A Structural Analysis of Wholesale Used-Car Auctions: Nonparametric Estimation and Testing of Dealers' Valuations

Kyoo il Kim and Joonsuk Lee*
UCLA

April 5, 2006

Abstract

Wholesale used-car markets have widely utilized ascending auctions for trading dealers' inventories. To study the latent demand structure of these markets, one needs to recover the underlying valuations of dealers from the bidding data observed at the auctions. Exploiting a new, rich data set on a wholesale used-car auction, we estimate the distribution of bidders' valuations nonparametrically under the symmetric independent private values (IPV) framework and nonparametrically test the validity of the IPV assumption. We develop a new nonparametric test of IPV for the case where the number of potential bidders is not observable to an econometrician by extending the work of Athey and Haile (2002). This is done utilizing and extending the methodology provided in Song (2005) for identifying and estimating the distribution of valuations nonparametrically when the number of potential bidders of ascending auctions is unknown in an IPV situation. Unlike previous work on ascending auctions, our estimation and testing methods use more information from observed losing bids by virtue of the rich structure of our data. We find that the null hypothesis of IPV is not rejected with our sample after controlling for observed auction heterogeneity and therefore our estimation result is a good approximation of the underlying distribution of dealers' valuations.

*Department of Economics, University of California, Los Angeles, CA 90095-1477, USA. Email: kyookim@ucla.edu and joonsuk@ucla.edu We are grateful to Dan Ackerberg, Sushil Bikhchandani, Jinyong Hahn, Joe Ostroy, John Riley, Jean-Laurent Rosenthal, Unjy Song, Mo Xiao and Bill Zame for helpful comments. Special thanks to Shijin Yoo for introducing us to the auction and to Ho-Soon Hwang and Hyun-Do Shin for allowing an access to the invaluable data. All errors are ours.

1 Introduction

Ascending-price auctions, or English auctions, are one of the oldest trading mechanisms that have been used for selling a variety of goods ranging from fish to airwaves. Wholesale used-car markets are among those that have utilized ascending auctions for trading inventories among dealers.

In this paper, we exploit a new, rich data set on a wholesale used-car auction to study the latent demand structure of the auction with a structural econometric model. The structural approach to study auctions assumes that bidders use equilibrium bidding strategies, which are predicted from game-theoretic models, and tries to recover bidders' private information from observables. The most fundamental issue of structural auction models is estimating the unobserved distribution of bidders' willingness to pay from the observed bidding data.[1,2]

We treat our data as a collection of independent, single-object, ascending auctions. So we are not studying any feature of multi-object or multi-unit auctions in this paper though there might be some substitutability and path-dependency in the auctions that we study. To reduce the complexity of our analysis, we defer those aspects to a later study. Also, to deal with the heterogeneity in the objects of our auctions, we first assume there is no unobserved heterogeneity in our data and then control for the observed heterogeneity in the estimation stage.[3]

In our model, there are symmetric, risk-neutral bidders while the number of potential bidders is assumed to be exogenous and unknown. Game-theoretic models of bidders' valuations can be classified according to informational structures.[4] We first assume that the information structure of our model follows the independent private-values (hereafter IPV) paradigm and then we develop a new nonparametric test with an unknown number of potential bidders to check the validity of the IPV assumption after we obtain our estimates of the distribution nonparametrically under the IPV paradigm by extending the work of Athey and Haile (2002). We find that the null hypothesis of IPV is not rejected in our sample and therefore our estimation result under the assumption remains a valid approximation of the underlying distribution of dealers' valuations.

To deal with the problem of unknown number of potential bidders, we utilize and extend the

[1] Paarsch (1992) first conducted a structural analysis of auction models. He adopted a parametric approach to distinguish between independent private-values and pure common-values models. For the first-price sealed-bid auctions, Guerre, Perrigne, Vuong (2000) provided an influential nonparametric identification and estimation results, which was extended by Li, Perrigne, and Vuong (2000, 2002) and Campo, Perrigne, and Vuong (2003).

[2] See Hendricks and Paarsch (1995) and Laffont (1997) for surveys of the early empirical works. For a survey of latest works in the literature, see Athey and Haile (2005).

[3] See Krasnokutskaya (2004) for a novel empirical work regarding unobserved auction heterogeneity.

[4] Valuations can be either from the private-values paradigm or the common-values paradigm. Private-values are the cases where a bidder's valuation depends only on her own private information while common-values refer to all the other general cases. IPV is a special case of private-values where all bidders have independent private information.

methodology provided in Song (2005) for the nonparametric identification and estimation of the distribution when the number of potential bidders of ascending auctions is unknown in an IPV setting. Unlike previous work on ascending auctions, the rich structure of our data, especially the absence of jump bidding, enables us to utilize more information from observed losing bids for our nonparametric estimation and testing. This is an important advantage of our data compared to any other existing studies of ascending auctions.

There is considerable research on the first-price sealed-bid auctions in the structural analysis literature; however, there are not as many on ascending auctions. One of the possible reasons for this scarcity is the discrepancy between the theoretical model of ascending auctions, especially the model of Milgrom and Weber (1982), and the way real-world ascending auctions are conducted. Another important reason preventing structural analysis of ascending auctions is the difficulty of getting a rich and complete data set that records ascending auctions.

Given these difficulties, our paper contributes to the empirical auctions literature by providing a new data set that is free from both problems and by conducting a sound structural analysis using recently developed econometric techniques.

