Identification and Estimation of a Search Model with Heterogeneous Consumers and Firms^{*†}

Mateusz Myśliwski NHH

May Rostom Bank of England

Fabio Sanches

Daniel Silva Jr

Sao Paulo School of Economics - FGV

City, University of London

Sorawoot Srisuma National University of Singapore

September 2024

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 first show the distribution of search cost can be identified by market shares and prices. Subsequently, we identify the production cost distribution using a similar strategy to Guerre, Perrigne and Vuong (2000) as the firms' decision problems resemble bidders' problems in a particular procurement auction. We prove the firm's cost density can be estimated at the same

[†]*E-mails*: mateusz.mysliwski@nhh.no; may.rostom@bankofengland.co.uk fmiessi@gmail.com; danielsjunior@gmail.com; s.srisuma@nus.edu.sg

^{*}This paper was previously circulated under the title *Identification and Estimation of a Search Model: a Procurement Auction Approach.* It also superscedes *Value of Information, Search and Competition in the UK Mortgage Market* by Myśliwski and Rostom (2022). We are grateful to the Editor, Associate Editor, and the referees for constructive comments and suggestions. We also 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, Áureo de Paula, Joris Pinkse (discussant), Fabien Postel-Vinay, Morten Ravn, Philip Reny, Xiaoxia Shi, Mikkel Sølvsten, Pai Xu and seminar participants at numerous universities and conferences for helpful comments and discussions. Any views expressed are solely those of the authors and do not represent the views of the Bank of England or have any relations to the Bank of England Policy. All errors are our own.

convergence rate as the optimal rate in Guerre et al. uniformly over any fixed subset on the interior of the support. The uniform convergence rate over any expanding support is slower due to a pole in the price pdf that is a feature of the equilibrium. Our simulation study confirms the theoretical features of the model. Our identification and convergence rate results also apply to two generalizations of the baseline search model that allow for: (i) vertically differentiated products; (ii) an intermediary. We apply the latter model to study loan search using UK mortgage data.

JEL CLASSIFICATION NUMBERS: C14, C57, D83

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

1 Introduction

Consumer search cost is one of the classic explanations for why homogeneous goods or services have different prices. Price dispersion can arise in equilibrium for a search model with minimal heterogeneity. An influential paper by Burdett and Judd (1983) showed that a continuous pricing rule can be generated by a mixed strategy Nash equilibrium in a fixed sample¹ search model with *complete information* consisting of infinitely many identical firms and consumers. There, firms are identical because they have the same marginal cost of production and consumers draw search costs from the same distribution. We refer the reader to a survey by Baye, Morgan, and Scholten (2006) for different rationalizations of price dispersions in search and other models.

Hong and Shum (2006) developed an empirical model of consumer search based on Burdett and Judd (1983). They showed, using just price data, the firms' marginal costs and parts of the distribution of consumer search costs can be identified nonparametrically. Having infinitely many firms is not essential for identification; Moraga-González and Wildenbeest (2008) employed the same strategy to identify an analogous model with a finite number of firms. Identification of the search cost distribution in these papers is, however, only partial. Moraga-González, Sándor and Wildenbeest (2013) showed the degree of partial identification can be reduced, or eliminated asymptotically, if additional price data from other equilibria are available.

¹In 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)).

In this paper we propose an empirical model of a 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 an *incomplete information* environment. We analyze both the theoretical and empirical aspects of this model. We make the following contributions:

(i) Provide a system of equations that characterize non-degenerate pure strategy Bayesian-Nash equilibria (BNE) in the model via a fixed-point. This is useful for solving the model, for the purpose of conducting counterfactual studies.

(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. Our identification strategy leads to closed-form expressions in terms of the observables that suggest easy to compute estimators.

(iii) Construct a nonparametric estimator for the density of the firm's marginal cost that achieves optimal convergence rate uniformly on any fixed subset in the interior of the support. A slight modification of this estimator can achieve any convergence rate that is slower than the optimal rate uniformly over a suitably expanding subset that increases towards the whole support asymptotically. The difference in uniform rates on fixed and expanding support is due to an interesting feature of the equilibrium price density that can be unbounded at the upper support. Our contribution here is a technical one, as we show how a two-step density estimator can obtain a near optimal convergence rate when it relies on a preliminary density estimator that estimates an unbounded density.

We provide three motivations for considering a model with incomplete information by way of incorporating heterogeneity through firms' marginal costs. First, unlike the competitive market framework in Burdett and Judd (1983), many consumer search applications in the literature involve only a small number of oligopolistic firms. It is desirable to allow for firms to have different production technologies that are private information in such setting. Second, heterogeneity amongst firms provides an agnostic way to account for a degree of product differentiation of goods and services that is unobservable (or deemed inconsequential at the time of purchase) to consumers but affects how firms set the price.² This is a useful concept that expand the scope of applications for search models. Third, relative to an incomplete information model, a complete information model has limited capacity in studying counterfactuals that involve changes to firms' production costs beyond a location shift. For example, consider a counterfactual scenario that is represented by a mean-preserving transformation to the firms' cost distribution, which could be motivated by increased uncertainties in supply chains or the adoptions of new technologies or regulations. No complete information model can be used to study the effect of such a counterfactual because the cost distribution is degenerate.

 $^{^{2}}$ For example, online sellers of second-hand books or music records have private information about the actual condition of the item that goes beyond the description provided in the offer because they physically own the product.

While we generalize the model in Moraga-González and Wildenbeest (2008), by allowing heterogeneous firms, due to the difference in assumptions on information structures, our model is closer to an earlier work by MacMinn (1980). MacMinn assumed firms differ by drawing private costs from a uniform distribution. He derived a partial equilibrium result in terms of the optimal pricing rule for firms when all consumers are pre-assigned to search a fixed number of times (cf. Pereira (2005)). Therefore, our analysis is a generalization MacMinn's setup because we do not impose any parametric assumption on the marginal cost distribution and we endogenize consumer search.

Hong and Shum (2006) relied on the constant profit condition that the mixed pricing strategy satisfies for identification in a complete information model. This approach is not applicable in an incomplete information framework. The decision problem for each firm in our model resembles that of a first-price procurement auction where each bidder has to form an expectation on the number and identity of her competitors.³ Moreover, Salz (2020, Appendix B.1) provided numerical evidence suggesting a search model like ours cannot be identified with price data alone. To make progress, we exploit the restriction imposed by market shares and price to identify consumers' search distribution. 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 demand for differentiated products (e.g., see Berry and Haile (2014)). We show that market shares relate to the equilibrium proportions of consumer search linearly in expectation conditional on price. These proportions can be recovered by solving a linear equation. Then, following the insight of Hong and Shum (2006), the proportions can be used to identify (a finite points of) the search cost distribution; they are also an important ingredient for identifying firm's marginal cost distribution. Our approach on the latter 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 do this by deriving the inverse of the equilibrium pricing function and using it to recover firms' latent costs from observed prices. Throughout, our identification strategy is constructive. We show the model parameters can be written explicitly in terms of the joint distribution of observed variables. They can then be estimated nonparametrically, in closed-form, without any numerical optimization.

The non-degenerate equilibrium price distribution has an interesting feature. Our analysis reveals its probability density function (pdf) generally has a pole at the upper support. I.e., the pdf asymp-

³The 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. Notable applications of this relation can be found the labor literature. For examples, Postel-Vinay and Robin (2002) and Cahuc, Postel-Vinay and Robin (2006) 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.

totes to infinity at that point. Intuitively, this happens because there are consumers who search once and will pay whatever price the firm charges up to their valuation of the good. Correspondingly, firms have an incentive to charge close to that price. Poles have also appeared in other structural models such as the equilibrium distributions of bids in a first price auction with binding reserved price (see Section 4 in GPV) and wages in a job search model (Bontemps, Robin and van den Berg (2000)). These authors show their respective estimators of bidder's valuation and firm's productivity densities attain the optimal uniform convergence rates derived in GPV on any fixed interval in the interior of the support. Our density estimator of the firm's costs can achieve the same rates. The pole, however, prevents the optimal convergence rates of these estimators to hold over intervals that expand towards it. In this case, we apply the estimation strategy proposed in Srisuma (2023) to construct an estimator that can achieve any uniform convergence rate that is slower than the optimal benchmark (without a pole), where uniformity is taken over a suitably expanding support that increases to the full support asymptotically.

We provide two extensions of our baseline model where our closed-form identification and convergence rates results mentioned above also apply. One is when there is vertical product differentiation known to consumers and firms, but it is not known to the econometrician. This model can be seen as an incomplete information counterpart to the search model in Wildenbeest (2011) that uses quality to explain systematic price differences between firms. The other is a model with an intermediary (i.e., a broker), where consumers with very high search costs purchase through an intermediary who conducts an exhaustive search for a fee. Our second extension is inspired by the empirical model used in Salz (2020).

There are noteworthy similarities and differences between our model with an intermediary with Salz's. Specifically, without an intermediary, his search model is the same as ours. Indeed, the numerical study by Salz that we mentioned earlier, which suggested one cannot identify our model with price data alone, was performed in a pure search setting. With an intermediary, Salz showed search cost distribution can be identified if the firms' cost distribution is known a priori. Such assumption is plausible in his application where the search and broker markets are separated. Specifically, his broker market is cleared through procurement auctions that he had data for, so he could identify the cost distribution directly using GPV's strategy. In contrast, we combine the search and broker market in a single framework because there are no separate physical markets in our application.

In spite of these connections between the models, Salz's identification approach and ours fundamentally differ. The difference is due to the assumptions on data availability. With the firms' cost distribution known from the broker market, Salz uses consumers' paid prices in the search and broker markets that are individualized (i.e., the same seller can offer different consumers different prices) to match with theoretical moments to estimate model primitives. Our approach, which works with or without intermediaries, uses market-level data on prices and shares as described previously. The notion of using market price in practice is innocuous in applications with posted price where firms charge all consumers the same price. Our approach is also amenable for applications with individualized prices. This, however, requires an additional parametric, though interpretable, assumption on the market price that is constructed from individual-level prices.

For the ease of notation and clarity of idea, the paper presents the identification arguments and theoretical results in the setting where there are no observable characteristics and each firm charges consumers the same price. The identification results immediately extend conditional on observables. It is worth highlighting that variables which affect firms' costs but not search costs provide exclusion restrictions that reduce the degree of partial identification by a similar argument of Moraga-González, Sándor and Wildenbeest (2013) as mentioned earlier. We will illustrate these points in our application, along with how to incorporate parametric assumptions that is coherent with our nonparametric identification strategy. Particularly, beyond employing a parametric assumption to handle individualized prices, imposing some parametric structures is a practical way to deal with the curse of dimensionality when there are many conditioning variables.

Our numerical studies consist of a Monte Carlo simulation and an empirical application. The simulation study confirms the pole exists and shows nonparametric estimators that ignore the pole can perform poorly in its vicinity. Our application uses mortgage data from the UK to estimate a model with mortgage brokers playing the role of intermediaries.

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 studies two extensions of our baseline model. Section 6 and 7 contain a Monte Carlo exercise and empirical application respectively. Section 8 concludes with some discussions. The proofs of all results not given in the main text can be found in the Appendix.

