mathbf{X}^\top \mathbf{Y}\right]}{\boldsymbol{\iota}^\top \mathbb{E}\left[\mathbf{X}^\top \mathbf{X}\right]^{-1} \mathbb{E}\left[\mathbf{X}^\top \mathbf{Y}\right]}, \tag{7}$$

where $\boldsymbol{\iota}$ denotes a $IM \times 1$ vector of ones. Note that $\mathbf{X}$ has full rank almost surely when P_{im} has a continuous distribution as columns in $\mathbf{X}_m$ form a polynomial basis of $\left\{(1 - F(P_{im}))^{l-1}\right\}_{l=1}^{I}$. Generally $\mathbf{q}$ is overidentified in the sense that it can be identified using $(\mathbf{Y}_m, \mathbf{X}_m)$ for any m when $F(\cdot)$ is known.

3.2 Supply Side

The optimal strategy derived in (1) relates the optimal price in terms of the latent variable. Although such expression is intuitive and natural from the theoretical analysis, it is not immediately useful for empirical purposes. (It is, however, useful for generating data in simulation studies!) We instead consider defining $\beta(\cdot)$ as a maximizer of the following function:

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

Taking a (partial) derivative of the above with respect to p gives,

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

We next use the insight from GPV by relating the distributions between the observed and unobserved variables. Particularly:

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

so that the first order condition implies

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

We then obtain the explicit form for $\beta^{-1}\left(\cdot\right)$ as,

$$\xi\left(p\right)=p-\frac{\sum_{k=1}^{I} q_k k\left(1-F\left(p\right)\right)^{k-1}}{f\left(p\right)\sum_{k=2}^{I} q_k k\left(k-1\right)\left(1-F\left(p\right)\right)^{k-2}}. \tag{8}$$

We can identify R_i from P_i, $f\left(\cdot\right)$, $F\left(\cdot\right)$ and $\left\{q_k\right\}_{k=1}^{I}$. Thus we can identify $\left\{R_i\right\}_{i=1}^{I}$ through $\left\{\xi\left(P_i\right)\right\}_{i=1}^{I}$, and subsequently identify $h\left(\cdot\right)$ with data from multiple markets.

3.3 Constructive Identification

Suppose we have a random sample for firms from multiple markets $\left\{\left(P_{im},Y_{im}\right)\right\}_{i,=1m=1}^{I,M}$. There is a natural corresponding estimation strategy by replacing unknown population quantities by sample analogs.

Estimation of $\mathbf{q}$

We first construct an estimator for $F\left(\cdot\right)$, such as the empirical CDF. We can estimate $\mathbf{q}$ using the sample counterpart of (7); by removing the expectation operators and replace $\mathbf{X}$ by its estimate $\widehat{\mathbf{X}}$ that replaces the unknown $F\left(\cdot\right)$ by some estimator $\widehat{F}\left(\cdot\right)$. Then

$$\widehat{\mathbf{q}}=\frac{\left(\widehat{\mathbf{X}}^{\top}\widehat{\mathbf{X}}\right)^{-1}\widehat{\mathbf{X}}^{\top}\mathbf{Y}}{\boldsymbol{\iota}^{\top}\left(\widehat{\mathbf{X}}^{\top}\widehat{\mathbf{X}}\right)^{-1}\widehat{\mathbf{X}}^{\top}\mathbf{Y}}.$$

Our estimator of $\mathbf{q}$ is a smooth functional of an estimator of $F\left(\cdot\right)$. Therefore $\widehat{\mathbf{q}}$ is expected to converge at the parametric rate of $\sqrt{M}$.

Estimation of $h\left(\cdot\right)$.

We first contruct an estimate for R_{im} by:

$$\widehat{R}_{im}=P_{im}-\frac{\sum_{k=1}^{I}\widehat{q}_k k\left(1-\widehat{F}\left(P_{im}\right)\right)^{k-1}}{\widehat{f}\left(P_{im}\right)\sum_{k=1}^{I}\widehat{q}_k k\left(k-1\right)\left(1-\widehat{F}\left(P_{im}\right)\right)^{k-2}}, \tag{9}$$

here $\widehat{f}(\cdot)$ and $\widehat{F}(\cdot)$ are some estimators for $f(\cdot)$ and $F(\cdot)$ respectively. We can then perform nonparametric density estimation for $h(\cdot)$ with $\left\{\widehat{R}_{im}\right\}_{i=1,m=1}^{I,M}$. When we estimate $f(\cdot)$ and $F(\cdot)$ nonparametrically it is expected that the rate of convergence of $\widehat{R}_{im}$ (and subsequently the estimator of $h(\cdot)$) will be determined by $\widehat{f}(\cdot)$; both $\widehat{\mathbf{q}}$ and $\widehat{F}(\cdot)$ converge at a faster rate.