Milgrom and Weber (1982) (hereafter MW) modeled ascending auctions as a button auction, an auction with full observability of bidders' actions and, most importantly, with the irrevocable exits assumption, an assumption that a bidder is not allowed to place a bid at any higher price once she drops out at a lower price. This assumption significantly restricts each bidder's strategy space and makes the auction game easier to analyze. Without this assumption for modeling ascending auctions one would have to consider dynamic features allowing bidders to update their information and therefore valuations continuously and to re-optimize during an auction. After MW, the button auction assumption was widely accepted by almost all the following studies on ascending auctions, both theoretical and empirical, because of its elegant simplicity.[5]

However, in almost all the real-world ascending auctions, we do not observe irrevocable exits. Recently, a paper by Haile and Tamer (2003) (hereafter HT) conducted an empirical study of ascending auctions without specifying such details as irrevocable exits. In their nonparametric analysis of ascending auctions with IPV, HT adopted an incomplete modelling approach with relaxing the button auction assumption and imposing only two axiomatic assumptions on bidding behavior. The first assumption of HT is that bidders do not bid more than they are willing to pay and their second assumption is that bidders do not allow an opponent to win at a price they are able to beat. With these two assumptions and known statistical properties of order statistics, HT

[5] There are few exceptions, e.g. Harstad and Rothkopf (2000) and Izmalkov (2003) in theoretical literature and Haile and Tamer (2003) in empirical literature. Also see Bikhchandani and Riley (1991, 1993) for extensive discussions of modeling ascending auctions.

nonparametrically estimated upper and lower bounds of the underlying distribution function.

The reason HT could only estimate bounds and could not make an exact interpretation of losing bids is because they did not impose the button auction assumption and allowed free forms of ascending auctions. Athey and Haile (2002) noted the difficulty of interpreting losing bids saying "*In oral 'open outcry' auctions we may lack confidence in the interpretation of losing bids below the transaction price even when they are observed.*"[6] Our main difference from HT is that we are able to provide a stronger interpretation of observed losing bids. Within ascending auctions, there are a few variants that differ in the exact way the auction is conducted. Among those variants, the distinction between one in which bidders call prices and another one in which an auctioneer raises prices has very important theoretical and empirical implications. The former is what Athey and Haile called as "open outcry" auctions and the latter is what we are going to exploit with our data. The main difference is that the former allows jumps in prices but the latter does not allow those jumps.

It is well known in the literature that it is empirically difficult to distinguish between private-values and common-values models with actual auction data. However, there have been a few recent attempts to develop these tests. Hong and Shum (2003) estimated and tested general private-values and common-values models in ascending auctions with a certain parametric modelling assumption using a quantile estimation method. Hendricks, Pinkse, and Porter (2003) developed a test based on data of the winning bids and the ex post values in first-price sealed-bid auctions. Athey and Levin (2001) also used an ex post data to test the existence of common values in first-price sealed-bid auctions. Most recently, Haile, Hong, and Shum (2003) developed a new nonparametric test for common-values in first-price sealed-bid auctions using Guerre, Perrigne, Vuong (2000)'s two-stage estimation method.

This paper contributes to this literature by developing the first nonparametric test of IPV in ascending auctions without requiring the information about the potential number of bidders.

The remainder of this paper is organized as follows. We provide briefly some background on the wholesale used-car markets in general and on the wholesale used-car auctions in Korea, where our data comes from, in the next section. We describe the data we use and give some summary statistics in Section 3. Section 4 presents the empirical model and our empirical strategy that includes identification and estimation. We develop our new nonparametric test in Section 5. We provide the empirical results in Section 6 and Section 7 discusses the implications and extensions of the results. Section 8 concludes and additional technical details are included in the appendix

[6] Bikhchandani, Haile, and Riley (2002) also show there are generally multiple equilibria even with symmetry in ascending auctions so that we have to be careful interpreting observed bidding data; however, they show that with private values and weak dominance, there exists uniqueness.

and the online supplementary technical appendix.

2 Wholesale Used-car Auctions

Any trade in wholesale used-car auctions is intrinsically intended for a future resale of the car. Used-car dealers participate in wholesale auctions to manage their inventories in response to either immediate demand on hand or anticipated demand in the future. The supply of used-car is generally considered not so elastic but the demand side is at least as elastic as the demand for new cars. In the wholesale used-car auctions, some bidders are general dealers who do business in almost every model; however, others may specialize in some specific types of cars. Also, some bidders are from large dealers, while many others are from small dealerships.[7]

For dealers, there exists at least two different situations about the latent demand structure. First, a dealer may have a specific demand, i.e. a pre-order, from a final buyer on hand. This can happen when a dealer gets an order for a specific used-car but does not have one in stock or when a consumer looks up the item list of an auction and asks a dealer to buy a specific car for her.[8] We call this as "demand-on-hand" situation. In this case, we make a conjecture that a dealer already knows how much money she can make if she gets the car and sells it to the consumer. Second, though a dealer does not have any specific demand on hand, but she anticipates some demand for a car in the auction in near future from the analysis of the market. We call this as "anticipated-demand" situation. In this case, we make a conjecture that a dealer is not so certain and confident about her anticipation of future market condition and has some incentives to find out other dealers' opinions about it.

The auction data in our paper comes from an offline auction house located in Suwon, Korea. Suwon is located within an hour drive south of Seoul, the capital of Korea. It opened in May 2000 and it is the first fully-computerized wholesale used-car auction house in Korea. The auction house engages only in the wholesale auctions and it has held wholesale used-car auctions once a week since its opening.

The auction house mainly plays the role of an intermediary as many auction houses do. While sellers can be anyone, both individuals and firms, who wants to sell her own car through the auction, only a used-car dealer who is registered as a member of the auction house can participate in the auction as a buyer. At the beginning, the number of total memberships was around 250, and now it has grown to about 350 and the set of the members is a relatively stable group of dealers, which

[7] These suggest that there may exist asymmetry among bidders in the auctions, which we are not studying in this paper, but plan to do in the future.

[8] The item list is publicly available on the website 3-4 days prior to each auction day and these kinds of delegated biddings seem to be not uncommon from casual observations.

makes the data from this auction more reliable to conduct meaningful analyses than those from online auctions in general.

Roughly, about a half of the members come to the auction each week. 600-1000 cars are auctioned on a single auction day and 40-50 per cent of those cars are actually sold through the auctions, which implies that a typical dealer who comes to the auction house on an auction day gets 2-4 cars/week on average. Bidders, i.e. dealers, in the auction have resale markets and, therefore, we can view this as if they try to win these used-cars auctioned only to resell them to the final consumers or, rarely, to other dealers.