2 Model

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 $G(\cdot)$ and $H(\cdot)$ that 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, Sándor and Wildenbeest (2013), and Sanches, Silva and Srisuma (2018). These only differ from Hong and Shum (2006) in that the latter assumed I is infinite. We describe the decision problem and the best response for the consumers in Section 2.1. The aforementioned papers assume firms have identical production costs 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 $[\underline{C}, \overline{C}] \subset \mathbb{R}^+$ with cdf $G(\cdot)$. A consumer with search cost c faces the following decision problem when a purchase is always made:

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

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:

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

 $\Delta_k(F)$ is non-increasing in k because $\mathbb{E}_F[P_{(1:k)}]$ is non-increasing in k. When price has a differentiable cdf, $\mathbb{E}_F[P_{(1:k)}]$ is strictly decreasing in k and

$$\Delta_k(F) = \int F(p) \left(1 - F(p)\right)^k dp.$$
(2)

It is optimal for a consumer to search once if she draws $c > \Delta_1(F)$ and to search k > 1 times 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} ,

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

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

Note that equation (3) holds irrespective whether the first price is free or not. The key feature of the best response in Lemma 1 is we can sort consumers so that those drawing higher costs cannot search more than those with lower costs. Such structure is also accommodated by non-linear cost functions that allow some economy or dis-economy of scale if one has a prior knowledge to impose them.⁴

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 $[\underline{R}, \overline{R}] \subset \mathbb{R}^+$ 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:

$$\max_{p} \Lambda\left(p, R_{i}; \mathbf{q}\right), \text{ where}$$
$$\Lambda\left(p, R_{i}; \mathbf{q}\right) = \left(p - R_{i}\right) \sum_{k=1}^{I} q_{k} \frac{k}{I} \mathbb{P}\left[P_{(1:k-1)} > p\right]$$

Here $\mathbf{q} = (q_1, \ldots, q_I)^{\top}$ 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. Note that when $q_I = 1$, the firm's decision problem is the same as the bidder's problem in a standard first-price procurement auction.

$$\sigma_{D}(F) = \begin{cases} 1 - G\left(\chi_{1}^{-1}\left(\Delta_{k}(F)/\chi_{2}(k)\right)\right) & \text{for } k = 1\\ G\left(\chi_{1}^{-1}\left(\Delta_{k-1}(F)/\chi_{2}(k-1)\right)\right) - G\left(\chi_{1}^{-1}\left(\Delta_{k}(F)/\chi_{2}(k)\right)\right) & \text{for } k > 1 \end{cases}$$

and the proof strategy for Proposition 1 remains applicable. To see how $\Delta \chi$ accommodates both economy and diseconomy of scale on search: (i) suppose χ_2 is increasing (dis-economy of scale case) then $\{\Delta_k/\chi_2(k)\}_{k=1}^{I-1}$ is strictly decreasing for any Δ_k that comes from a non-degenerate price distribution; (ii) if χ_2 is decreasing (economy of scale case) then $\{\Delta_k/\chi_2(k)\}_{k=1}^{I-1}$ can be strictly decreasing for some price distribution that needs to be determined on a case-by-case basis.

⁴For example, suppose $\Delta \chi(c,k) = \chi_1(c) \chi_2(k)$ where χ_1 is strictly increasing and χ_2 is positive. This includes the linear cost as a special case when $\chi_1(c) = c$ and $\chi_2(k) = 1$. If $\{\Delta_k/\chi_2(k)\}_{k=1}^{I-1}$ is strictly decreasing, then equation (3) in Lemma 1 can be generalized to:

Firm's Best Response

We consider a pricing strategy $\beta : [\underline{R}, \overline{R}] \to [\underline{P}, \overline{P}] \subset \mathbb{R}$ that is strictly increasing almost everywhere and satisfies $\beta(\overline{R}) = \overline{R}$. The latter is the zero profit 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 $\Lambda^*(\cdot; \mathbf{q})$ 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 its inverse, $\beta^{-1}(\cdot; \mathbf{q})$ by $\xi(\cdot; \mathbf{q})$.

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

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

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

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

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

It can be verified that $\beta(\cdot; \mathbf{q})$ is continuous and non-decreasing on $[\underline{R}, \overline{R}]$ as well as satisfying $\beta(\overline{R}) = \overline{R}$. Furthermore, suppose $H(\cdot)$ is differentiable with a positive pdf, $h(\cdot)$. Differentiating the expression above gives,

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

This shows $\beta(\cdot; \mathbf{q})$ is strictly increasing on $[\underline{R}, \overline{R}]$ whenever $q_1 < 1$, and when $q_1 = 1$, $\beta(r; \mathbf{q}) = \overline{R}$ for all r.

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

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} , $\sigma_S(\mathbf{q})$ is the cdf of $\beta(R_i; \mathbf{q})$ where $\beta(\cdot; \mathbf{q})$ 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 $\sigma_D(\cdot)$ are defined in Lemmas 1 and 2 respectively.

An equilibrium with degenerate price distribution always exists in our model. This occurs when all consumers search once and all firms set the monopoly price, i.e. $\beta_M(r; \mathbf{q}_M) = \overline{R}$ for all r. I.e., \mathbf{q}_M is such that $q_{1M} = 1$ and $q_{kM} = 0$ for $k \neq 1$ (cf. Diamond (1971)). Such equilibrium is not suitable in many applications where prices differ. We focus on the case when $\beta(\cdot; \mathbf{q})$ is strictly increasing. Theorem 1 characterizes such equilibria by \mathbf{q} that satisfies (3) and (4) simultaneously.

THEOREM 1. In a symmetric equilibrium $(\mathbf{q}, F) \in \mathbb{S}^{I-1} \times \mathcal{F}$, where the equilibrium pricing strategy is strictly increasing, \mathbf{q} satisfies the following system of equations:

$$q_{k} = \begin{cases} 1 - G\left(\int F(p)\left(1 - F(p)\right)dp\right) & \text{for } k = 1\\ G\left(\int F(p)\left(1 - F(p)\right)^{k-1}dp\right) - G\left(\int F(p)\left(1 - F(p)\right)^{k}dp\right) & \text{for } 1 < k < I \end{cases},$$

$$(6)$$

$$(m) = H\left(f(p; \mathbf{q})\right) \text{ for } all \ p \in [P, \overline{P}]$$

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

The characterization above shows that an equilibrium can be summarized by a fixed-point, which is useful for solving the model. In general there may be multiple equilibria. We are not aware of a 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} is observed market share and P_{im} is price of firm *i* in market *m* such that the data is generated from a single equilibrium. We distinguish Y_{im} from the *theoretical* market share. We formally define both 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 $q_1 \in (0, 1)$ so that $P_{im} = \beta(R_{im}) \equiv \beta(R_{im}; \mathbf{q})$ where $\beta(\cdot; \mathbf{q})$ has been defined in (4) and $\{R_{im}\}_{i=1,m=1}^{I,M}$ is i.i.d. with a continuous density that is positive and finite almost everywhere on $[\underline{R}, \overline{R}]$;

(ii) $\{(Y_{1m},\ldots,Y_{Im})\}_{m=1}^{M}$ is i.i.d. such that the joint distribution of (Y_{im},P_{im}) satisfies,

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

Assumption D(i) assumes observed prices are a random sample from an equilibrium of a search model. $q_1 < 1$ ensures $\beta(\cdot; \mathbf{q})$ is strictly increasing and price has a continuous distribution. Having $q_1 = 0$ does not affect our identification strategy and it simplifies our asymptotic analysis. However, we expect $q_1 > 0$ to be the norm in many applications and it has an econometric implication (due to the pole in the price density (see Lemma 4(b))). We assume the more difficult case is on hand rather than having to treat two separate cases. Equation (7) in Assumption D(ii) states a defining property of market shares as the RHS of the equation has the interpretation of the ex-ante probability of firm *i* winning a sale by setting price to be P_{im} . Assumption D(ii) also assumes shares across markets are i.i.d., but it allows shares to be correlated within a market.

Observed and theoretical market shares generally differ. The former is an aggregation of decisions from a finite number of consumers and the latter aggregates decisions from a continuum of consumers. To define these shares formally, omitting the market index, we denote the search cost for consumer bby c_b and let $C_k := [\Delta_k, \Delta_{k-1}]$ for k > 0 where $\{\Delta_k\}_{k=0}^{I-1}$ denote the search costs where consumers are indifferent in making k and k+1 searches. Let D_{bi} , ℓ_{bi} , and ℓ_{biA} be binary variables that takes value 1 if consumer b respectively purchases from firm i, searches only at firm i, and searches at firm i along with k-1 other firms in the set $\mathcal{A} \in \mathcal{I}_k^i := \{\mathcal{A} = \bigcup_{j \in \mathcal{I} \setminus \{i\}} \{j\} \mid |\mathcal{A}| = k-1\} \subseteq \mathcal{I} := \{1, \ldots, I\}$. The micro-foundation for market share of firm i is based on an individual's purchasing decision,

$$D_{bi} = \ell_{bi} \mathbf{1} \left[c_b > \Delta_1 \right] + \sum_{k=2}^{I} \sum_{\mathcal{A} \in \mathcal{I}_k^i} \ell_{bi\mathcal{A}} \mathbf{1} \left[c_b \in C_k, P_i < \min_{j \in \mathcal{A}} \left\{ P_j \right\} \right].$$
(8)

When each firm has an equal chance of being found, it is easy to verify that $\mathbb{E}[D_{bi}|P_i]$ leads to the expression on the RHS of (7). When Y_{im} is defined as an average of D_{bi} over B i.i.d. consumers, (7) follows. Moreover, as $B \to \infty$, by the law of large numbers, Y_{im} converges to the theoretical market share:

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

where $C_k^I := \frac{I!}{(I-k)!k!}$. Note that $\mathbb{E}\left[\overline{Y}_{im}|P_{im}\right] = \mathbb{E}\left[Y_{im}|P_{im}\right]$, because $\sum_{\mathcal{A}\in\mathcal{I}_{ik}}\mathbb{E}\left[P_{im} < \min_{j\in\mathcal{A}}\left\{P_{jm}\right\}|P_{im}\right] = C_{k-1}^{I-1}\left(1 - F\left(P_{im}\right)\right)^{k-1}$ and $C_{k-1}^{I-1}/C_k^I = \frac{k}{I}$. The discrepancy between Y_{im} and \overline{Y}_{im} thus represents an approximation error of the model. The error, $\epsilon_{im} := Y_{im} - \overline{Y}_{im}$, exhibits a property of a classical measurement error since $\mathbb{E}\left[\epsilon_{im}|P_{im}\right] = 0$. Note that $\left(\overline{Y}_{im}, \overline{Y}_{jm}\right)$ are correlated as they are jointly determined by prices and we expect (Y_{im}, Y_{jm}) to be correlated as a consequence.

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

$$Y_{im} = X_{im}^{\top} \mathbf{q} + \varepsilon_{im}, \tag{9}$$

where ε_{im} satisfies $\mathbb{E}[\varepsilon_{im}|P_{im}] = 0$. We can then identify **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 **q** is identified. We now treat both **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 **q** are identified. In particular, note that Δ_k is a functional of $F(\cdot)$ for all k (see (1) and (2)) and **q** is identified from Lemma 3.

Our identification strategy for **q** is different to the method Hong and Shum (2006) used to identify a complete information model. Once **q** is identified, however, we identify each $G(\Delta_k)$ in the same way as them. This approach only partially identifies $G(\cdot)$, as we can identify it at $\{\Delta_k\}_{k=1}^{I-1}$. The degree of non-identification can be reduced if there is 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. Moraga-González, Sándor and Wildenbeest (2013) propose this identification strategy in the complete information context. The same idea applies to our setting. We illustrate in Section 7 how to find and exploit such variation empirically.

3.2 Firms

We can identify $H(\cdot)$ by inverting latent production costs from observed prices. Lemma 4 provides key properties of the price density and gives an explicit formula for the inverse of the pricing strategy in terms of **q** and the price distribution.

LEMMA 4. Suppose Assumption D(i) holds. Then:

(a) $\inf_{p \in [\underline{P}, \overline{P}]} f(p) > 0;$ (b) $\lim_{p \to \overline{P}} f(p) = \infty;$ (c) the inverse of the equilibrium pricing strategy, $\xi : [\underline{P}, \overline{P}] \to [\underline{R}, \overline{R}],$ takes the following form

$$\xi(p) = p - \frac{\sum_{k=1}^{I} q_k k \left(1 - F(p)\right)^{k-1}}{f(p) \sum_{k=2}^{I} q_k k \left(k - 1\right) \left(1 - F(p)\right)^{k-2}},$$
(10)

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

When $q_1 = 0$, it can be shown that $f(\cdot)$ is bounded away from zero on $[\underline{P}, \overline{P})$ and $f(\overline{P})$ does not need to be infinite for $\xi(\overline{P}) = \overline{R}$.