4 Main Results

We present two Theorems. Theorem 1 shows that the theoretical search model imposes testable restrictions on the distribution of the observed prices. Theorem 2 gives a convergence rate for $\widehat{h}(\cdot)$.

4.1 Nonparametric Restrictions on the Data

Let $\mathcal{P}$ denote the set of strictly increasing CDFs with support in $\mathbb{R}$. Let $\mathbf{F}(\cdot)$ denote the joint CDF of equilibrium prices.

THEOREM 1. *Let $I \geq 2$. Let $\mathbf{F}(\cdot) \in \mathcal{P}^I$ with support $\left[\underline{P}, \overline{P}\right]^I$. There exists a distribution of marginal cost with CDF $H(\cdot)$, with an increasing CDF $H(\cdot) \in \mathcal{P}$ such that $\mathbf{F}(\cdot)$ is the joint CDF of the equilbrium prices in the search model if and only if:*

C1. $\mathbf{F}(p_1, ..., p_K) = \prod_{i=1}^{I} F(p_i)$;

C2. *The function $\xi(\cdot)$ defined in (8) is strictly increasing on $[\underline{P}, \overline{P}]$, and its inverse is differentiable on $[\underline{R}, \overline{R}] = \left[\xi(\underline{P}), \xi\left(\overline{P}\right)\right]$.*

Moreover, when $H(\cdot)$ exists, it is unique with support $[\underline{R}, \overline{R}]$ and satisfies $H(r) = F\left(\xi^{-1}(r)\right)$ for all $r \in [\underline{R}, \overline{R}]$. In addition, $\xi(\cdot)$ is the quasi-inverse of the equilibrium strategy in the sense that $\xi(p) = \beta^{-1}(p)$ for all $r \in [\underline{P}, \overline{P}]$.

Our Theorem 1 is analogous to Theorem 1 in GPV.

4.2 Large Sample Properties

In order to study the rate of convergence of our estimators we need to know some regularity properties of the objects to be estimated. We begin with some regularity assumptions on the distribution of the underlying cost.

ASSUMPTION A.

(i) For any observe price P: there exists R such that

$$P = R + \frac{\sum_{k=1}^{I} q_k k \int_{s=R}^{\overline{R}} (1 - H(s))^{k-1} ds}{\sum_{k=1}^{I} q_k k (1 - H(R))^{k-1}},$$

for $\mathbf{q}$ *that satisfies Proposition 1, and there is an observable* Y *that satisfies Assumption I;*

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

The equilibrium restrictions imply the following properties for the observed price distribution.

PROPOSITION 2. *Under Assumption A:*

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

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

(iii) $\lim_{p \to \overline{P}} f(p) = \infty$, *furthermore* $0 < \lim_{p \to \overline{P}} \frac{f(p)}{(\overline{P}-p)^{-1}} < \infty$;

(iv) $F(\cdot)$ *admits upto* $\tau + 1$ *continuous derivatives on* $\left[\underline{P}, \overline{P}\right]$;

(v) $f(\cdot)$ *admits upto* $\tau + 1$ *continuous derivatives on* $\left(\underline{P}, \overline{P}\right)$.

The findings we want to highlight here are (iii) and (v). The former reveals that $f(\cdot)$ has a pole at the upper boundary. Kernel density estimation in a neighborhood of a *pole* has to be treated with care (e.g. see Section 5 in Marron and Ruppert (1994)). We suggest a transformation to deal with this issue below.[6] The latter suggests that the implied observed PDF is smoother than the latent PDF; similar findings are also found in GPV based on the same rationale by an inspection of (i).

Suppose we have data $\{(P_{im}, Y_{im})\}_{i=1,m=1}^{I,M}$. We assume to have some preliminary estimators for $\mathbf{q}$, $F(\cdot)$, and $f(\cdot)$ that converge to zero at some rates as $M \to \infty$. Let $\eta_{0,M} = \left(\frac{\log M}{M}\right)^{\frac{\tau+1}{2\tau+3}}$. So that $\eta_{0,M}$ is the optimal rate of convergence for density estimation with $\tau + 1$ continuous derivatives (see Stone (1982)).