The auction house generates its revenue by collecting fees from both a seller and a buyer when a car is sold through the auction. The fee is 2.2% of the transaction price, both for the seller and the buyer. The auction house also collects a fixed fee, about 50 US dollars, from a seller when the seller consigns her car to the auction house. The auction house's objective should be long-term profit maximization. Since there exists repeated relationship between dealers and the house, it may be important for the house to build favorable reputation.[9] There also exists a competition among three similar wholesale used-car auction houses in Korea. In this paper, we ignore any effect from the competition and consider the auction house as a single monopolist for simplicity.

3 Data

3.1 Description of the Data

The auction game itself is an interesting and unique variant of ascending auctions. There is a reserve price set by an original seller with consultations from the auction house. An auction starts from an opening bid, which is somewhere below the reserve price. The opening bids are made public on the auction house's website two or three days before an auction day. After an auction starts, the current price of the auction increases by a small, fixed increment, about 30 US dollars for all price ranges, whenever there are two or more bidders at the price who press their buttons beneath their desks in the auction hall.[10]

In the auction hall, there are big screens in front that display pictures and key descriptions of a car of the current auction. The current price is also displayed and updated real-time. Next to the screens, there are two important indicators for information disclosure. One of them resembles a traffic light with green, amber, and red lights, and the other is a sign that turns on when the

[9] An original seller's goal is to sell her car at the highest possible price near the time she wants to sell it after considering the trade-off between the price and the possibility of being sold.

[10] When the current price is below the reserve price, there only need to be one bidder to press the button for an auction to continue.

current price rises above the reserve price, which means that the reserve price is made public once the current price reaches the level.

The three-colored lights indicate the number of bids at the current price. The green light signals three or more bids, yellow means two bids, and red indicates that there is only one bid at the current price while the current price is above the reserve price. When the current price is below the reserve price, they are indicating two or more, one, and zero bids respectively. This indicator is needed because, unlike the usual open out-cry ascending auctions, the bidders in this auction do not see who are pressing their buttons and therefore do not know how many bidders are placing their bids at the current price. With the colored lights, bidders only get somewhat incomplete information on the number of current bidders and they never observe the identities of the current bidders.

There is a short length of a single time period, such that all the bids made in the same period are considered as made at the same price. A bidder can indicate her willingness to buy at the current price by pressing her button at any time she wants, i.e. exit and reentry are freely allowed. The auction ends when three seconds have passed after there remains only one active bid at the current price. When an auction ends at the price above the reserve price, the item goes to the bidder who presses last, but when it ends at the price below the reserve price, the item is not sold.

Our data set consists of 59 weekly auctions from a period between October 2001 and December 2002. While all kinds of used-cars including passenger cars, buses, trucks, and other special-purpose vehicles, are auctioned in the auction house during the period, we select only passenger cars to ensure a minimum homogeneity of the objects in the auction. For those 59 weekly auctions, there were 28334 passenger-car auctions. However, we use only data from those auctions where a car is sold and there exist at least four bidders above the reserve price in an auction. After we apply this criteria, we have 5184 cars, 18.3% of all the passenger-cars auctioned. Although this step restricts our sample from the original data set, we do this because we need four meaningful bids, which means they should be above the reserve price, to conduct our nonparametric test and we think our analysis remains meaningful since what are important in generating most of the revenue for the auction house are those cars where there are at least four bidders above the reserve price.[11]

In the actual estimation stage of our econometric analysis of the sample, we are forced to forgo additional cars out of those 5184 auctions because of the ties among the second, third, and fourth highest bids.[12,13] We remove the total of 1358 cars and we use the final sample of 3826 auctions.[14]

[11] Since we maintain the assumption of no unobserved auction heterogeneity and the exogeneity of the number of bidders, the sample selection is not a concern in this paper.

[12] We need to do this solely because there exists a technical difficulty in our estimation procedure.

[13] We never observe the first highest bids in ascending auctions separately. They are always the same as the second highest bids (plus the minimum increment.)

[14] The second-highest bid and the third-highest bid are tied in 842 auctions. In 624 auctions, the third-highest bid

Among those, 1676 cars are from the maker Hyundai, the biggest Korean carmaker, 1186 cars are from the maker Daewoo, now owned by General Motors, and the remaining 964 cars are from the other makers like Kia, Sssangyong, and Samsung while most of them are from Kia, which merged with Hyundai in 1999. Some important summary statistics of the sample are provided in Table 1 and the market shares of major carmakers in the sample are presented in Table 2.

Available data includes the detailed bid-level, button-pressing log data for every auction. Auction covariates, very detailed characteristics of cars auctioned, are also available. The covariates available includes each car's make, model, production date, engine-size, mileage, rating - the auction house inspects each car and gives a 10-0 scaled rating to each car, which can be thought of a summary statistic of the car's exterior and mechanical condition, transmission-type, fuel-type, color, options, purpose, body-type etc. We also observe the opening bids and the reserve prices of all auctions. Some bidder-specific covariates such as identities, locations, 'members since' date are also available. And the date of title change is available for each car sold, which may be used as a proxy for the resale date in a future study. Last, we only observe the information on 'who' come to the auction house at 'what time' of an auction day for a very rough estimate on the potential set of bidders for an auction.

Here is a descriptive snapshot of a typical auction day, September 4th, 2002, which is randomly picked. A total of 567 cars were auctioned on that day, 386 cars of which were passenger cars and the remaining 181 were full-size vans, trucks, buses, etc. Since this auction day was the first week of the month, there was relatively a small number of cars (the number of cars auctioned greatest in the last auctions of the month.) 248 cars (43 percent) were sold through the auction and, among those unsold, at least 72 cars sold afterwards through post-auction bargaining, or re-auctioning next week, etc. The log data shows that the first auction of the day started at 13:19 PM and the last auction of the day ended at 16:19 PM. It only took 19 seconds per auction and 43 seconds per transaction. 152 ID cards (132 dealers since some dealers have multiple IDs) were recorded as entered the house. On average each ID placed bids for 7.82 auctions during the day. There were 98 bidders who won at least one car but 40 bidders did not win a single car. On average, each bidder won 1.8 cars and there are three bidders who won more than 10 cars. Among 386 passenger cars, there were at least one bid in 218 auctions. Among those 218 auctions, 170 cars were successfully sold through auctions and 48 were unsold.