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 look to estimate $h(\cdot)$ at the best possible convergence rate. We focus on this problem because $h(\cdot)$ is the most difficult object to estimate in our model in the sense that its estimator has the slowest convergence rate amongst estimators of other identifiable objects. Along the way, we will discuss estimators of other parameters in the model. Particularly, 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 $[\underline{R}, \overline{R}]$. 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 $[\underline{R}, \overline{R}]$ as the sample size increases. In this case we will provide another estimator for $h(\cdot)$ that can converge at any slower rate than the one achievable over a fixed support. The reason for a slower convergence rate is the latter accounts for the pole $(\lim_{p\to\overline{P}} f(p) = \infty)$. 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 (10)). 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 up to $\tau + 1$ continuous derivatives on $|\underline{R}, \overline{R}|$ for some $\tau \geq 1$.

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

Lemma 5 says that $f(\cdot)$ has the same degree of smoothness as $H(\cdot)$ everywhere other than at \overline{P} . We next define estimators for $(\mathbf{q}, f(\cdot), F(\cdot))$ and discuss their convergence rates under Assumptions D and R.

An estimator for $F(\cdot)$

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

$$\widehat{F}(p) = \frac{1}{MI} \sum_{m=1}^{M} \sum_{i=1}^{I} \mathbf{1} \left[P_{im} \le p \right] \quad \text{for all } p.$$
(11)

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 $[\underline{P}, \overline{P}]$. Then by the continuous mapping theorem, $\sup_{p \in [\underline{P}, \overline{P}]} \left|\widehat{F}(p) - F(p)\right| = O_p\left(1/\sqrt{M}\right)$.

An estimator for q

We suggest to estimate \mathbf{q} by least squares. Let $\mathbf{Y}_m = (Y_{1m}, \dots, Y_{Im})^\top$, $\mathbf{e}_m = (\varepsilon_{1m}, \dots, \varepsilon_{Im})^\top$ 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 $\mathbf{Y} = [\mathbf{Y}_1^\top : \cdots : \mathbf{Y}_M^\top]^\top$, $\mathbf{X} = [\mathbf{X}_1^\top : \cdots : \mathbf{X}_M^\top]^\top$ and $\mathbf{e} = [\mathbf{e}_1^\top : \cdots : \mathbf{e}_M^\top]^\top$ respectively. Then a vector version of (9) 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,

$$\widehat{\mathbf{q}} = \left(\widehat{\mathbf{X}}^{\top} \widehat{\mathbf{X}} \right)^{-1} \widehat{\mathbf{X}}^{\top} \mathbf{Y},$$

$$= \mathbf{q} + \mathbf{a}_{M} + \mathbf{b}_{M}, \text{ where}$$

$$\mathbf{a}_{M} = \left(\mathbf{X}^{\top} \mathbf{X} \right)^{-1} \mathbf{X}^{\top} \mathbf{e},$$

$$\mathbf{b}_{M} = \left(\left(\widehat{\mathbf{X}}^{\top} \widehat{\mathbf{X}} \right)^{-1} \widehat{\mathbf{X}}^{\top} - \left(\mathbf{X}^{\top} \mathbf{X} \right)^{-1} \mathbf{X}^{\top} \right) \mathbf{Y}.$$

$$(12)$$

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 $\|\mathbf{b}_M\| = O_p\left(1/\sqrt{M}\right)$. 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 $\hat{\mathbf{q}}$ and $\hat{F}(\cdot)$ based on the constructive identification result in Proposition 1. Such estimator will be a smooth functional of $\hat{F}(\cdot)$ 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) proposed 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 et al. (2018) for further details.

An estimator for $f(\cdot)$

Consider the following kernel density estimator for $f(\cdot)$,

p

$$\widehat{f}(p) = \frac{1}{MIb_{f,M}} \sum_{m=1}^{M} \sum_{i=1}^{I} K\left(\frac{P_{im} - p}{b_{f,M}}\right) \quad \text{for all } p,$$
(13)

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

$$\sup_{\in [\underline{P}+\delta,\overline{P}-\delta]} \left| \widehat{f}(p) - f(p) \right| = O(\eta_M^*) \text{ a.s.},$$
(14)

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 (11) to (13):

(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

$$\widehat{R}_{im} = \begin{cases}
P_{im} - \frac{\sum_{k=1}^{I} \widehat{q}_k k \left(1 - \widehat{F}(P_{im})\right)^{k-1}}{\widehat{f}(P_{im}) \sum_{k=1}^{I} \widehat{q}_k k (k-1) \left(1 - \widehat{F}(P_{im})\right)^{k-2}} & \text{for } P_{im} \in [\underline{P} + \delta, \overline{P} - \delta] \\
+\infty & \text{otherwise}
\end{cases} .$$
(15)

When $\widehat{R}_{im} < \infty$, \widehat{R}_{im} is the estimator of R_{im} based on on the feasible version of (10). 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 the convergence rate of \widehat{R}_{im} is determined by $\sup_{p \in [\underline{P}+\delta, \overline{P}-\delta]} |\widehat{f}(p) - f(p)|$. We effectively omit \widehat{R}_{im} when $P_{im} \notin [\underline{P}+\delta, \overline{P}-\delta]$ for the purpose of estimating $h(\cdot)$. The omission does not prevent us from attaining the desired convergence rates because the probability that $\widehat{R}_{im} = +\infty$ is 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:

$$\widehat{h}(r) = \frac{1}{MIb_{h,M}} \sum_{m=1}^{M} \sum_{i=1}^{I} K\left(\frac{\widehat{R}_{im} - r}{b_{h,M}}\right) \quad \text{for any } r,$$
(16)

where $K(\cdot)$ is a kernel function $b_{h,M}$ is the bandwidth. Under the conditions of Theorem 2, the uniform convergence rate of $\hat{h}(\cdot)$ is determined by the convergence rate of \widetilde{R}_{im} .

THEOREM 2. Suppose Assumptions D and R hold. Assume the following properties for components in (16):

- (i) $K(\cdot)$ be a symmetric τ -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

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

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 τ -times continuously differentiable density function.

Next, we extend the study of uniform convergence rate for an estimator of $h(\cdot)$ over $[\underline{R} + \varsigma_M, \overline{R} - \varsigma_M]$ for some $\varsigma_M = o(1)$. This requires us to provide a rate for an estimator of $f(\cdot)$ over intervals that expand towards $[\underline{P}, \overline{P}]$, which turns out to be an unusual problem⁵ because $f(\cdot)$ has a pole at \overline{P}

 $^{{}^{5}}$ One of the main assumptions commonly used in deriving uniform convergence rates of a kernel density estimator is boundedness of the underlying density. E.g., see Andrews (1995), Masry (1996), Fan and Yao (2003) and Hansen (2008).

(Lemma 4(b)). Particularly, we should expect a slower convergence rate in the vicinity of a pole since the asymptotic variance and bias of kernel density estimator are pointwise proportional to the underlying density and its derivatives respectively.

To this end, we apply the log-transformation approach suggested in Srisuma (2023) to estimate $f(\cdot)$. His proposal is based on the observation that a log-transformed random variable will have a bounded density under a mild regularity condition. In the context of our application, without having to specify the divergence rate of $f(\cdot)$, the log-transformed price will have a bounded density as long as the limit of $(\overline{P} - p) f(p)$ as $p \to \overline{P}$ exists^{6,7}. Subsequently, we can estimate $f(\cdot)$ at any convergence rate that is slower than η_M^* uniformly over a suitably expanding support (cf. (14)) using the back-transformed estimator, which in turn allows us to estimate $h(\cdot)$ at any rate slower than $(\frac{\log M}{M})^{\frac{\tau}{2\tau+3}}$ over an expanding support in a similar way as we have shown in Theorem 2.

Let $P_{im}^{\dagger} \equiv -\ln\left(\overline{P} - P_{im}\right)$. Denote the pdf of P_{im}^{\dagger} by $f^{\dagger}(\cdot)$ that is positive on $\left[-\ln\left(\overline{P} - \underline{P}\right), \infty\right)$. By a change of variables, we have $f(p) = \frac{f^{\dagger}(-\ln(\overline{P}-p))}{\overline{P}-p}$ for $p \in [\underline{P}, \overline{P}]$. Lemma 7 says that $\lim_{p\to\overline{P}} (\overline{P} - p) f(p)$ exists, which ensures $f^{\dagger}(\cdot)$ and its derivatives are bounded by Lemma 8. We estimate $f^{\dagger}(\cdot)$ using a standard kernel density estimator with $\left\{P_{im}^{\dagger}\right\}_{i=1,m=1}^{I,M}$ that will achieve the η_M^* -convergence rate. We then multiply this estimator by $(\overline{P} - p)^{-1}$ to get the back-transformed estimator. Since $p \mapsto (\overline{P} - p)^{-1}$ increases to infinity as $p \to \overline{P}$, the back-transformation slows down the uniform convergence rate as the support expands.

LEMMA 7. Suppose Assumption D(i) holds. Then $\lim_{p\to\overline{P}} (\overline{P}-p) f(p)$ exists.

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

(a) $f^{\dagger}(\cdot)$ is bounded;

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

⁶When $\lim_{p\to\overline{P}} (\overline{P} - p) f(\overline{p})$ exists, it must be 0. We show the limit is indeed 0 in the proof of Lemma 7. This corresponds to the rate that $(\overline{P} - p)^{-1}$ diverges as $p \to \overline{P}$ is an upper bound on the divergence rate of $f(\cdot)$; cf. Proposition 2(a) in Srisuma (2023).

⁷We refer the reader to Srisuma (2023) for further discussions on why the rate that $(\overline{P} - p)^{-1}$ diverges as $p \to \overline{P}$ is a suitable choice for an upper bound on the divergence rate of $f(\cdot)$ at \overline{P} ; examples of primitive conditions that imply the existence of $\lim_{p\to\overline{P}} (\overline{P} - p) f(p)$; and, other ways to construct estimators that can obtain the same uniform convergence rate as the one based on a log-transformation.

Formally, the back-transformed estimator is defined as follows:

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

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

$$(17)$$

By using a $(\tau + 1)$ -th higher order kernel and set $b_{f^{\dagger},M}$ to be proportional to $\left(\frac{\log M}{M}\right)^{\frac{1}{2\tau+3}}$, we have $\left|\widehat{f^{\dagger}}\left(p^{\dagger}\right) - f^{\dagger}\left(p^{\dagger}\right)\right| = O\left(\eta_{M}^{*}\right)$ a.s. uniformly over any fixed inner proper subset of $\left[-\ln\left(\overline{P} - \underline{P}\right), \infty\right)$. Since $\widetilde{f}\left(p\right) - f\left(p\right) = \frac{\widehat{f^{\dagger}}\left(-\ln\left(\overline{P} - \underline{p}\right)\right) - f^{\dagger}\left(-\ln\left(\overline{P} - \underline{p}\right)\right)}{\overline{P} - p}$, we have $\left|\widetilde{f}\left(p\right) - f\left(p\right)\right| = O\left(\eta_{M}^{*}\right)$ a.s. uniformly over any fixed inner subset of $\left[\underline{P}, \overline{P}\right]$.

There are two factors that affect the uniform convergence rate of $\tilde{f}(\cdot)$ over $[\underline{P} + \delta'_M, \overline{P} - \delta''_M]$ when both δ'_M and δ''_M are o(1). One is the boundary bias from estimating $f^{\dagger}(p)$ when p^{\dagger} lies within a $b_{f^{\dagger},M}$ -neighborhood from $-\ln(\overline{P} - \underline{P})$. The other is the divergence of $(\overline{P} - p)^{-1}$ diverges as $p \to \overline{P}$. We can avoid the boundary effect at the lower boundary by limiting the rate δ'_M goes to zero according to $\overline{P} - \exp\left(\ln(\overline{P} - \underline{P}) - b_{f^{\dagger},M}\right) = o(\delta'_M)$. We can choose the divergence rate from the back-transformation. Suppose $p \to \overline{P}$ at the rate δ''_M , then

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

Thus, we can always find an interval expanding to $[\underline{P}, \overline{P}]$ that $\tilde{f}(\cdot) - f(\cdot)$ converges uniformly over at any rate that is slower than η_M^* . We state this as a proposition.