ASSUMPTION B. *Suppose* $\{(P_{im}, Y_{im})\}_{i=1,m=1}^{I,M}$ *satisfies Assumption A. There exists estimators:* $\widehat{\mathbf{q}}$, $\widehat{F}(\cdot)$, *and* $\widehat{f}(\cdot)$ *such that:*

(i) $\|\widehat{\mathbf{q}} - \mathbf{q}\| = O\left(1/\sqrt{M}\right)$ *a.s.;*

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

(iii) *For any positive sequence* ε'_M *that decreases to* 0 *there exists some positive sequence* δ'_M *that decreases to zero such that* $\sup_{p \in [\underline{P}+\delta'_M, \overline{P}-\delta'_M]} \left|\widehat{f}(p) - f(p)\right| = o\left(\frac{\eta_{0,M}}{\varepsilon'_M}\right)$ *a.s.;*

[6] There are also other auction models that have unbounded densities. E.g. in a first price auction with a reserve price (see GPV) and in models with selective entry (see Gentry, Li and Lu (2015)).

(iv) There exist some positive sequences $\{\delta_M\}$ *and* $\{\eta_M\}$ *that decrease to zero such that* $\eta_{0M} = o(\eta_M)$, $\sup_{p\in[\underline{P}+\delta_M,\overline{P}-\delta_M]}\left|\widehat{f}(p) - f(p)\right| = O(\eta_M)$ *a.s.*

Estimators for $\mathbf{q}$ and $F(\cdot)$ that converge at a parametric rate are going to be available under weak conditions. We will focus on the uniform convergence properties of a kernel estimator for $\widehat{f}(\cdot)$. Studying uniformity over the entire support of P_{im} is difficult as it has a compact support. It is well-known that kernel estimators have problems at (and near) the boundaries; e.g. see Chapter 2.11 in Wand and Jones (1990). On the other hand if we consider any fixed inner subset of $[\underline{P},\overline{P}]$ then a kernel density estimator can achieve the convergence rate $\eta_{0,M}$ under standard constructions. For example by using a $\tau+1$ order kernel and set bandwidth to be proportional to $b_{0,M} \equiv \left(\frac{\log M}{M}\right)^{\frac{1}{2\tau+3}}$; see Härdle (1991). But these rates cannot be maintained when we allow the support to expand to $[\underline{P},\overline{P}]$ as sample size grows. Existing results on the uniform convergence rates for kernel estimators over expanding supports assume densities are bounded (e.g. see Masry (1996) and Hansen (2008)). They are therefore not immediately applicable to us due to the pole at $\overline{P}$.

Assumption B(iii) says that any decreasing function of M converging to zero slower than $\eta_{0,M}$ can serve as an *upper bound* for $\sup_{p\in[\underline{P}+\delta'_M,\overline{P}-\delta'_M]}\left|\widehat{f}(p) - f(p)\right|$ for some $\delta'_M = o(1)$. This is possible, for instance, with a kernel estimator using a transformation method. From Proposition 2(iii) we know $f(p)$ behaves similarly to $\left(\overline{P}-p\right)^{-1}$ for p close to $\overline{P}$. Then let us consider $P^{\dagger}_{im} \equiv -\ln\left(\overline{P}-P_{im}\right)$. The support of $P^{\dagger}_{im}$ is $[-\ln\left(\overline{P}-\underline{P}\right),\infty)$. Denote the PDF of $P^{\dagger}_{im}$ by $f^{\dagger}(\cdot)$. By a change of variable, we have,

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

Then it follows that $f^{\dagger}(\cdot)$ is bounded and, in particular, $f^{\dagger}\left(-\ln\left(\overline{P}-p\right)\right)$ is flat as $p\to\overline{P}$. Furthermore it has the same smoothness as $f(\cdot)$[7]. Consider the following estimators,

$$\begin{aligned}\widehat{f}(p) &= \frac{\widehat{f^{\dagger}}\left(-\ln\left(\overline{P}-p\right)\right)}{\overline{P}-p}, \text{ where}\\ \widehat{f^{\dagger}}\left(p^{\dagger}\right) &= \frac{1}{MIb^{\dagger}_M}\sum_{m=1}^{M}\sum_{i=1}^{I}K\left(\frac{P^{\dagger}_{im}-p^{\dagger}}{b^{\dagger}_M}\right) \quad \text{for any } p^{\dagger},\end{aligned}$$

and $K(\cdot)$ is a kernel function with a bandwidth $b^{\dagger}_M$. Thus it can be shown that $\widehat{f^{\dagger}}(\cdot)$ converges uniformly at rate $\eta_{0,M}$ over some expanding support when we use a $\tau+1$ higher order kernel coupled with bandwidth $b_{0,M}$. The division by $\overline{P}-p$ slows down the rate of convergence for $\widehat{f}(\cdot)$ at the upper boundary. This can be controlled to be as slow as we like by letting δ'_M go to zero slowly. There is also a bias issue at the lower boundary. This can be avoided by setting $b^{\dagger}_M = o(\delta'_M)$.