3.2 Interpretation of Losing Bids

As we briefly mentioned in the Introduction, a strong interpretation of losing bids in our data is one of the main strong points of this study that distinguishes it from other work on ascending auctions.

and the fourth-highest bid are tied, and for 108 auctions the second-, third-, and fourth-highest bids are all tied.

Our basic idea is that with the assumption of IPV and with some institutional features of this auction such as a fixed discrete increments of prices raised by an auctioneer, we are able to make strong interpretations of losing bids above the reserve price and therefore identify the corresponding order statistics of valuations from observed bids.

We are able to do this based on some equilibrium implications of the auction game in the spirit of Haile and Tamer (2003). So, in the end, without imposing the button auction model explicitly, we are able to treat some observed bids as if they come from a button auction. Then, we use this information from observed bids to estimate the exact underlying distribution of valuations.[15] Within the private values paradigm, in ascending auctions without irrevocable exits, the last price at which each bidder shows her willingness to win, i.e. each bidder's final exit price, can be directly interpreted as her private value, and with symmetry, to relevant order statistics because it is a weakly dominant strategy, and with high probabilities a strictly dominant strategy, for a bidder to place a bid at the highest price she can afford.[16] Here, we assume there is no 'real' cost associated with each bidding action, i.e. an additional button pressing requires a bidder to consume zero or negligible additional amount of energy, given that the bidder is interested in a car being auctioned.

The basic setup in our auction game model is that we have a collection of single-object ascending auctions with all symmetric, risk-neutral bidders. The number of potential bidders in any auction is assumed to be exogenous and the number of potential bidders is never observed to any bidders nor an econometrician. For the information structure, we assume independent private values (IPV), which means a bidder's bidding strategy does not depend on any other bidder's action or private information in this auction we study, i.e. an ascending auction with exogenous rise of prices with fixed discrete increments with no jump. Therefore, with IPV, the fact that any bidder has very limited observability of other current bidders' identities or bidding activities does not cause any big problem in this auction.

Within this setting, while it is obviously weakly dominant strategies for a bidder to press her bidding button at the last chance she can press, it does not necessarily guarantee that a bidder's actual observed last press corresponds to the bidder's true maximum willingness to pay. A strictly positive potential benefit from pressing at the last price a bidder can afford will ensure that our strong interpretation of losing bids is a close approximation and such a chance for a strictly positive benefits can be present when there exists a possibility that all the other remaining bidders simultaneously exit at the price, which a bidder is considering to press her button at. When the

[15] Actually, HT noted that if the true underlying model is the button auction, then their two bounds collapsed to a single distribution, which is also the exact estimate.

[16] In common values model, this is not the case and the analysis is much more complicated because a bidder may try not to press her button unless it is absolutely necessary because of strategic consideration to conceal her information. See Riley (1988) and Bikhchandani and Riley (1991, 1993).

auction price rises continuously as in Milgrom and Weber (1982)'s button auction, with continuous distribution of valuations, the probability of such event is zero at any price; however, when price rises in a discrete manner with a minimum increment as in this auction we study, this is not a measure-zero event, especially for the higher bidders like the second-, third-, and fourth-highest bidders as well as the winner.[17,18]

4 Model and Estimation

4.1 Empirical Model and Identification

This section describes the basic set-up of an IPV model we analyze. Consider a Wholesale Used-Car Auction (hereafter, WUCA) of a single object with the number of risk-neutral potential bidders, $N \geq 2$, drawn from $p_n = Pr(N = n)$. Each potential bidder i has the valuation V^i, which is independently drawn from the absolutely continuous distribution $F(\cdot)$ with support $\mathcal{V} = [\underline{v}, \bar{v}]$. Each bidder knows only his valuation but the distribution $F(\cdot)$ and the distribution p_n are common knowledge. By the design of WUCA, we can treat it as a button auction if we disregard the minimum increment. The minimum increment (about 30 dollars) in WUCA is small relative to the average car value (around 3,000 dollars) sold in WUCA, which is about one percent of the average car value.

Hence, in what follows, we simply disregard the existence of the minimum increment in WUCA to make our discussion simple and the bounds estimation implied by the minimum increment is handled in Section 7.2. Therefore, if we observe the number of potential bidders and any i^{th} order statistic of the valuation (identical to i^{th} order statistic of the bids), then we can identify the distribution of valuations from the cumulative density function (CDF) of the i^{th} order statistic as done in many previous literatures.[19] Define the CDF as

$$G^{(i:n)}(v) = H(F(v); i:n) = \frac{n!}{(i-1)!(n-i)!} \int_0^{F(v)} t^{i-1}(1-t)^{n-i} dt$$

Then, we obtain the distribution of the valuations $F(\cdot)$ from

$$F(v) = H^{-1}(G^{(i:n)}(v); i:n)$$

[17] More precisely, for this argument, we need to model the auction game such that each bidder can only press the button at any price simultaneously and at most once and the auction stops when only one bidder presses her button at a price. Actually, although this is not an exact modeling of the auction we study, we think this modeling is an acceptable and close approximation of the real auction.

[18] Actual observed bidding may depend on each bidder's subjective probability regarding the simultaneous exits of the all active bidders since a bidder does not know who the current active bidders are.

[19] See Arnold et al. (1992) and David (1981) for extensive statistical treatments on order statistics.