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

We can then estimate R_{im} on using $\tilde{f}(\cdot)$, and use it to estimate $h(\cdot)$ as done previously in (15) and (16) respectively. Specifically, for any $\delta_M > 0$ let

$$\widetilde{R}_{im} = \begin{cases}
P_{im} - \frac{\sum_{k=1}^{I} \widehat{q}_k k \left(1 - \widehat{F}(P_{im})\right)^{k-1}}{\widetilde{f}(P_{im}) \sum_{k=1}^{I} \widehat{q}_k k (k-1) \left(1 - \widehat{F}(P_{im})\right)^{k-2}} & \text{for } P_{im} \in [\underline{P} + \delta_M, \overline{P} - \delta_M] \\
+\infty & \text{otherwise}
\end{cases}, \quad (18)$$

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

$$(19)$$

The following results are similar to Lemma 6 and Theorem 2. They differ in that the rates do not reach the optimal rate and they hold over a sequence of expanding intervals.

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

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

THEOREM 3. Suppose Assumptions D and R hold. Assume the following properties for components in (19):

(i) $K(\cdot)$ be a symmetric τ -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 $\eta_M = O(b_{h,M}^2)$, 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 $\tilde{h}(\cdot)$ is derived over an expanding support that avoids the boundary effect as well as anticipating the pole effect. We highlight the condition Theorem 3 imposes, which is not required in Theorem 2, is $\eta_M = O(b_{h,M}^2)$. This is a mild condition to handle the pole. In particular, this condition is not restrictive when $\tau \geq 2$. To see this, suppose $\eta_M = \eta_M^* \phi_M$ for some ϕ_M with $\lim_{M\to\infty} \phi_M = \infty$, then $\eta_M = O(b_{h,M}^2)$ is equivalent to $\phi_M (\frac{\log M}{M})^{\frac{\tau-1}{2\tau+3}} = O(1)$. Since we are only be interested in ϕ_M that diverges to infinity slowly, we can choose it to diverge at an arbitrarily slow rate.

5 Extensions

We consider two extensions of the nonparametric identification arguments presented in Section 3. The first allows products to be vertically differentiated that can be useful for modeling systematic price differences across firms. The second introduces an intermediary that can search on behalf of consumers at a fee. Our discussions in this section will focus on identification. We will show the estimation strategy and the results on convergence rates of developed in Section 4 are applicable to these settings.

5.1 Vertical Product Differentiation

Let firm *i*'s product is characterized by $\nu_i \in \mathbb{R}$, which is a measure of differentiated quality. The econometrician will observe market share and prices but not quality of the products. The main modeling assumption employed by Wildenbeest (2011), in a complete information model, 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,

$$U_i := \nu_0 + \nu_i - P_i, \tag{20}$$

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:

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

A purchase is always made so that ν_0 does not enter our analysis, just as it does not in the model with a homogeneous product. For the moment suppose firms set prices such that $\{U_i\}_{i=1}^{I}$ is a random sample. Then, for $k \ge 1$, 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 $L(\cdot)$.

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,

$$\Upsilon_k(L) := \mathbb{E}_L\left[U_{(k:k)}\right] - \mathbb{E}_L\left[U_{(k-1:k-1)}\right].$$
(21)

We set $U_{(0:0)}$ to be 0. The consumer's best response is to search once if $c > \Upsilon_1(L)$, and search k > 1times if $\Upsilon_{k-1}(L) < c \leq \Upsilon_k(L)$. 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 components:

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

where R_{0i} has cdf $H_0(\cdot)$ supported on $\mathcal{R}_0 \in [\underline{R}_0, \overline{R}_0]$ for some $\overline{R}_0 > \underline{R}_0 > 0$. We denote the support of R_i by $\mathcal{R}_i := [\nu_i + \underline{R}_0, \nu_i + \overline{R}_0]$. 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.

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 $\mathcal{P}_i = [\nu_i + \underline{P}_0, \nu_i + \overline{P}_0]$ and $\beta_i(\cdot; \mathbf{q}) = \nu_i + \beta_0(\cdot; \mathbf{q})$ and $\beta_0(\cdot; \mathbf{q}) : \mathcal{R}_0 \to \mathcal{P}_0 = [\underline{P}_0, \overline{P}_0]$. 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 $x_i \in [-\overline{R}_0, -\underline{R}_0]$, 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 $\mu(-\overline{R}_0; \mathbf{q}) = \overline{R}_0$. Under this premise, we can apply the arguments used to obtain (4) to show that

$$\mu\left(x\left(r_{0i}\right);\mathbf{q}\right) = x\left(r_{0i}\right) - \frac{\sum_{k=1}^{I} q_k k \int_{s=r_{0i}}^{\overline{R}_0} \left(1 - H_0\left(s\right)\right)^{k-1} ds}{\sum_{k=1}^{I} q_k k \left(1 - H_0\left(r_{0i}\right)\right)^{k-1}},$$
(22)

for any $r_{0i} \in \mathcal{R}_0$ and $x(r_{0i}) := -r_{0i}$. 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,

$$\beta_0(r_{0i}; \mathbf{q}) = r_{0i} + \frac{\sum_{k=1}^{I} q_k k \int_{s=r_{0i}}^{\overline{R}_0} (1 - H_0(s))^{k-1} ds}{\sum_{k=1}^{I} q_k k (1 - H_0(r_{0i}))^{k-1}}.$$
(23)

 $\beta_0(\cdot; \mathbf{q})$ 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 (10) in Lemma 5.

Equilibrium

We define a quasi-symmetric equilibrium where players using pricing strategies that are affine translation from each other. Theorem 4 gives a characterization of the equilibrium (cf. Theorem 1). Its proof and the definition of a quasi-symmetric equilibrium are in the Appendix.

THEOREM 4. In a quasi-symmetric equilibrium $(\mathbf{q}, F_0) \in \mathbb{S}^{I-1} \times \mathcal{F}_0$, where the equilibrium pricing strategies are strictly increasing, \mathbf{q} satisfies the following system of equations:

$$q_{k} = \begin{cases} 1 - G\left(\int F_{0}(p)\left(1 - F_{0}(p)\right)dp\right) & \text{for } k = 1\\ G\left(\int F_{0}(p)\left(1 - F_{0}(p)\right)^{k-1}dp\right) - G\left(\int F_{0}(p)\left(1 - F_{0}(p)\right)^{k}dp\right) & \text{for } 1 < k < I \end{cases}$$

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 $\beta_0(\cdot; \mathbf{q})$.

5.1.1 Identification

We assume our data satisfy the following conditions.

Assumption DE1. $\{(Y_{im}, P_{im})\}_{i=1,m=1}^{I,M}$ 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 (23) and $\{R_{0im}\}_{i=1,m=1}^{I,M}$ is i.i.d. with positive and finite density almost everywhere on $[\underline{R}_0, \overline{R}_0]$;

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

$$\mathbb{E}\left[Y_{im}|P_{0im}\right] = \sum_{k=1}^{I} q_k \frac{k}{I} \left(1 - F_0\left(P_{0im}\right)\right)^{k-1}.$$
(24)

Assumption DE1 has an analogous interpretations to Assumption D. 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: (i) use $\{P_{0im}\}_{i=1,m=1}^{I,M}$ to identify $f_0(\cdot)$ and $F_0(\cdot)$; (ii) identify **q** from (24) (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

$$R_{0im} = P_{0im} - \frac{\sum_{k=1}^{I} q_k k \left(1 - F_0\left(P_{0im}\right)\right)^{k-1}}{f_0\left(P_{0im}\right) \sum_{k=2}^{I} q_k k \left(k-1\right) \left(1 - F_0\left(P_{0im}\right)\right)^{k-2}},$$
(25)

cf. (10), which in turn identifies $H_0(\cdot)$ (cf. Proposition 2).

We, however, do not observe $\{\nu_i\}_{i=1}^{I}$. 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 $U_{im} = \nu_i - P_{im}$, so for all *i* and *j*:

$$U_{im} - U_{jm} = P_{0jm} - P_{0im} = \omega_{jm} - \omega_{im},$$
(26)

where ω_{im} denotes $P_{im} - \mathbb{E}[P_{im}]$, thus the second equality above follows from $\mathbb{E}[P_{0im}] = \mathbb{E}[P_{0jm}]$.

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 (26) 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. We use it to identify the consumer search distribution.

LEMMA 10. Suppose Assumption DE1 holds. Then $F_0(P_{0im}) = W(\omega_{im})$ for all i and m.

PROPOSITION 5. Suppose Assumption DE1 holds. Then $G(\Upsilon_k)$ is identified for k = 1, ..., I - 1. PROOF. By Lemma 10, any **q** that satisfies (24) also satisfies

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

We can then identify \mathbf{q} in closed-form as done in Lemma 3. From (21), 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 (25) 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 DE1 holds. Then the distribution of ρ_{im} is identified. PROOF. Replace $(P_{0im}, f_0(P_{0im}), F_0(P_{0im}))$ in the RHS of (25) by $(\omega_{im}, w(\omega_{im}), W(\omega_{im}))$ to construct ρ_{im} . Then apply Lemma 10.

Propositions 5 and 6 show we can use $\{\omega_{im}\}_{i=1,m=1}^{I,M}$ instead of price to identify the demand and supply side parameters in the same way as done in Sections 3.1 and 3.2 respectively. Analogous estimators and results on convergence rates discussed in Section 4 are therefore immediately applicable.

It is worth commenting that not knowing $\{\nu_i\}_{i=1}^{I}$ does not limit the scope of counterfactual studies relative to the model with homogenous goods. This is because consumers in our model bear the cost of quality differences and get compensated in equal amount in terms of 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} .

5.2 Intermediary

Next, we assume there is an intermediary, which we will also refer to as *broker*. Our model is closely related to Salz (2020). We adopt his key assumption that a consumer with very high cost consumers prefer to pay a broker to search. The broker then performs an exhaustive search and gives such consumer the lowest price.

Consumer's Best Responses

Previously we saw that it is optimal for a consumer that draws $c > \Delta_1(F)$ to search once. If a broker charges ϕ to conduct an exhaustive search for such consumer, the cost that makes a consumer indifferent between searching once or delegating search solves: $\mathbb{E}_F[P] + c = \phi + \mathbb{E}_F[P_{(1:I)}]$. Let us denote such cost by $\Delta_0(F) := \phi + \mathbb{E}_F[P_{(1:I)}] - \mathbb{E}_F[P]$. We take ϕ to be exogenous. Following Salz (2020), we focus on an equilibrium where $c > \Delta_0(F)$ has positive probability and $\Delta_0(F) > \Delta_1(F)$. This allows us to generalize Lemma 1 and define a consumer's best response to be a map $\tilde{\sigma}_D : \mathcal{F} \to \mathbb{S}^I$ such that for any F in \mathcal{F} ,

$$\widetilde{\sigma}_{D}(F) = \begin{cases} 1 - G(\Delta_{k}(F)) & \text{for } k = 0\\ G(\Delta_{k-1}(F)) - G(\Delta_{k}(F)) & \text{for } 0 < k < I\\ G(\Delta_{k}(F)) & \text{for } k = I \end{cases}$$