[7] For any $p^{\dagger}\in[-\ln\left(\overline{P}-\underline{P}\right),\infty)$, $f^{\dagger}\left(p^{\dagger}\right) = \exp\left(-p^{\dagger}\right)f\left(\overline{P}-\exp\left(-p^{\dagger}\right)\right)$.

Assumption B(iv) then assumes an existence of an estimator for $f(\cdot)$ that converges uniformly over $\left[\underline{P}+\delta_M, \overline{P}-\delta_M\right]$ at an *achievable rate* η_M. We can extend the argument given for B(iii) and make η_M arbitrarily close to $\eta_{0,M}$. More specifically, we can set $\delta_M = \overline{P} - \varepsilon_M$ for some decreasing positive sequence $\{\varepsilon_M\}$ such that $b_{0,M} = o(\delta_M)$. Then B(iv) holds with $\eta_M = \frac{\eta_{0M}}{\varepsilon_M}$.

Now that we have some estimators that satisfy Assumption B, we turn to $\widehat{R}_{im}$ as defined in equation (9). We shall use a modified version of $\widehat{R}_{im}$ for the second stage estimation since we only have the desired uniform convergence rate for $\widehat{f}(\cdot)$ over an expanding support. For some positive sequence $\{\delta_M\}$ that decrease to zero, let

$$\widetilde{R}_{im} = \begin{cases} \widehat{R}_{im} \text{ for } P_{im} \in \left[\underline{P}+\delta_M, \overline{P}-\delta_M\right] \\ +\infty \qquad \text{otherwise} \end{cases}. \tag{10}$$

When $\widetilde{R}_{im} < \infty$, $\widetilde{R}_{im}$ is a smooth function of $\widehat{\mathbf{q}}, \widehat{F}(\cdot)$ and $\widehat{f}(\cdot)$. Therefore we can obtain its convergence rate that is determined by $\sup_{p\in\left[\underline{P}+\delta_M,\overline{P}-\delta_M\right]}\left|\widehat{f}(p) - f(p)\right|$.

Lemma 1. *Under Assumptions A and B, for the same $\{\delta_M\}$ and $\{\eta_M\}$ in B(d),*

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

We define explicitly a kernel estimator for $h(\cdot)$ here:

$$\widehat{h}(r) = \frac{1}{MIb_M}\sum_{m=1}^{M}\sum_{i=1}^{I} K\left(\frac{\widetilde{R}_{im} - r}{b_M}\right) \quad \text{for any } r.$$

As before, $K(\cdot)$ is a kernel function with a bandwidth b_M. We can use Lemma 1 to quantify the estimation error that arises from using $\widetilde{R}_{im}$ instead of R_{im}, and obtain the convergence rate for $\widehat{h}(\cdot)$.

Theorem 2. *Under Assumptions A and B, and for the same $\{\delta_M\}$ and $\{\eta_M\}$ as in B(d), let: (i) $K(\cdot)$ be a symmetric $(\tau+1)-$order kernel with support $[-1,1]$; (ii) $K(\cdot)$ is twice continuously differentiable on $[-1,1]$; (iii) $\{b_M\}$ for some positive real numbers decreasing to zero such that $\delta_M = O(b_M)$. Then for any sequence $\{\varsigma_M\}$ of positive real numbers decreasing to zero such that $b_M = o(\varsigma_M)$,*

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

Theorem 2 shows that $\widehat{h}(\cdot)$ converges at a slower rate than $\widehat{f}(\cdot)$ by a factor of b_M^{-1}. We have argued that the convergence rate for the latter can be made arbitrarily close to $\eta_{0,M}$. Therefore choosing an appropriate choice of b_M will ensure $\widehat{h}(\cdot)$ converge uniformly at a rate arbitrarily close to $\frac{\eta_{0,M}}{b_{0M}} = \left(\frac{\log M}{M}\right)^{\frac{\tau}{2\tau+3}}$, which is the optimal rate of convergence for a related density function derived in Theorem 3 of GPV.