However, in the auction we consider, we do not know the exact number of potential bidders in a given auction and the number of potential bidders varies over different auctions. Nonetheless we can still identify the distribution of valuations $F(\cdot)$ following the methodology proposed by Song (2005), since we observe several order statistics in a given auction. Song (2005) showed that an arbitrary absolutely continuous distribution $F(\cdot)$ is nonparametrically identified from observations of any pair of order statistics from an I.I.D. sample, even when the sample size, n, is unknown and stochastic. The idea is that we can reinterpret the density of the k_1^{th} highest value $\widetilde{V}$ conditional on the k_2^{th} highest value V as the density of the $(k_2-k_1)^{th}$ order statistic from a sample of (k_2-1) following $F(\cdot)$. Denote the probability density function (PDF) of the i^{th} order statistic of the n sample (I.I.D.) as $g^{(i;n)}(v)$

$$g^{(i:n)}(v) = \frac{n!}{(i-1)!(n-i)!}[F(v)]^{i-1}[1-F(v)]^{n-i}f(v) \tag{1}$$

and the joint density of the i-th and the j-th order statistics of the n sample (I.I.D.) for $n \geq j \geq i \geq 1$ as $g^{(i,j:n)}(\widetilde{v}, v)$

$$g^{(j,i:n)}(\widetilde{v}, v) = \frac{n![F(v)]^{i-1}[F(\widetilde{v})-F(v)]^{j-i-1}[1-F(\widetilde{v})]^{n-j}f(v)f(\widetilde{v})}{(i-1)!(j-i-1)!(n-j)!}I_{\{\widetilde{v}>v\}}. \tag{2}$$

Then, the density of $\widetilde{V}$ conditional on V, $p_{k_1|k_2}(\widetilde{v}|V=v)$ can be written

$$\begin{aligned}
& p_{k_1|k_2}(\widetilde{v}|v) = \frac{g^{(n-k_1+1,n-k_2+1:n)}(\widetilde{v}, v)}{g^{(n-k_2+1:n)}(v)} \qquad (3) \\
= \;& \frac{(k_2-1)!}{(k_2-k_1-1)!(k_1-1)!}\frac{(F(\widetilde{v})-F(v))^{k_2-k_1-1}(1-F(\widetilde{v}))^{k_1-1}f(\widetilde{v})}{(1-F(v))^{k_2-1}}I_{\{\widetilde{v}\geq v\}} \\
= \;& \frac{(k_2-1)!}{(k_2-k_1-1)!(k_1-1)!}((1-F(v))F(\widetilde{v}|v))^{k_2-k_1-1} \\
& \times\frac{((1-F(v))(1-F(\widetilde{v}|v)))^{k_1-1}f(\widetilde{v}|v)(1-F(v))}{(1-F(v))^{k_2-1}} \cdot I_{\{\widetilde{v}\geq v\}} \\
= \;& \frac{(k_2-1)!}{(k_2-1-k_1)!(k_2-1-(k_2-k_1))!} \\
\times \;& F(\widetilde{v}|v)^{k_2-1-k_1}(1-F(\widetilde{v}|v))^{k_1-1}f(\widetilde{v}|v)\cdot I_{\{\widetilde{v}\geq v\}} \\
\equiv \;& g^{(k_2-k_1:k_2-1)}(\widetilde{v}|v),
\end{aligned}$$

where $f(\widetilde{v}|v)(g^{(\cdot)}(\widetilde{v}|v))$ denotes the truncated density of $f(\cdot)(g^{(\cdot)}(\cdot))$ truncated at v from below and $F(\widetilde{v}|v)$ denotes the truncated distribution of $F(\cdot)$ truncated at v from below[20]. This interpretation comes from the probability density function, $g^{(i;n)}(v)$ of the i^{th} order statistic of the n sample for $n=k_2-1$ and $i=k_2-k_1$ in our case noting $\lim_{v\to\underline{v}} p_{(k_2,k_1)}(\widetilde{v}|v) = \lim_{v\to\underline{v}} g^{(k_2-k_1:k_2-1)}(\widetilde{v}|v) =$

[20] To be precise, $f(\widetilde{v}|v) = \frac{f(\widetilde{v})}{1-F(v)}$, $F(\widetilde{v}|v) = \frac{F(\widetilde{v})-F(v)}{1-F(v)}$, and $g^{(k_2-k_1:k_2-1)}(\widetilde{v}|v) = \frac{g^{(k_2-k_1:k_2-1)}(\widetilde{v})}{1-G^{(k_2-k_1:k_2-1)}(v)}$.

$g^{(k_2-k_1:k_2-1)}(\widetilde{v})$. Then, the identification of the distribution of valuations is straightforward by Theorem 1 in Athey and Haile (2002) saying that the parent distribution is identified whenever the distribution of any order statistic (here $k_2 - k_1$) with a known sample size (here $k_2 - 1$) is identified.

4.2 Auction Heterogeneity

In practice, the valuation of objects sold in WUCA (as in other auctions) varies according to several observed characteristics such as car types, makes, mileage, year, and etc. We want to control the effect of these observables on the individual valuation to obtain the homogeneity of the idiosyncratic factors such as tastes, cost shocks, or demand shocks. For this purpose, we assume the following nonparametric form of the valuation $\mathbb{V}_i(X_t)$ as

$$\ln \mathbb{V}_i(X_t) = W(l^*(X_t)) + V_{ti}, \tag{4}$$

where $W(\cdot)$ is a known link function, X_t is a vector of observable characteristics of the auction t, and i denotes the bidder i. We assume that V_{ti} is independent of X_t. Thus, we do assume the additively(or multiplicatively) separable structure of the value function, which is preserved by equilibrium bidding. In the auction we consider, ignoring the minimal increment, we have

$$\ln B(\mathbb{V}_{ti}, X_t) = W(l^*(X_t)) + b(V_{ti}),$$

where $\mathbb{V}_{ti}$ is the valuation of a bidder i on an auction t, $B(\mathbb{V}_{ti}, X_t)$ is a bidding function of a bidder i with observed heterogeneity of an auction t and $b(V_{ti})$ is a bidding function of homogeneous auctions. Under the IPV assumption, we have $B(\mathbb{V}, X) = \mathbb{V}(X)$ and $b(V) = V$ as before.
Here we impose parametric assumptions on the shape of $l(\cdot)$ but the distribution of V denoted by $F(\cdot)$ is nonparametrically specified. In what follows, we also assume $W(\cdot)$ is an identity function. Thus, we have

$$\ln \mathbb{V}_i(X_t) = X_t\beta + V_{ti}, \tag{5}$$

for a $\dim(X_k) \times 1$ parameter β.