Firm's Best Response

Let us denote the proportion of consumers that use a broker by $q_0 < 1$. Firm *i* chooses price to maximize $\widetilde{\Lambda}(p, R_i; q_0, \mathbf{q})$ where:

$$\widetilde{\Lambda}(p, R_i; q_0, \mathbf{q}) := (1 - q_0) (p - R_i) \sum_{k=1}^{I} q_k \frac{k}{I} \mathbb{P}\left[P_{(1:k-1)} > p\right] + q_0 (p - R_i) \mathbb{P}\left[P_{(1:I-1)} > p\right],$$

where **q** now represents a vector of proportions of consumers that search a different number of firms, conditioning on them searching, so that $(q_0, \mathbf{q}) \in [0, 1) \times \mathbb{S}^{I-1}$. The two additive components on the RHS in the display above are firm *i*'s expected payoffs from the consumers that search and use an intermediary respectively. It will be useful to define the vector $\widetilde{\mathbf{q}}(q_0, \mathbf{q}) := (\widetilde{q}_k(q_0, \mathbf{q}))_{k=1}^I \in \mathbb{S}^{I-1}$ with the following components,

$$\widetilde{q}_k(q_0, \mathbf{q}) = (1 - q_0) q_k \text{ for } 0 < k < I \text{ and } \widetilde{q}_I(q_0, \mathbf{q}) = (1 - q_0) q_I + q_0.$$
(27)

Then, we have for all (p, r) and (q_0, \mathbf{q}) ,

$$\widetilde{\Lambda}(p,r;q_0,\mathbf{q}) = (p-r)\sum_{k=1}^{I} \widetilde{q}_k(q_0,\mathbf{q}) \frac{k}{I} \mathbb{P}\left[P_{(1:k-1)} > p\right] = \Lambda\left(p,r;\widetilde{\mathbf{q}}(q_0,\mathbf{q})\right),$$

where $\Lambda(\cdot)$ is the same function used in Section 2.2. Subsequently, the solution to the maximization problem above is given by $\beta(\cdot; \tilde{\mathbf{q}}(q_0, \mathbf{q}))$ (see (4)) where $\beta(\cdot; \tilde{\mathbf{q}}(q_0, \mathbf{q}))$ is strictly increasing when $\tilde{q}_1(q_0, \mathbf{q}) < 1$. Notably, its inverse takes the same form as (10) given in Lemma 5 and the characteristics of the equilibrium price distribution of Lemma 7 applies when $\tilde{q}_1(q_0, \mathbf{q}) > 0$.

Equilibrium

We characterize an equilibrium for a search model with an intermediary as follows.

THEOREM 5. In a symmetric equilibrium $(q_0, \mathbf{q}, F) \in [0, 1) \times \mathbb{S}^{I-1} \times \mathcal{F}$, where $\Delta_0(F) > \Delta_1(F)$ and the equilibrium pricing strategies are strictly increasing, (q_0, \mathbf{q}) satisfies the following system of equations:

$$q_{0} = 1 - G(\Delta_{k}(F)) \quad \text{for } k = 0$$

$$(1 - q_{0}) q_{k} = G(\Delta_{k-1}(F)) - G(\Delta_{k}(F)) \quad \text{for } 0 < k < I ,$$

$$(1 - q_{0}) q_{I} = G(\Delta_{k}(F)) \quad \text{for } k = I$$

where $F(p) = H(\xi(p; \widetilde{\mathbf{q}}(q_0, \mathbf{q})))$ for all $p \in [\underline{P}, \overline{P}]$ and $\xi(\cdot; \widetilde{\mathbf{q}}(q_0, \mathbf{q}))$ is the inverse of $\beta(\cdot; \widetilde{\mathbf{q}}(q_0, \mathbf{q}))$.

Identification

We can apply the same identification strategy as in Section 3 when q_0 is known. We make the following assumptions.

ASSUMPTION DE2. $\{(Y_{im}, P_{im})\}_{i=1,m=1}^{I,M}$ is a sequence of random variables such that for a known $q_0 \in (0,1)$:

(i) there exists $(\mathbf{q}, F) \in \mathbb{S}^{I-1} \times \mathcal{F}$ with $q_1 \in (0, 1)$ so that $P_{im} = \beta(R_{im}; \widetilde{\mathbf{q}}(q_0, \mathbf{q}))$ where $\beta(\cdot; \mathbf{q})$ and $\widetilde{\mathbf{q}}(q_0, \mathbf{q})$ are defined in (4) and (27) respectively, and $\{R_{im}\}_{i=1,m=1}^{I,M}$ is i.i.d. with positive and finite density almost everywhere on $[\underline{R}, \overline{R}]$;

(ii) $\{(Y_{1m},\ldots,Y_{Im})\}_{m=1}^{M}$ is i.i.d. such that the joint distribution of (Y_{im},P_{im}) satisfies,

$$\mathbb{E}\left[Y_{im}|P_{im}\right] = \sum_{k=1}^{I} \widetilde{q}_k \left(q_0, \mathbf{q}\right) \frac{k}{I} \left(1 - F\left(P_{im}\right)\right)^{k-1}.$$
(28)

The known q_0 assumption is suitable when the proportion of consumers who used a broker can be identified directly from the data. Examples of this include Salz (2020) and our application in Section 7. Other than assuming $q_0 \in (0, 1)$, the discussions on Assumption D are applicable to the remainder of Assumption DE2.

Under Assumption DE2, $\tilde{\mathbf{q}}(q_0, \mathbf{q})$ is identified as long as $\mathbb{E}\left[X_{im}X_{im}^{\top}\right]$ has full rank $(X_{im}$ is defined as in Section 3.1). We can then identify $\{G(\Delta_k)\}_{k=0}^{I-1}$ as in Proposition 1.

PROPOSITION 7. Suppose Assumption DE2 holds and $\mathbb{E}\left[X_{im}X_{im}^{\top}\right]$ has full rank. Then $G(\Delta_k)$ is identified for $k = 0, 1, \ldots, I - 1$.

PROOF. First apply Lemma 3 to identify $\tilde{\mathbf{q}}(q_0, \mathbf{q})$. Since q_0 is known, \mathbf{q} is identified from $\tilde{\mathbf{q}}(q_0, \mathbf{q})$. Subsequently, $G(\Delta_0) = 1 - q_0$ and $G(\Delta_k) = 1 - q_0 - (1 - q_0) \sum_{k'=1}^k q_{k'}$ for k > 0.

Once $\tilde{\mathbf{q}}(q_0, \mathbf{q})$ is known, we can identify $H(\cdot)$. The argument in Section 3.2 applies directly, because the inverse of $\beta(\cdot; \tilde{\mathbf{q}}(q_0, \mathbf{q}))$ takes the same form as (10) and the price distribution is identified.

6 Monte Carlo Study

The purpose of this section is to numerically investigate theoretical features of our model. We consider a simple design with three firms. Consumers draw costs from a distribution with $\operatorname{cdf} G(c) = \sqrt{c}$ for $c \in [0, 1]$. Firms draw costs from a uniform distribution on [0, 1]. We solved for the equilibrium of the game by iterating the system of equations in (6). We tried different initial values and found only one equilibrium that generates price dispersion with $\mathbf{q} = (0.7852, 0.0455, 0.1693)$. 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 $h(\cdot)$. We estimate $F(\cdot)$ and \mathbf{q} using the estimators described in Section 4. 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 can use data outside of the interval that are closer to the boundaries if they can be well estimated. Following the suggestion of Hickman and Hubbard (2015), who estimated a first-price auction model, we employ a boundary corrected kernel and 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). We note that our estimation problem is more challenging than a pure auction setup because we have to estimate $f(\cdot)$, which has a pole, and the estimation of $h(\cdot)$ has additional sampling errors from estimating \mathbf{q} and $F(\cdot)$.



We consider two estimators for $f(\cdot)$. $\widehat{f}_1(\cdot)$ is an estimator based on KZ that accounts for the boundary effects but ignores the presence of the pole. $\widehat{f}_{2}(\cdot)$ uses the transformation described in equation (17) to accommodate the pole and applies boundary correction at the lower boundary. 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 and 2 plot the mean and the 5th and 95th percentiles for each $(\widehat{f}_1(\cdot), \widehat{f}_2(\cdot))$ against the true price pdf. We see that $\widehat{f}_{1}(\cdot)$ performs quite well near the lower support point but not near the pole. $\widehat{f}_{2}(\cdot)$ performs much better near the pole. A careful inspection, however, shows the bias of $\widehat{f}_{2}(\cdot)$ is generally larger than that of $\widehat{f}_1(\cdot)$ away from the pole. We next estimate $h(\cdot)$. The plots in Figures 3 and 4 contain the mean and the 5th and 95th percentiles of KZ boundary corrected estimators, $(\hat{h}_1(\cdot), \hat{h}_2(\cdot))$, that correspond respectively to $(\widehat{f}_1(\cdot), \widehat{f}_2(\cdot))$. These figures also include analogous plots from an infeasible KZ boundary corrected estimator constructed from the estimated costs when the true $f(\cdot)$ is used, while q and $F(\cdot)$ are still estimated, to highlight the effect of density estimation. The infeasible estimator is generally the superior estimator as expected, although it is worth noting that even the infeasible estimator still suffer from the boundary effect. For the feasible estimators, $\hat{h}_2(\cdot)$ has lower bias over its lower half of the support compared to the upper half due to substantial bias from estimating $f(\cdot)$ near the pole. In contrast, $\hat{h}_2(\cdot)$ 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 $\hat{h}_j(\cdot)$ inherits characteristics of $\hat{f}_j(\cdot)$. Therefore it is clear one should account for the boundary effect at the lower support as well as the pole at the upper support. One way to proceed in practice is to perform some kind of model averaging. As-



ymptotically, such estimator will be consistent since $(\hat{f}_1(\cdot), \hat{f}_2(\cdot))$, and subsequently $(\hat{h}_1(\cdot), \hat{h}_2(\cdot))$, are consistent estimators on the interior of the support.

7 Empirical Application

7.1 Data and setup

This section aims to illustrate how adaptations of our baseline model can be applied to a rich dataset through the lens of a search model. The particular adaptations are: (i) incorporating observable firm and consumer heterogeneity as well as market heterogeneity; (ii) introducing an intermediary who provides information about prices in exchange for a fee (see section 5.2); (iii) showing how some parts of the estimation procedure can be conducted parametrically to deal with the curse of dimensionality.

Our application is on the mortgage search of British households, which is related our other work: the technical report in Myśliwski and Rostom (2022, MR hereafter). The data come from the Product Sales Database containing loan-level administrative data for all new mortgages in the UK. We focus on fixed rate mortgage products offered with two-, three-, and five-year durations; and to loan sizes less than £1M(illion) from 2016 and 2017. The sellers in our model are the six biggest banks that account for around 75% of all mortgages issued in the UK. Transaction data from these banks is representative of the British mortgage market as the their products are standardized compared to niche mortgagors, such as credit unions and building societies; the mortgage products we consider are available to all borrowers across the country. The omission of smaller lenders is also consistent with our structural model as, even in aggregate, they may have different cost distributions to the biggest lenders due to tighter budget constraints and more stringent capital requirements (Benetton (2021)). For further background details, we refer the reader to MR for descriptions of the UK mortgage market, dataset, and the descriptive evidence of price dispersion that motivates the use of a search model.

Like MR, markets in our application are defined temporally as 24 months over 2016 and 2017.⁸ Our empirical strategy, however, differs from MR as we combine parametric and nonparametric approaches for estimation. Specifically, we impose parametric assumptions on the price distribution and the firm's cost distribution, and we aggregate household prices and shares for each lender, conditional on loan and borrower characteristics, each month in order to estimate the consumer's cost distribution and inverting the firms' production costs according to our nonparametric identification strategy. Our estimation sample involves 1.3M individual transactions.

Since fixed rate mortgage prices are two-part tariffs with an initial fee and a (fixed) interest rate, we follow Allen, Clark, and Houde (2019)) and construct price based on monthly mortgage cost defined as $P = iL + \frac{Fee}{N}$, where *i* is the interest rate in the initial, fixed-rate period, *L* is the size of the loan, *Fee* is the up-front fee, and *N* is the initial period of the mortgage contract (24, 36, or 60 months). We detrend prices to remove dispersion from macroeconomic shocks (e.g. changes to the Bank of England's interest rate) and deflate them to January 2016 levels. We further normalize this measure to correspond to the median-sized loan in the data (around £150k). Even with the detrending and normalization, loans can still differ of observable characteristics which we label as z_l . Moreover, we allow for different borrower types (e.g. younger/older, more/less wealthy households) to draw their search costs from different distributions. Those characteristics are denoted by z_b . Table 1 lists out z_b and z_l .