5 Possible Extensions

We briefly discuss how to extend our model and methodology. First we generalize Assumption I by allowing for possibly nonparametric relation between Y_i and the probability that firm i wins the sale with price P_i. Then we consider an asymmetric game where firms have different probabilities of being found.

5.1 Relaxing Assumption I

We anticipate that Assumption I will be the most convenient in applications. However, the mathematical structure of the search problem is conducive for a nonparametric generalization. In what follows let $\mathbf{x}_{im}$ be a $I \times 1$ vector such that $(\mathbf{x}_{im})_k = \frac{k}{I}(1 - F(P_{im}))^{k-1}$.[8]

Assumption I'. *There exists a function* $\phi : R \to R$ *such that*

$$\mathbb{E}[Y_{im}|P_{im}] = \phi\left(\mathbf{x}_{im}^{\top}\mathbf{q}\right). \tag{11}$$

Assumption I is a parametric special case of Assumption I' when $\phi(\cdot)$ is an identity function multiplied by an unknown scale. More generally Assumption I' only imposes that: *Y_i is a (possibly unknown) function of the probability firm i wins with price P_{im}.* When $\phi(\cdot)$ is parametrically specified, whether $\mathbf{q}$ is identifiable depends on the parametric specification. A sufficient, but not necessary, condition for identification is strict monotonicity of $\phi(\cdot)$.[9] When $\phi(\cdot)$ is unknown (11) imposes a semiparametric index restriction. Ichimura (1993, Theorem 4.1) provides a set of conditions for identification of an index model like ours. Note that we cannot apply, at least without any modification, the average derivative argument of Powell, Stock and Stoker (1989) to identify $\mathbf{q}$ as our model does not satisfy their boundary conditions (see their Assumption 2). When $\mathbf{q}$ is identified, regardless whether $\phi(\cdot)$ is known or not, we would expect the estimator for $\mathbf{q}$ to converge sufficiently fast to not affect the convergence rate for $\widehat{f}(\cdot)$ and subsequently $\widehat{h}(\cdot)$ under general conditions.

[8] In principle we can also allow $\mathbf{w}_{im}$ to be other known functions of $\{P_{im}\}_{i=1}^{I}$. But $\mathbf{q}$ has a structural meaning so it is natural to use powers of the price hazard functions as in Assumption I.

[9] Let $\phi^{-1}(\cdot)$ denote the inverse of $\phi(\cdot)$. Given that $\mathbb{E}\left[\mathbf{x}_{im}\mathbf{x}_{im}^{\top}\right]$ has full rank a sufficient we can write $\mathbf{x}_{im}^{\top}\mathbf{q} = \phi^{-1}(\mathbb{E}[Y_{im}|P_{im}])$, so that

$$\mathbf{q} = \frac{\mathbb{E}\left[\mathbf{x}_{im}\mathbf{x}_{im}^{\top}\right]^{-1}\mathbb{E}\left[\mathbf{x}_{im}\phi^{-1}(\mathbb{E}[Y_{im}|P_{im}])\right]}{\boldsymbol{\iota}^{\top}\mathbb{E}\left[\mathbf{x}_{im}\mathbf{x}_{im}^{\top}\right]^{-1}\mathbb{E}\left[\mathbf{x}_{im}\phi^{-1}(\mathbb{E}[Y_{im}|P_{im}])\right]}.$$

5.2 Asymmetric Search Probabilities

Consider a situation when firms have different probabilities of being searched. When a consumer sets out to visit k firms, for $\ell_i \in \{1, \ldots, I\}$, we denote the probability that the set of firms $\{\ell_1, \ldots, \ell_k\}$ get visited by $\omega_{\ell_1 \ldots \ell_k}$. Since there is no need to keep track of different permutations of the same combination of firms, we only define $\omega_{\ell_1 \ldots \ell_k}$ for $\ell_1 < \ldots < \ell_k$. Let $\mathcal{I}_k \equiv \{\{\ell_1, \ldots, \ell_k\} : \ell_j \in \mathcal{I} \text{ and } \ell_j < \ell_{j+1} \text{ for all } j\}$, and $\mathcal{I}_k^i \equiv \{\{\ell_1, \ldots, \ell_k\} \in \mathcal{I}_k : \ell_j = i \text{ for some } j\}$. I.e. $\mathcal{I}_k$ is the set of indices for all combinations of k firms. $\mathcal{I}_k^i$ is the set of indices for all combinations of k firms that always include firm i. Let $\mathcal{C}_k^I \equiv \frac{I!}{(I-k)!k!}$ denote the combinatorial number from choosing k objects from a set of I. Note that $\mathcal{I}_k$ and $\mathcal{I}_k^i$ have cardinality $\mathcal{C}_k^I$ and $\mathcal{C}_{k-1}^I$ respectively. Note that:

$$\omega_{\ell_1 \ldots \ell_{i-1} \ell_{i+1} \ldots \ell_k} = \sum_{\{\ell_1, \ldots, \ell_k\} \in \mathcal{I}_k^i}^{I} \omega_{\ell_1 \ldots \ell_k} \text{ for all } i, k.$$

Using a similar argument to previously, in equilibrium it can be shown that the optimal pricing strategy for firm i, $\beta_i(\cdot)$ becomes:

$$\beta_i(r) = r + \frac{\sum_{k=1}^{I} q_k \sum_{\{\ell_1,\ldots,\ell_k\} \in \mathcal{I}_k^i} \omega_{\ell_1 \ldots \ell_k} \int_{s=r}^{\overline{R}} \prod_{j:1 \le j \le k,\ \ell_j \ne i} \left(1 - H\left(\xi_{\ell_j}\left(\beta_i(s)\right)\right)\right) ds}{\sum_{k=1}^{I} q_k \sum_{\{\ell_1,\ldots,\ell_k\} \in \mathcal{I}_k^i} \omega_{\ell_1 \ldots \ell_k} \prod_{j:1 \le j \le k,\ \ell_j \ne i} \left(1 - H\left(\xi_{\ell_j}\left(\beta_i(r)\right)\right)\right)},$$

over the region of r where $\beta_i(\cdot)$ is strictly increasing[10]. Here $\xi_j(\cdot)$ denotes the inverse of $\beta_j(\cdot)$. It is clear that we have asymmetric pricing functions that have been induced by differing probabilities of being searched.

We can also write down the inverse function for p that corresponds to where $\beta_i(\cdot)$ is strictly increasing,

$$\xi_i(p) = p - \frac{\sum_{k=1}^{I} q_k \sum_{\{\ell_1,\ldots,\ell_k\} \in \mathcal{I}_k^i} \omega_{\ell_1 \ldots \ell_k} \prod_{j:1 \le j \le k,\ \ell_j \ne i} \left(1 - F_{\ell_j}(p)\right)}{\sum_{k=2}^{I} q_k \sum_{\{\ell_1,\ldots,\ell_k\} \in \mathcal{I}_k^i} \omega_{\ell_1 \ldots \ell_k} \left(\sum_{j:1 \le j \le k,\ \ell_j \ne i} f_{\ell_j}(p) \prod_{j':1 \le j' \le k,\ \ell_{j'} \ne i,\ \ell_{j'} \ne j} \left(1 - F_{\ell_{j'}}(p)\right)\right)}.$$

We can extend Assumption I (and I') accordingly and replicate our earlier identification strategies. Particularly, we will need

$$\mathbb{E}[Y_i | P_i] = \sum_{k=1}^{I} q_k \sum_{\{\ell_1,\ldots,\ell_k\} \in \mathcal{I}_k^i} \omega_{\ell_1 \ldots \ell_k} \int \prod_{j:1 \le j \le k,\ \ell_j \ne i} \left(1 - F_{\ell_j}(p)\right) dF_i(p).$$

[10] The support of optimal prices now differ between (some) firms.

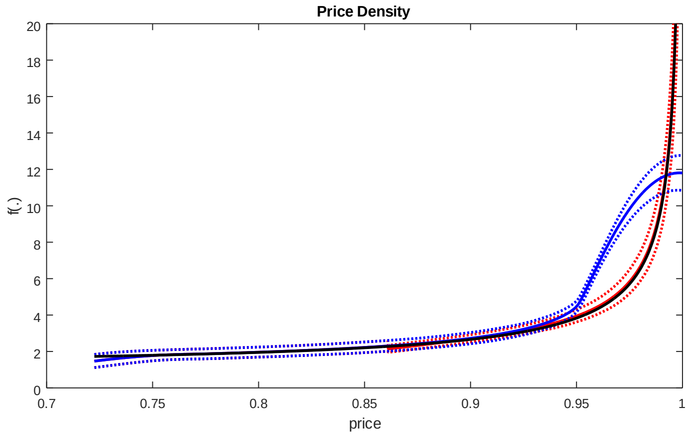


Figure 1:

6 Simulation

We consider a simple design for a game of search with three firms. 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 (5) to solve for the equilibrium of the game.