4.3 Estimation

4.3.1 Distribution of Valuations

Here we use the semi-nonparametric (SNP) estimation procedure developed by Gallant and Nychka (1987) and Coppejans and Gallant (2002).

In particular, we implement a particular sieve estimation of the unknown density function using a Hermite series. First, we approximate the function space, $\mathcal{H}$, containing the true density function with a sieve space of the Hermite series, $\mathcal{H}_T$. Once we set up the objective function based on a Hermite series approximation of the unknown density function, then the estimation procedure is just a finite dimensional parametric problem. In particular, we use the maximum likelihood methods. What remains is to specify the particular rate in which a sieve space, $\mathcal{H}_T$, gets closer to $\mathcal{H}$ achieving the consistency of the estimator. We will specify several regularity conditions for this in the Technical Appendix.

Since we observe at least the second-, third- and fourth-highest bids in each auction of WUCA. We can estimate several different versions of the distributions of valuations ($F(\cdot)$), since any pair of order statistics can identify the parent distribution according to Song (2005) under the number of potential bidders unknown or unobserved. Here, we use two pairs of order statistics (second-, fourth-) and (third-, fourth-) highest bids and obtain two different values of $F(\cdot)$, which provides us an opportunity to test the hypothesis that WUCA is the IPV. This testable implication comes from the fact that under the IPV, value of $F(v)$ implied by the distributions of different order statistics must be identical for all v.[21]

Once we show that the IPV assumption holds, then we can combine several order statistics to identify $F(\cdot)$ extending Song (2005) to the case of more than three bids observed. This version of estimator is better than the version that use a pair of order statistics in the sense that we are using more information. In Appendix, we establish the consistency and the convergence rate of this SNP estimator.

4.3.2 Simple Two Step Estimation

Though the estimation procedure considered up to now is a feasible one-step method and may be more efficient, we rather use a two-step estimation method as follows so that we can avoid the computational burden involved in estimating $l^*(\cdot)$ and $F(\cdot)$ at the same time[22]. First, we estimate the following first-stage regression.

$$\ln \mathbb{V}_{tj} = X_t{}'\beta + \epsilon_{tj},$$

Then, construct the residuals for each order statistics as

$$\widehat{v}_{tj}(k_1) = \ln \mathbb{V}_{tj} - X_t'\widehat{\beta}.$$

In the second step, based on the estimated pseudo values $\hat{v}_{tj}$, we estimate $f(\cdot)$ based on (23). To implement the estimation we proposed here, we need to choose the optimal length of series denoted by K^* using a cross-validation strategy. In particular, we follow the Coppejans and Gallant (2002)'s method. For detailed discussion, see Appendix B.1.

[21] For detailed discussion, see Athey and Haile (2002).

[22] We may also consider nonparametric estimation of $l^*(\cdot)$. In this case, a two step procedure will be very relevant.

5 Nonparametric Testing

As noted in the previous section, we can test the IPV assumption, since several versions of distribution of valuations are identified under availability of three order statistics. In particular, one is from the pairs of the second- and fourth- highest bids and the second one is from the pairs the third- and fourth- order statistics. Therefore, by comparing $\hat{f}_1(\cdot)$ and $\hat{f}_2(\cdot)$, we can test the following hypothesis H_0

$$\begin{aligned} H_0 &: \text{ WUCA is an IPV auction} \\ H_A &: \text{ WUCA is not an IPV auction} \end{aligned} \tag{6}$$

since under H_0, there should be no significant difference between $\hat{f}_1(\cdot)$ and $\hat{f}_2(\cdot)$. Denote f_{01} to be the true density of the parent distribution that generates the second- and fourth- highest order statistics and f_{02} to be the true density of the parent distribution for the third- and fourth- order statistics. Then, formally we can test (6) by comparing the two density functions $f_{01}(\cdot)$ and $f_{02}(\cdot)$. Under the null we have

$$H_0 : f_{01}(\cdot) = f_{02}(\cdot) \tag{7}$$

against $H_A : f_{01}(\cdot) \neq f_{02}(\cdot)$.

5.1 Tests based on Means or Higher Moments

We can test (6) based on the means or higher moments implied by f_1 and f_2 as

$$\begin{aligned} H_0^j\ (IPV) &: \ \mu_1^j = \mu_2^j \\ H_A^j\ (NIPV) &: \ \mu_1^j \neq \mu_2^j, \quad j = 1, 2, \ldots, J \end{aligned} \tag{8}$$

where $\mu_k^j = \int_c^\infty v^j f_k(v)dv$, $k = 1, 2$, since (6) implies (8) and (8) implies (6) as $J \to \infty$. We can compare several estimates of moments implied by $\hat{f}_1$ and $\hat{f}_2$ and test the significance difference of each pair by constructing a standardized test statistics. One difficulty is to reflect the fact that we used pre-estimated functions in obtaining $\hat{f}_1$ and $\hat{f}_2$ in calculating the asymptotic variance of each moment estimate.