7.2 Estimation

We observe borrowers with different prices for the same borrower-loan type (z_b, z_l) . First, we use the transaction prices to estimate the equilibrium price distribution under the assumption that it has a Beta distribution conditional on (z_b, z_l, Yr) , where Yr denotes the Year 2017 dummy, which we parameterize with a 3-way interactions of (z_b, z_l, Yr) . Then, we estimate the search cost distribution nonparametrically following our identification strategy in the previous sections. This requires us to aggregate consumers' prices in each market. In what follows, for any (z_b, z_l, Yr) , let $Y_{im}^{(z_b, z_l, Yr)}$

⁸While we do not formally test the random sampling assumption, by performing joint significance tests on a series of AR models, we fail to reject the no-autocorrelation hypothesis in market shares and prices for the majority of borrower-seller characteristic configurations for each lender.

and $P_{im}^{(z_b,z_l,Yr)}$ denote market share and price of lender *i* in month *m*, respectively, where we use the median price to represent $P_{im}^{(z_b,z_l,Yr)}$. Such representation seems reasonable for our application, for example, as Benetton (2021) has argued the UK mortgage market is characterized by posted rather than individualized prices since mortgage offers are based on observables with little room for negotiation. Variation in prices for each lender-market-product combination is therefore due to discretization of loan characteristics rather unobserved product differentiation. Nevertheless, we have explored other representative market prices as a robustness check by using we the average price or several other price quantiles. Our results are largely robust to different representations for $P_{im}^{(z_b,z_l,Yr)}$.

Variable	ole Values				
	Borrower characteristics: z_b				
Age	< 30, 30+				
Income	Below/above median				
FTB status	First-time-buyer (FTB), Non-FTB				
Location	Urban, Rural				
	Loan characteristics: z_l				
LTV	$\leq 70, 71\text{-}75, 76\text{-}80, 81\text{-}85, 86\text{-}90, 91\text{-}95$				
Deal length	2-, 3-, 5-year				
Term	< 10, (10, 15], (15, 20], (20, 25], (25, 30], (30, 35]				
Loan value	4 quartiles				

Table 1. List of observable heterogeneities in the model.

We estimate our model following these steps:

1. Estimate $F(\cdot|z_b, z_l, Yr)$. Assume transaction price has a Beta distribution conditional on (z_b, z_l, Yr) with the location and precision parameters governed by $\theta_1(z_b, z_l, Yr)$ and $\theta_2(z_b, z_l, Yr)$, respectively:

$$\theta_k\left(z_b, z_l, Yr\right) = \theta_{0k} + \theta_{1k}^{\top} z_b + \theta_{2k}^{\top} z_l + \theta_{3k} Yr + \theta_{4k}^{\top} z_b \circ z_l + \theta_{5k}^{\top} z_b \circ Yr + \theta_{6k}^{\top} z_l \circ Yr + \theta_{7k}^{\top} z_b \circ z_l \circ Yr,$$

for k = 1, 2. With an abuse of notation, we use \circ between two variables to denote a vector of all possible interactions of their components. Take price to be i.i.d. within each year and independent across years, we estimate $(\theta_1^{\top}, \theta_2^{\top})^{\top}$ by maximum likelihood⁹.

⁹We performed the estimation in R using the **betareg** package. This procedure reparameterizes the usual two shape parameters of a beta distribution into the mean and a precision parameter. For more details, see: https://cran.r-project.org/web/packages/betareg/index.html

Furthermore, since the beta distribution is supported on a unit interval, we first rescaled price and transformed it back. Specifically, we recovered the original pdf and cdf by: (i) let $\widetilde{P}_{im}^{(z_b,z_l)} = \frac{P_{im}^{(z_b,z_l)} - \underline{P}^{(z_b,z_l)}}{\overline{P}^{(z_b,z_l)} - \underline{P}^{(z_b,z_l)}}$ where $\left(\underline{P}^{(z_b,z_l)}, \overline{P}^{(z_b,z_l)}\right)$

- 2. Estimate $\widetilde{\mathbf{q}}(z_b, z_l, Yr)$. Use $\left\{ \left(Y_{im}^{(z_b, z_l, Yr)}, P_{im}^{(z_b, z_l, Yr)} \right) \right\}_{i=1,m=1}^{I,M}$ to perform constrained least squares based on minimizing $\left\| \mathbf{Y}^{(z_b, z_l, Yr)} - \widehat{\mathbf{X}}^{(z_b, z_l, Yr)} q \right\|^2$ subject to q summing to 1 and $q \ge 0$ to obtain $\widehat{\mathbf{q}}(z_b, z_l, Yr)$. $Y^{(z_b, z_l, Yr)}$ is the vector of market shares and $\left(\widehat{\mathbf{X}}_m^{(z_b, z_l, Yr)} \right)_{ik} = \frac{k}{I} \left(1 - \widehat{F} \left(P_{im} | z_b, z_l, Yr \right) \right)^{k-1}$ with $\widehat{F}(\cdot | z_b, z_l, Yr)$ taken from Step 1.
- 3. Estimate $G(\cdot|z_b)$. Estimate $G(\Delta_0(z_b, z_l, Yr)|z_b)$ by $1 \hat{q}_0(z_b, z_l, Yr)$ and $G(\Delta_k(z_b, z_l, Yr)|z_b)$ by $1 - \hat{q}_0(z_b, z_l, Yr) - (1 - \hat{q}_0(z_b, z_l, Yr)) \sum_{k'=1}^k \hat{q}_{k'}(z_b, z_l, Yr)$ for k > 0. $\hat{q}_0(z_b, z_l, Yr)$ is the proportion of borrowers using a broker and $\hat{q}_k(z_b, z_l, Yr) = \frac{(\hat{\mathbf{q}}_{(z_b, z_l, Yr)})_k}{1 - \hat{q}_0(z_b, z_l, Yr)}$ for $1 \le k < I$ and $q_I(z_b, z_l, Yr) = \frac{(\hat{\mathbf{q}}_{(z_b, z_l, Yr)})_I - \hat{q}_0(z_b, z_l, Yr)}{1 - \hat{q}_0(z_b, z_l, Yr)}$ where $\hat{\mathbf{q}}(z_b, z_l, Yr)$ is from Step 2.
- 4. Estimate $\Delta_k(z_b, z_l, Yr)$. Use $\widehat{F}(\cdot|z_b, z_l, Yr)$ from Step 1 to simulate prices and estimate $\Delta_k(z_b, z_l, Yr)$ by $E_{\widehat{F}(\cdot|z_b, z_l, Yr)}[P_{(1:k)}] - E_{\widehat{F}(\cdot|z_b, z_l, Yr)}[P_{(1:k+1)}]$ for k > 0 and $\Delta_0(z_b, z_l, Yr)$ by $\widehat{\phi}(z_b, z_l, Yr) + E_{\widehat{F}(\cdot|z_b, z_l, Yr)}[P_{(1:I)}] - E_{\widehat{F}(\cdot|z_b, z_l, Yr)}[P]$ where $\widehat{\phi}(z_b, z_l, Yr)$ is the median broker fee.
- 5. Estimate $h(\cdot|z_l, Yr)$. Let $\widehat{R}_{im}^{(z_b, z_l, Yr)} = P_{im}^{(z_b, z_l, Yr)} \frac{\sum_{k=1}^{I} (\widehat{\mathbf{q}}(z_b, z_l, Yr))_k k(1 \widehat{F}(P_{im}^{(z_b, z_l, Yr)}|z_b, z_l, Yr))^{k-1}}{\widehat{f}(P_{im}^{(z_b, z_l, Yr)}|z_b, z_l, Yr) \sum_{k=1}^{I} (\widehat{\mathbf{q}}(z_b, z_l, Yr))_k k(k-1)(1 \widehat{F}(P_{im}^{(z_b, z_l, Yr)}|z_b, z_l, Yr))^{k-1}}$ where $\left(\widehat{f}(\cdot|z_b, z_l, Yr), \widehat{F}(\cdot|z_b, z_l, Yr)\right)$ are from Step 1 and $\widehat{\mathbf{q}}(z_b, z_l, Yr)$ is from Step 2. Take $\widehat{R}_{im}^{(z_b, z_l, Yr)}$ to be i.i.d. across i for each m, whose distribution is independent across m, and $\widehat{R}_{im}^{(z_b, z_l, Yr)}$ has a Beta distribution conditional on (z_l, Yr_m) with the location and precision parameters governed by $\pi_1(z_l, Yr)$ and $\pi_2(z_l, Yr)$, respectively, such that, for k = 1, 2:

$$\pi_k (z_l, Yr) = \pi_{0k} + \pi_{1k}^{\top} z_l + \pi_{2k} Yr + \pi_{3k}^{\top} z_l \circ Yr.$$

Other than the parametric estimation of price and lender's cost distributions, we follow closely the nonparametric identification strategy from Step 2 to construct $\widehat{R}_{im}^{(z_b,z_l,Yr)}$ in Step 5. In particular, we do not specify $G(\cdot|z_b)$ to have a particular parametric distribution. Note that, relating to our discussion in Section 3.2, the variation in (z_l, Yr) allows us to identify $G(\cdot|z_b)$ at more points than the case without lender or market heterogeneity. Specifically, for a given z_b , we have estimates of $G(\Delta_k (z_b, z_l, Yr) | z_b)$ for $k = 1, \ldots, 5$ for each (z_l, Yr) combination. Following Sanches et al. (2018), for example, we can pool estimates of $\{G(\Delta_k (z_b, z_l, Yr) | z_b)\}_{k=1}^5$ across all (z_l, Yr) combinations by series estimation.

are the min and max prices; (ii) perform betareg with $\left\{\widetilde{P}_{im}^{(z_b,z_l)}\right\}_{i=1,m=1}^{I,M}$ to estimate $\widetilde{f}(\cdot|z_b,z_l)$ and $\widetilde{F}(\cdot|z_b,z_l)$; (iii) estimate $f(\cdot|z_b,z_l)$ and $F(\cdot|z_b,z_l)$ using $f(p|z_b,z_l) = \frac{1}{\overline{P}^{(z_b,z_l)} - \underline{P}^{(z_b,z_l)}} \widetilde{f}\left(\frac{p - \underline{P}^{(z_b,z_l)}}{\overline{P}^{(z_b,z_l)} - \underline{P}^{(z_b,z_l)}}|z_b,z_l\right)$ and $F(p|z_b,z_l) = \widetilde{F}\left(\frac{p - \underline{P}^{(z_b,z_l)}}{\overline{P}^{(z_b,z_l)} - \underline{P}^{(z_b,z_l)}}|z_b,z_l\right)$ respectively. In terms of our modeling choice, we choose the Beta distribution because it allows for poles at the boundaries. We model the equilibrium price distribution using an interactive 3-way model of (z_b, z_l, Yr) that allows for rich interactions between variables. We assume the search cost distribution is the same across the years, while the firm's costs in 2016 and 2017 can differ. These particular selections are made in parts due to estimation feasibility¹⁰ and for brevity's sake¹¹.

7.3 Results

Given the richness of our models, we will be selective in presenting our results. Before discussing the primitives, we highlight main observations from our estimates for the price distribution, which is a reduced form object in our model. We found that, in levels, most of the loan characteristics have the expected sign relative to the baseline category. For examples, larger and riskier loans with longer crediting periods are generally more expensive, and those with higher income also pay less, which is consistent with the notion that wealthier individuals are more financially literate and know they should be comparing prices. A comprehensive descriptive study of the interactive effects is beyond the scope of our empirical illustration¹², which focuses on the model primitives.

For the search cost distributions, Table 2 presents the search cost quartiles for 16 different borrower types (referred to as z_b -comb(inations). Age: L (below 30)/H (over 30). Inc(ome): L (below median)/H (above median). FTB (first time buyer status): Yes/No. Urb(an): U (urban area)/R (rural area). Columns 5-7 contains the median search cost in £/month in the initial period. Column 8 expressed in relative terms (divided by median and average monthly payment, respectively). Bootstrap standard errors in parentheses based on 500 replications.