We generate the data by drawing prices from (1) with $\mathbf{q} = (0.7852, 0.0455, 0.1693)$. We generate Y_i according to Assumption I with $\lambda = 1$. We generate the data for 333 markets, so $IM = 999$. We follow the estimation strategy described in Section 3.3. In particular we the empirical price CDF to estimate $\widehat{F}(\cdot)$. We employ different bias correction and transformation techniques to estimate the densities. For the bias correction we use the procedure proposed in Karunamuni and Zhang (2008, henceforth KZ) that have recently been shown to be effective when applied to auction models (see Hickman and Hubbard (2015), and Li and Liu (2015)). We use the Epanechnikov kernel along with the forms of the plug-in bandwidths suggested in KZ. Since KZ's technique does not accommodate unbounded densities we also use the transformation we suggested in Section 4.2 to address the upper support. We combine it with the KZ's estimator to correct the lower support. We repeat the experiment 10000 times. The parametric estimators work extremely well. We only show the graphs for the density estimation.

Figure 1 shows the true price density, and the mean, 2.5th and 97.5th percentiles (percentiles using dotted lines) of the boundary corrected kernel estimator of KZ (in blue) and the kernel estimator that transforms the data to deal with the pole (in red). It is clear that standard boundary correction procedure will not be sufficient to deal with unbounded densities. On the other hand the transformation method seems to serve the purpose very well.

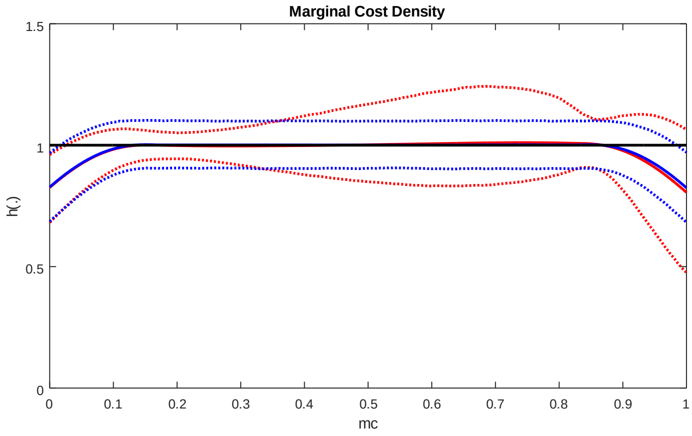
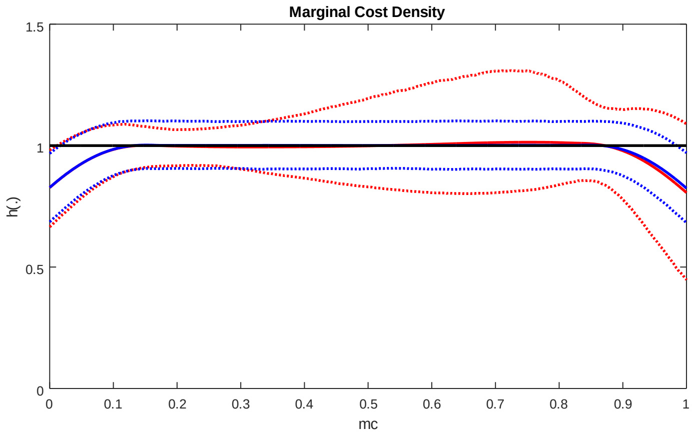
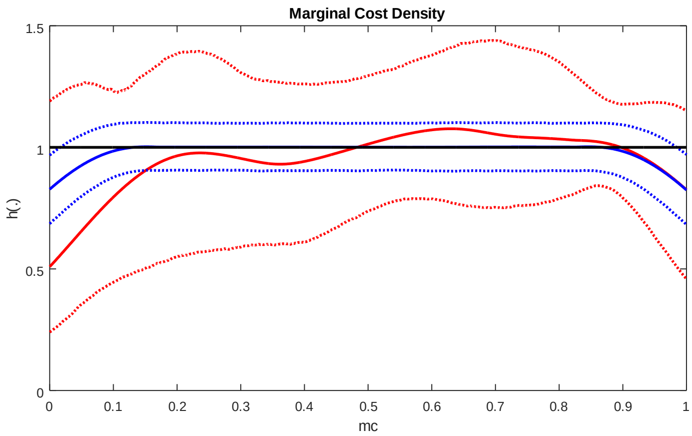