5.2 Comparison of Densities Using the Pseudo Kullback-Leibler Divergence Measure

Here we are interested in testing the equivalence of two densities f_{01} and f_{02} where these densities are estimated from (29) and (31), respectively using the SNP estimation. A natural measure to compare f_{01} and f_{02} will be the integrated squared error given by $I_s(f_{01}(z), f_{02}(z)) = \int_{\mathcal{V}} (f_{01}(z) - f_{02}(z))^2 dz$

noting f_{01} and f_{02} have the same support $\mathcal{V}$. Under the null we have $I_s(f_{01}(z),f_{02}(z)) = 0$. Li (1996) develops a test statistic of this sort when both densities are estimated using a kernel method. Other possible measures for the distance of two density functions are the Kullback-Leibler (KL) information distance or the Hellinger metric. For testing of serial independence, it is well known that test statistics based on these two measures have better small sample properties than those based on a squared distance in terms of a second order asymptotic theory. The KL measure is entertained in Ullah and Singh (1989), Robinson (1991), and Hong and White (2005) when they effectively test the affinity of two densities. The KL measure is defined as $I_{KL} = \int_{\mathcal{V}} (\ln f_{01}(z) - \ln f_{02}(z))\, f_{01}(z)dz$ or $I_{KL} = \int_{\mathcal{V}} (\ln f_{02}(z) - \ln f_{01}(z))\, f_{02}(z)dz$ which are equally zero under the null and have positive values under the alternative as shown in Kullback and Leibler (1951). However, it is noted that the KL information distance is not a proper distance measure, since it is not symmetric although it still serves as a valid discrepancy measure. Kim (2005) proposes a variation of the Kullback-Leibler measure which is symmetric and nonnegative as

$$I(f_{01}, f_{02}) = \int_{\mathcal{V}} (\ln f_{01}(z) - \ln f_{02}(z))\, f_{01}(z)dz + \int_{\mathcal{V}} (\ln f_{02}(z) - \ln f_{01}(z))\, f_{02}\,(z)\,dz \tag{9}$$

which has zero value under the null but is strictly positive under the alternative by construction. It is also symmetric, $I(f_{01}, f_{02}) = I(f_{02}, f_{01})$.

We could construct a test statistic as a sample analogue of (9):

$$\begin{aligned} I(\widehat{f}_1, \widehat{f}_2) &= \int_{\mathcal{V}} (\ln \widehat{f}_1(z) - \ln \widehat{f}_2(z))\widehat{f}_1(z)dz + \int_{\mathcal{V}} (\ln \widehat{f}_2(z) - \ln \widehat{f}_1(z))\widehat{f}_2(z)dz \\ &= \int_{\mathcal{V}} (\ln \widehat{f}_1(z) - \ln \widehat{f}_2(z))d\widehat{F}_1(z) + \int_{\mathcal{V}} (\ln \widehat{f}_2(z) - \ln \widehat{f}_1(z))d\widehat{F}_2(z) \end{aligned}$$

where $d\widehat{F}_1(z) = \widehat{f}_1(z)dz$ and $d\widehat{F}_2(z) = \widehat{f}_2(z)dz$. Now suppose $\{\widetilde{v}_t\}_{t=1}^T$ are the second-highest order statistics. Then, one may expect that for example, $\int_{\mathcal{V}} (\ln \widehat{f}_1(z) - \ln \widehat{f}_2(z))d\widehat{F}_{o1}(z) \approx \frac{1}{T}\sum_{t=1}^T \left(\ln \widehat{f}_1(\widetilde{v}_t) - \ln \widehat{f}_2(\widetilde{v}_t)\right)$ but this cannot be true since v_t follows the distribution of the second-highest order statistic not of F_{o1}. We, however, argue that

$$\frac{1}{T}\sum_{t=1}^T \left(\ln \widehat{f}_1(\widetilde{v}_t) - \ln \widehat{f}_2(\widetilde{v}_t)\right) \tag{10}$$

is still valid in terms of comparing two densities since under the null, the following object still equals to zero

$$\int_{\mathcal{V}} (\ln f_{o1}(z) - \ln f_{o2}(z))\, g^{(n-1:n)}(z)dz \tag{11}$$

where $g^{(n-1:n)}(z)$ is the density of the second-highest order statistic among T. This is true for any given order statistics. Based on (10), we develop a test statistic that tests (7) in Appendix E. For more details on the proposed test statistic, see Appendix E.

5.3 Combining Several Order Statistics

Once we show the several versions of estimates for the distribution of valuations are statistically not different each other, we may obtain a better estimate by combining these. One way to do this is to consider the joint density function of two or more order statistics conditional on a certain order statistic. Assume that we have the k_1^{th}, k_2^{th}, and k_3^{th}-highest order statistics, which are the $(n-k_1-1)^{th}$, $(n-k_1-1)^{th}$, $(n-k_1-1)^{th}$ order statistics respectively $(1 \le k_1 < k_2 < k_3 \le n)$. Denote the joint density of these three order statistics as $\tilde{g}^{(k_1,k_2,k_3:n)}(\cdot)$[23]

$$\begin{aligned}\tilde{g}^{(k_1,k_2,k_3:n)}(\widetilde{v},\widetilde{\widetilde{v}},v) &= \frac{n!}{(n-k_3)!(k_3-k_2-1)!(k_2-k_1-1)!(k_1-1)!}\\ &\times F(v)^{n-k_3}f(v)[F(\widetilde{\widetilde{v}})-F(v)]^{k_3-k_2-1}f(\widetilde{\widetilde{v}})[F(\widetilde{v})-F(\widetilde{\widetilde{v}})]^{k_2-k_1-1}f(\widetilde{v})[1-F(\widetilde{v})]^{k_1-1},\end{aligned}$$

where $\widetilde{V}$ denotes the k_1^{th}-, $\widetilde{\widetilde{V}}$ denotes k_2^{th}- and V denotes k_3^{th}-highest order statistics. Using this joint density function with (1), we obtain the conditional joint density of the k_1^{th} and the k_2^{th}-highest order statistics conditional on the k_3^{th}-highest statistics as