The results reveal several interesting patterns: first, there are substantial differences between median search costs depending on different demographics. Younger buyers have on average lower median search costs, apart from low-income first-time buyers. This is true both in rural and urban areas. This shows that younger households are on average more financially literate. Secondly, on

¹⁰Estimates from a model takes a richer form than an interactive 4-way effect in the price distribution, which allows for another level of interaction with either a borrower's or lender's characteristic, begin to become unstable with respect to the initial value used in the optimization procedure.

¹¹For results that are stable, qualitative results on search cost and firm's cost distributions are the largely the same whether search cost distribution can vary with time or not. Here, we illustrate exclusion of time effects does not go to waste as it benefits identification of the search cost distribution.

¹²The multi-way interactive effects show many rationalizable patterns, but these are hard to succinctly summarize as there is sometimes non-uniformity in signs and directions with respect to a particular combination of interacting characcteristics. The latter point is not surprising given the complexity of what goes into a mortgage decision and the richness of our dataset, and our model has 160 parameters for the mean and precision apiece. The estimates of the price distribution are available upon request.

average, it is low income households which are affected by the information frictions to the largest extent. The differences between first and subsequent time buyers are also intuitive – first-time mortgagors have typically higher search costs and their distributions exhibit more dispersion. Generally, for all demographic groups, search costs amount to about 7-16% of monthly interest payments. Those numbers are an important benchmark for both policy makers who might want to improve information provision to disadvantaged households, as well as intermediaries in the market who need to know the value of information to optimally set their fee structures. Finally, we comment that while we notice some discrepancies between the results in Table 2 and the corresponding table in MR, the relative magnitudes of search cost all fall within a similar range, and the main qualitative conclusions remain unchanged despite differences in our use of parametric assumptions.

	z_b-	comb		1st quartile	2nd quartile	3rd quartile	% Median
Age	Inc	FTB	Urb				
\mathbf{L}	\mathbf{L}	Υ	\mathbf{R}	48.5(0.3)	51.5(0.3)	$53.1 \ (0.3)$	13.4%
Η	\mathbf{L}	Υ	R	39.2(1.5)	45.9(1.5)	54.8(1.5)	12.3%
\mathbf{L}	Η	Υ	R	23.8(0.7)	28.2(1.2)	34.8(1.9)	7.7%
Η	Η	Υ	R	43.5(14.9)	45.6(18.8)	54.6(23.2)	13.2%
\mathbf{L}	\mathbf{L}	Ν	R	29.6(1.0)	35.9(1.0)	45.3(1.1)	10.9%
Η	\mathbf{L}	Ν	R	34.7(3.2)	45.8(2.8)	52.7(2.8)	14.2%
\mathbf{L}	Η	Ν	R	32.5(1.7)	35.8(1.9)	37.4(2.5)	11.5%
Η	Η	Ν	R	41.4(9.4)	48.9(0.4)	51.9(0.4)	16.0%
\mathbf{L}	\mathbf{L}	Υ	U	41.6(15.9)	50.1 (4.7)	59.4(5.2)	12.9%
Η	\mathbf{L}	Y	U	35.4(18.5)	46.7(21.7)	114.8(25.6)	12.3%
\mathbf{L}	Η	Y	U	20.7(14.5)	25.7(20.6)	30.7(25.7)	6.9%
Η	Η	Υ	U	30.5(2.4)	38.9(2.5)	46.0(3.1)	11.0%
\mathbf{L}	\mathbf{L}	Ν	U	21.7(2.4)	27.0(2.7)	32.4(3.1)	8.1%
Η	\mathbf{L}	Ν	U	25.0(0.4)	32.4(0.3)	35.5(0.4)	10.0%
\mathbf{L}	Η	Ν	U	23.8(0.2)	27.6(0.3)	29.0(0.3)	9.0%
Η	Η	Ν	U	18.9(0.6)	34.9(0.7)	37.7(0.7)	11.6%

Table 2. Summary of quantiles search cost distributions.

On the supply-side, we estimate distributions of marginal costs of providing loans of different types and evaluate the extent of banks' market power by analysing distributions of price-cost margins. Notable observations from our estimates are in line with what is known about the mortgage market - e.g. that riskier loans (i.e. those with longer crediting term and higher LTV ratio) are costlier to supply. Figure 5 illustrates the latter by showing how the cost distribution generally shifts along the











For the most standard loan in the data (interest rate fixed for 2 years, crediting term between 20 and 25 years and total value just above the median), we plot the estimated beta distributions to show that lender's cost of providing that loan can be up to twice as much if LTV exceeds 85% compared to one with LTV in the range of 70-80%. The most risky loans, where the downpayment is below 10% are even more costly. We also observe that, generally, the distributions of the firms' costs appear to be higher in 2017 relative to 2016.

We end this illustration with a plot of the distribution of percentage markups implied by the model estimates in Figure 7. Similarly to the results obtained using a fully nonparametric approach and a larger sample of lenders and loans in MR, the distribution is right-skewed with mean just



above 10%.

While our goal here was to show that the model can produce meaningful estimates and can be used to analyze competition and quantify the extent of information frictions in a real-world, important industry, we note that the scope for its applications is much broader. In our other work (MR), we explain how banks' market power is affected by the presence of intermediaries and look at changes in markups in a counterfactual without brokers. Equipped with the estimates of all model primitives, particularly the estimates of $H(\cdot)$ and $G(\cdot)$, the model can also be used to answer other policy-relevant questions, such as effects of bank mergers, introduction of LTV caps etc.¹³

8 Some Discussions

Our paper focuses on the theoretical and methodological aspects of identifying and estimating a search model using aggregate data: market shares and market-level prices. The exposition of our theoretical models assumes each firm sets the same price to all consumers. The posted price framework, which is common in applications with retail data (e.g., online books (Hong and Shum (2006)), computer memory chips (Moraga-Gonzalez and Wildenbeest (2008)), is consistent with this assumption. Our estimation strategy is amenable to applications where the same firm could offer consumers different prices for the same (or similar) products/services. To proceed, one would need to construct aggregate price, P_{im}^{agg} , from individualized prices and impose the following restriction,

$$\mathbb{E}\left[Y_{im}|P_{im}^{agg}\right] = \sum_{k=1}^{I} q_k \frac{k}{I} \left(1 - F\left(P_{im}^{agg}\right)\right)^{k-1}.$$
(29)

¹³The main counterfactual study in MR was to calculate the value of information provided by mortgage brokers, i.e. the change in consumer welfare based on new simulated equilibrium prices when intermediaries are removed from the market. This application is related to the counterfactual study in Salz, who analyzed value of the brokers in New York City's trade-waste market.

We illustrated this in our application, where we used the median price as the representative market price based on the argument by Benetton (2021) that the UK mortgage market is well-represented by posted prices. Statistically, (29) can be thought of as a parametric restriction since it may not be implied by the model. Nevertheless, our proposal is appealing for two practical reasons. First, (29) can be interpreted as an approximation to the relation between market shares and representative prices analogous to Assumption D(ii). Second, our estimation strategy is very easy to implement in a variety of contexts. It can be applied to the pure search model, with or without product differentiation, as well as a model with an intermediary. Moreoever, computationally, estimation is based on our constructive identification strategy that is transparent and straightforward to apply.

The difficulty in estimating a search model like ours directly with individualized prices is not unique to our approach. It is also apparent in Salz (2020), even in the special case when $H(\cdot)$ can be treated as known. More specifically, with the presence of intermediary aside, while he showed the search cost distribution is nonparametric identified, his identification strategy is not constructive in the sense there are no direct estimators for the model primitives. For estimation, Salz took a full-solution approach¹⁴ with simulated method of moments based on matching moments of observed and theoretical prices; some parametric assumptions on the cost distribution and moment selections are also necessary for feasible estimation. We note there are other ways to estimate the model. For example, one could use equation (13) in Appendix B.2.1 in Salz (2020) that resembles the condition we have on market shares and prices (cf. (29)). However, exploiting that would require one to numerically solve differential equations (see page 60 in Salz (2020)) repeatedly, which may be a formidable computational task in real applications.

The point of the discussion above is to elucidate the challenge in estimating a search model nonparametrically when prices are individualized. This is perhaps not surprising. From a statistical standpoint, working with individualized prices means we have to account for a complicated missing data mechanism since only the purchased price is observed for each consumer amongst an unknown number of other prices. Economically, observing different prices offered by the same firm could imply that products may not be homogeneous, which brings us into the realms of product differentiation. Then, in building a structural econometric model, it places us somewhere between a very simple posted price framework with no supply-side heterogeneity like Hong and Shum (2006) and, at the other extreme, a sophisticated model of differentiated product search in Moraga-González, Sándor,

¹⁴When the empirical model admits multiple equilibria, a full-solution estimation method must account for multiple equilibria explicitly. This can substantially add to the computational burden. In contrast, our estimation procedure remains valid without change for the same reason highlighted by the two-step estimators in the literature on dynamic game estimation. We refer the reader to the Introduction of Srisuma (2013) for a discussion on the advantages of two-step estimators.

and Wildenbeest (2023) that requires a richer data environment for identification than price and market shares alone (e.g., needs data on individual search behaviour). Therefore, the parametric approach we have shown provides pragmatic ways to study search markets where products are not too heterogeneous with relatively modest requirements on data (i.e., prices and shares). Nevertheless, our model still paves the way to answer interesting counterfactual questions. E.g., Salz (2020) and Myśliwski and Rostom (2022) study the value of intermediaries in their respective applications. Other possible applications include quantification of effects on market price caused by changes to search costs (in the same vein as Choi, Dai, and Kim (2018), Moraga-González, Sándor, and Wildenbeest (2023)) or to production costs or market structure where the role of heterogeneous firms can be emphasized.

In the introduction, we gave intuitive and empirically motivated scenarios for when one may favor the incomplete information model over the complete information one. Another way to distinguish between the models is through statistical testing. Since we are not aware of an existing test for this purpose, an interesting direction for research would be to establish testable restrictions for each model and develop tests for them. One way to proceed in this direction could be along the lines of Grieco (2014), who showed, in the context of an entry game, how complete and incomplete information models can provide different testable restrictions. Analogous to his approach, common and private costs can also be present simultaneously in a search model like ours.

We end the paper by discussing possible extensions to our work. One modeling extension we can consider is for firms to have different probabilities of being found by consumers. Our results readily extend to this case if we assume an equilibrium exists where the optimal pricing strategies of firms are strictly increasing and share the same support. We are unable to prove such an equilibrium exists, but we are optimistic that it does based on positive results from the literature on asymmetric firstprice auctions¹⁵. The challenge stems from the fact that the (quasi-)inverse of the optimal pricing strategies are solutions to a system of nonlinear differential equations that are difficult to analyze.¹⁶

On the econometrics, our nonparametric identification strategy readily extends to include observed heterogeneity. All the assumptions made in Sections 2 to 5 can be written to condition on covariates (that can also include the number of firms). As in the auction model of Guerre et al. (2000), however, the nonparametric rate of convergence for the conditional distributions with continuous variables will be slower than that of the unconditional ones. In this case, we expect the quantile

¹⁵Some 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)).

 $^{^{16}}$ It is not trivial even to show the existence of such equilibrium numerically. For instance, in a related problem, numerical studies of the equilibrium in asymmetric auctions are a current topic of research - e.g. see the discussion in Fibich and Gavish (2011).

regression approach of Gimenes and Guerre (2020), recently developed to mitigate the dimensionality issue in the auction literature, to be applicable to our search model. 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 $\hat{h}(\cdot)$ and $\tilde{h}(\cdot)$ 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 derived 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 all of the propositions. 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 (9), we have $\mathbb{E}[Y_{im}|X_{im}] = X_{im}^{\top}\mathbf{q}$. Multiply both sides by X_{im} and take expectation yields $\mathbb{E}[X_{im}Y_{im}] = \mathbb{E}[X_{im}X_{im}^{\top}]\mathbf{q}$. Since $\mathbb{E}[X_{im}X_{im}^{\top}]$ has full rank, the proof follows from solving for \mathbf{q} .