Marginal Cost Density
h(.)
mc
0
0.5
1
1.5
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Figure 2:


Marginal Cost Density
h(.)
mc
0
0.5
1
1.5
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1


Marginal Cost Density
h(.)
mc
0
0.5
1
1.5
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

We next consider three similar plots of the density estimation of marginal cost PDF (using KZ). In Figure 2 - 4, we use the true R_{im} to estimate the density (in blue) as the benchmark. The other density estimators (in red) in other figures contain estimated components. Those in Figures 2 and 3 are also infeasible as they estimate R_{im} using the unknown $f(\cdot)$: the former only estimates $\mathbf{q}$ and the latter in addition estimates $F(\cdot)$. The result for the feasible estimator using $\widetilde{R}_{im}$ as defined in (10) is in Figure 4. Again, we plot the mean and the percentiles using solid and dotted lines respectively.

Note that the boundary correction method of KZ does not completely eliminate the bias at the boundary even for the estimator that uses R_{im}. This is expected. There is in fact some improvements since density estimation without any bias correction will, in this case, converge to 0.5 at both boundaries. The mean of the bandwidth use in these figures is around 0.17, and the estimator performs much better in the interior of the support away from the boundary by at least a bandwidth. Figures 2 - 4 also show that the main source of estimation error can be traced to the estimation of the price PDF. This is not unexpected given that the PDF is the most difficult object to estimate in the entire problem.

7 Concluding Remarks

Hong and Shum (2006) and a series of papers by Moraga-González et al. show that we can identify the demand side of the market using just observed prices alone. We show when other market data, such as market shares, are available we can allow firms to be heterogenous and identify the supply side as well.

We characterize the equilibrium in a search game with heterogenous consumers and firms that supports price dispersion. We provide conditions to identify the model and propose a way to estimate the model primitives. We show that the density of the unobserved marginal cost can be estimated to converge at an arbitrary close to, but not achieving, the optimal rate derived in related auction models (such as Guerre, Perrigne and Vuong (2000)). The reason can be traced to the fact that the density of the equilibrium price has a pole at the upper support.

References

[1] Berry, S., J. Levinsohn and A. Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, **63**, 841-890.

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

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

[4] Gentry, M., T.Li and J. Lu (2015): "Identification and Estimation in First-Price Auctions with Risk-Averse Bidders and Selective Entry," *Working Paper, LSE.*

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

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

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

[8] 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.

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

[10] Hortaçsu, A., and C. Syverson (2004): "Product Differentiation, Search Costs, and Competition in the Mutual Fund Industry: A Case Study of S&P 500 Index Funds,", *Quarterly Journal of Economics*, May 2004

[11] Ichimura, H. (1993): "Semiparametric Least Squares (SLS) and Weighted SLS Estimation of Single-Index Models," *Journal of Econometrics*, **58**, 71–120.

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

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

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

[15] Marron, J.S. and D. Ruppert (1994): "Transformations to Reduce boundary Bias in Kernel Density Estimation," *Journal of the Royal Statistical Society Series B*, **56**, 653-671.

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

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

[18] Moraga-González, J.L. and M. Wildenbeest (2007): "Maximum Likelihood Estimation of Search Costs," *European Economic Review*, **52**, 820-48.

[19] Moraga-González, J.L., Z. Sándor and M. Wildenbeest (2010): "Nonsequential Search Equilibrium with Search Cost Heterogeneity," *Working Paper, University of Indiana.*

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

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

[22] Powell, J., J. Stock and T. Stoker (1989): "Semiparametric Estimation of Index Coefficients," *Econometrica*, **57**, 1403–1430.

[23] Salz, T. (2017): "Intermediation and Competition in Search Markets: An Empirical Case Study," *Working Paper, Columbia University*

[24] Sanches, F., D. Silva Junior and S. Srisuma (2016): "Minimum Distance Estimation of Search Costs using Price Distribution," *forthcoming in Journal of Business and Economic Statistics.*

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

[26] Wand, M.P. and M.C. Jones (1999): *Kernel Smoothing.* Chapman & Hall/CRC Monographs on Statistics & Applied Probability.