PROOF OF LEMMA 4. By inspecting (4) and (5), $\beta(\cdot)$ is strictly increasing and continuously differentiable on $[\underline{R}, \overline{R})$ therefore $\beta^{-1}(\cdot) (= \xi(\cdot))$ exists. Using a change of variables argument, we have $f(p) = \frac{h(\beta^{-1}(p))}{\beta'(\beta^{-1}(p))}$. From (5), we can write $f(p) = \psi(\beta^{-1}(p))$, where $\psi(\cdot)$ is a real-value function defined on $[\underline{R}, \overline{R})$ such that

$$\psi(r) = \frac{\left(\sum_{k=1}^{I} q_k k \left(1 - H\left(r\right)\right)^{k-1}\right)^2}{\left(\sum_{k=2}^{I} q_k k \left(k - 1\right) \left(1 - H\left(r\right)\right)^{k-2}\right) \left(\sum_{k=1}^{I} q_k k \int_{s=r}^{\overline{R}} \left(1 - H\left(s\right)\right)^{k-1} ds\right)}.$$
(30)

Part (a) follows from $\inf_{p \in [\underline{P}, \overline{P}]} f(p) = \inf_{r \in [\underline{R}, \overline{R}]} \psi(r) \ge \frac{q_1^2}{\left(\sum_{k=2}^{I} q_k k(k-1)\right) \left(\overline{R} \sum_{k=1}^{I} q_k k\right)} > 0$. Part (b) follows from $\lim_{n \to \infty} -f(n) = \lim_{k \to \infty} -q_k(n) = \infty$. To obtain the expression in part (c), we know that $\beta(r)$ is

from $\lim_{p\to\overline{P}} f(p) = \lim_{r\to\overline{R}} \psi(r) = \infty$. To obtain the expression in part (c), we know that $\beta(r)$ is the maximizer of the following function,

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

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

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

Note that the cdf and pdf of P_{im} and R_{im} are related through $F(p) = H(\xi(p))$ and $f(p) = \xi'(p) h(\xi(p))$ respectively. Substitute these in and impose the first-order condition leads to,

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

Rearranging the relation above gives (10).

PROOF OF LEMMA 5. 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 (30), is a smooth functional of $H(\cdot)$ for all $r \in [\underline{R}, \overline{R})$.

PROOF OF LEMMA 7. Using the strict monotonicity property of $\beta(\cdot)$, for any $p \in [\underline{P}, \overline{P}]$ there exists a unique $r \in [\underline{R}, \overline{R}]$ such that $\beta(r) = p$

$$\left(\overline{P} - p\right)f(p) = \beta'\left(\widetilde{r}\right)\left(\overline{R} - r\right)f(\beta\left(r\right)) = \left(\overline{R} - r\right)h\left(\widetilde{r}\right)\frac{f(\beta\left(r\right))}{f(\beta\left(\widetilde{r}\right))},$$

for some $\tilde{r} \in (r, \overline{R})$. The first equality comes from replacing (p, \overline{P}) with $(\beta(r), \beta(\overline{R}))$ and, since $\beta'(\cdot)$ is continuously differentiable, applying the mean value theorem. The second equality uses the relation $f(\beta(\tilde{r})) = \frac{h(\tilde{r})}{\beta'(\tilde{r})}$, which holds through a change of variables argument. It can be verified from (30) that $\lim_{r\to\overline{R}} (\overline{R}-r) f(\beta(r)) = \lim_{\tilde{r}\to\overline{R}} (\overline{R}-\tilde{r}) f(\beta(\tilde{r}))$ is bounded away from zero and infinity, therefore $\lim_{p\to\overline{P}} (\overline{P}-p) f(p) = \lim_{\tilde{r}\to\overline{R}} (\overline{R}-\tilde{r}) h(\tilde{r})$. Since $h(\cdot)$ is bounded, it follows that $\lim_{p\to\overline{P}} (\overline{P}-p) f(p)$ exists and is equal to 0.

PROOF OF LEMMA 8. For part (a), given that $f^{\dagger}(p^{\dagger}) = \exp(-p^{\dagger}) f(\overline{P} - \exp(-p^{\dagger}))$ for all $p^{\dagger} \in [-\ln(\overline{P} - \underline{P}), \infty)$, boundedness follows if we can show $p \mapsto (\overline{P} - p) f(p)$ is continuous on $[\underline{P}, \overline{P})$ and $\lim_{p\to\overline{P}} (\overline{P} - p) f(p)$ exists. The former is a consequence of Lemma 5 and the latter follows from Lemma 7. For part (b), it suffices to show $f^{\dagger}(\cdot)$ has the same degree of smoothness as $f(\cdot)$. This holds for the same reason as Corollary 2 of Srisuma (2023).

PROOF OF LEMMA 9. From (18) 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 $\Psi\left(\mathbf{q}, f(P_{im}), F(P_{im})\right)$ 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 4(a), we know $\inf_{p \in [\underline{P} + \delta_M, \overline{P} - \delta_M]} \widetilde{f}(p) > c_0$ 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

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

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

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

so that

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

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 (19),

$$\widetilde{h}(r) - h(r) = J_{1}(r) + J_{2}(r) + J_{3}(r), \text{ where}$$

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

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

$$J_{3}(r) = \frac{1}{MIb_{h,M}} \sum_{m=1}^{M} \sum_{i=1}^{I} K\left(\frac{R_{im} - r}{b_{h,M}}\right) - h(r).$$

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

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

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

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

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

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

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

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

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

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

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 12 and 13 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 (22) instead of prices.

LEMMA 12. Suppose Assumption D' holds. Then the consumer's best response is a map $\sigma_D : \mathcal{L} \to \mathbb{S}^{I-1}$ such that for any L in \mathcal{L} ,

$$\sigma_D(L) = \begin{cases} 1 - G(\Upsilon_k(L)) & \text{for } k = 1\\ G(\Upsilon_{k-1}(L)) - G(\Upsilon_k(L)) & \text{for } k > 1 \end{cases}$$
(31)

LEMMA 13. 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 (22).

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 theorem follows immediately from here. \blacksquare

References

- Allen, J., R. Clark and J. F. Houde (2019): "Search Frictions and Market Power in Negotiated Price Markets", *Journal of Political Economy*, **127**, 1550–1598.
- [2] Andrews, D.W.K. (1995): "Nonparametric Kernel Estimation for Semiparametric Models," *Econometric Theory*, 11, 560–596.
- [3] Athey, S. and P. Haile (2007): "Nonparametric Approaches to Auctions," in Handbook of Econometrics, ed. by J. J. Heckman and E. Leamer, North-Holland.
- [4] 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.

- [5] Bénabou, R. (1993): "Search Market Equilibrium, Bilateral Heterogeneity, and Repeat Purchases," *Journal of Economic Theory*, 60, 140-158.
- [6] Benetton, M. (2021): "Leverage Regulation and Market Structure: A Structural Model of the U.K. Mortgage Market," *Journal of Finance*, 76, 2997–3053.
- [7] 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.
- [8] 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.
- [9] Burdett, K., and K. Judd (1983): "Equilibrium Price Dispersion," *Econometrica*, **51**, 955–969.
- [10] Choi, M., A. Y. Dai and K. Kim (2018): "Consumer Search and Price Competition", Econometrica, 86, 1257-1281.
- [11] Diamond, P.A. (1971): "A Model of Price Adjustment," Journal of Economic Theory, 3, 156– 168.
- [12] 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.
- [13] Fan, J. and Q. Yao (2003): Nonlinear Time Series: Nonparametric and Parametric Methods, Springer-Verlag.
- [14] Fibich, G. and N. Gavish (2011): "Numerical Simulations of Asymmetric First-Price Auctions," *Games and Economic Behavior*, **73**, 479-495.
- [15] Gimenes, N. and E. Guerre (2022): "Quantile Regression Methods for First-Price Auction: A Signal Approach," *Journal of Econometrics*, **226**, 224-247.
- [16] Guerre, E., I. Perrigne and Q. Vuong (2000): "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 68, 525-574.
- [17] Hansen, B. (2008): "Uniform Convergence Rates for Kernel Estimation with Dependent Data," *Econometric Theory*, 24, 726-748.

- [18] Hansen, B. and S. Lee (2019): "Asymptotic Theory for Clustered Samples," Journal of Econometrics, 210, 268-290.
- [19] Härdle, W. (1991): Smoothing Techniques with Implementation in S. New York: Springer Verlag.
- [20] 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.
- [21] Hong, H. and M. Shum (2006): "Using Price Distribution to Estimate Search Costs," RAND Journal of Economics, 37, 257-275.
- [22] Honka, E. and P. Chintagunta (2017): "Simultaneous or Sequential? Search Strategies in the US Auto Insurance Industry," *Marketing Science*, 36, 21-42.
- [23] Karunamuni R.J. and S. Zhang (2008): "Some Improvements on a Boundary Corrected Kernel Density Estimator," *Statistics & Probability Letters*, 78, 499-507
- [24] Lebrun, B. (1999): "First Price Auctions in the Asymmetric N Bidder Case," International Economic Review, 40, 125-142.
- [25] Li, H. and N. Liu (2015): "Nonparametric Identification and Estimation of Double Auctions with Bargaining,", Working Paper, Shanghai University of Finance and Economics.
- [26] 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.
- [27] MacMinn, R.D. (1980): "Search and Market Equilibrium," Journal of Political Economy, 88, 308-327.
- [28] Masry, E. (1996): "Multivariate Local Polynomial Regression for Time Series: Uniform Strong Consistency and Rates," *Journal of Time Series Analysis*, 17, 571-599.
- [29] McAfee, R.P. and J. McMillan (1988): "Search Mechanisms," Journal of Economic Theory, 44, 99-123.
- [30] Milgrom, P. and I. Segal (2002): "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 70, 583-601.
- [31] Moraga-González, J.L. and M. Wildenbeest (2008): "Maximum Likelihood Estimation of Search Costs," *European Economic Review*, **52**, 820-48.

- [32] Moraga-González, J.L., Z. Sándor and M. Wildenbeest (2012): "Consumer Search and Prices in the Automobile Market," Working Paper, University of Indiana.
- [33] Moraga-González, J.L., Z. Sándor and M. Wildenbeest (2013): "Semi-nonparametric Estimation of Consumer Search Costs," *Journal of Applied Econometrics*, 28, 1205-1223.
- [34] 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.
- [35] Moraga-González, J.L., Z. Sándor and M. Wildenbeest (2023): "Consumer Search and Prices in the Automobile Market," *Review of Economic Studies*, **90**, 1394-1440.
- [36] Myśliwski, M. and M. Rostom (2022): "Value of Information, Search and Competition in the UK Mortgage Market," Staff Working Paper 967, Bank of England.
- [37] Pereira, P. (2005): "Multiplicity of Equilibria In Search Markets with Free Entry and Exit," International Journal of Industrial Organization, 23, 325-339.
- [38] Postel-Vinay, F. and J.-M. Robin (2002): "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, 70, 2295-2350.
- [39] Salz, T. (2020): "Intermediation and Competition in Search Markets: An Empirical Case Study," forthcoming in the Journal of Political Economy.
- [40] 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.
- [41] Srisuma, S. (2013): "Minimum Distance Estimators for Dynamic Games," Quantitative Economics, 4, 549-583.
- [42] Srisuma, S. (2023): "Uniform Convergence Rates for Nonparametric Estimators of a Density Function and its Derivatives when the Density has a Known Pole," *Working Paper*.
- [43] Stigler, G. (1961): "The Economics of Information," Journal of Political Economy, 69, 213-225.
- [44] Wildenbeest, M. R. (2011): "An Empirical Model of Search with Vertically Differentiated Products," RAND Journal of Economics, 42, 729-757.
- [45] Stone, C.J. (1982): "Optimal Rate of Convergence for Nonparametric Regressions," Annals of Statistics, 10, 1040-1053.