$$\begin{aligned}p_{(k_1,k_2)|k_3}(\widetilde{v},\widetilde{\widetilde{v}}|v) &= \frac{(k_3-1)!}{(k_3-k_2-1)!(k_2-k_1-1)!(k_1-1)!}\\ &\times\frac{[F(\widetilde{\widetilde{v}})-F(v)]^{k_3-k_2-1}f(\widetilde{\widetilde{v}})[F(\widetilde{v})-F(\widetilde{\widetilde{v}})]^{k_2-k_1-1}f(\widetilde{v})[1-F(\widetilde{v})]^{k_1-1}}{[1-F(v)]^{k_3-1}}\\ &= \frac{(k_3-1)!}{(k_3-k_2-1)!(k_2-k_1-1)!(k_1-1)!}[(1-F(v))F(\widetilde{\widetilde{v}}|v)]^{k_3-k_2-1}\\ &\times f(\widetilde{\widetilde{v}}|v)(1-F(v))[(F(\widetilde{v}|v)-F(\widetilde{\widetilde{v}}|v))(1-F(v))]^{k_2-k_1-1}\\ &\times\frac{f(\widetilde{v}|v)(1-F(v))[(1-F(\widetilde{v}|v))(1-F(v))]^{k_1-1}}{[1-F(v)]^{k_3-1}}\\ &= \frac{(k_3-1)!}{(k_3-k_2-1)!(k_2-k_1-1)!(k_1-1)!}F(\widetilde{\widetilde{v}}|v)^{k_3-k_2-1}f(\widetilde{\widetilde{v}}|v)\\ &\times[F(\widetilde{v}|v)-F(\widetilde{\widetilde{v}}|v)]^{(k_3-k_1)-(k_3-k_2)-1}f(\widetilde{v}|v)[1-F(\widetilde{v}|v)]^{(k_3-1)-(k_3-k_1)}\\ &= g^{(k_3-k_1,k_3-k_2:k_3-1)}(\widetilde{v},\widetilde{\widetilde{v}}|v),\end{aligned}\tag{12}$$

where $F(\cdot|v)$ and $f(\cdot|v)(g(\cdot|v))$ are the truncated CDF and PDF truncated at x respectively. The last equality comes from the joint density of the j-th and i-th order statistics $(n \ge j > i \ge 1)$,

$$g^{(j,i:n)}(a,b) = \frac{n![F(b)]^{i-1}[F(a)-F(b)]^{j-i-1}[1-F(a)]^{n-j}f(b)f(a)}{(i-1)!(j-i-1)!(n-j)!}I_{\{a>b\}}$$

where a and b are the j-th and i-th order statistics, respectively by matching $j=k_3-k_1$, $i=k_3-k_2$, and $n=k_3-1$. Therefore we can interpret $p_{(k_1,k_2|k_3)}(\cdot)$ as the joint density of $(k_3-k_1)^{th}$ and

[23] We use this notation $\tilde{g}^{(\cdot)}$ to distinguish it from $g^{(\cdot)}$ so that k_i denotes the k_i^{th} highest order statistics

$(k_3 - k_2)^{th}$ order statistics from a sample of size equals to $(k_3 - 1)$. When $(k_1, k_2, k_3) = (2, 3, 4)$, (12) becomes

$$p_{(2,3)|4}(\widetilde{v}, \widetilde{\widetilde{v}}|v) = \frac{6f(\widetilde{\widetilde{v}})f(\widetilde{v})[1 - F(\widetilde{v})]}{[1 - F(v)]^3} \tag{13}$$

Based on (13), we can estimate the distribution of valuations, $f_0(\cdot)$, similarly with the method proposed in Appendix D. The resulting estimator is more efficient than $\hat{f}_1(\cdot)$ or $\hat{f}_2(\cdot)$ in the sense that it uses more information than the others.

6 Empirical Results

6.1 Benchmark Monte Carlo

In this section, we perform several Monte Carlo experiments to illustrate the validity of our estimation strategy. First, we generate artificial data of $T = 1000$ auctions as follows. The number of potential bidders, $\{N_i\}$, are drawn from a Binomial distribution with $(n, p) = (50, 0.1)$ for each auction $(i = 1, \ldots, T)$. N_i potential bidders are assumed to value the object according to:

$$\ln V_{ij} = \alpha_1 X_{1i} + \alpha_2 X_{2i} + \alpha_3 X_{3i} + v_{ij}, \tag{14}$$

where $\alpha_1 = 1$, $\alpha_2 = -1$, $\alpha_3 = 0.5$, $X_{1i} \sim N(0, 1)$, $X_{2i} \sim Exp(1)$, $X_{3i} = X_{1i} \cdot X_{2i} + 1$, and $v_{ij} \sim Gamma(9, 3)$[24]. $X_{\cdot i}$'s represent the observed auction heterogeneity and v_{ij} is bidder j's private information in auction i, whose distribution is our primary interests here. To consider the case of biding reserve prices, we also generate the reserve prices equation as

$$\ln R_i = \alpha_1 X_{1i} + \alpha_2 X_{2i} + \alpha_3 X_{3i} + \eta_i,$$

where $\eta_i \sim Gamma(9, 3) - 2$. Note that by construction V_{ij} and R_i are independent conditional on $X_{\cdot i}$'s. Artificial actual bidders bid only when those V_{ij} are greater than R_i. Here we assume our imaginary researcher do not know the presence of potential bidders with valuations below R_i. Thus, in each experiment, she has a data set of $X_{\cdot i}$'s, and the second-, the third-, and the fourth-highest among actual bidder's bids. Auctions with fewer than four actual bidders are dropped. Hence, our research has the sample size less than $T = 1000$ on average $T = 680$ with 50 repetitions. Our researcher estimates α_1, α_2, α_3 and $f_v(\cdot)$ by varying the smoothness (K) of the SNP estimator, from 0 to 7 without knowing the specification of the distribution of v_{ij} in (14).[25,26]

[24] Note that for $X \sim Gamma(9, 3)$, $E(X) = 3$ and $Var(X) = 1$.

[25] For the actual Monte Carlo experiments in this paper, we used the simple two-stage estimation with linear model because of a computational burden.

[26] Among K between 0 to 7, we choose $K = 6$ because it performs best in this Monte Carlo experiments.

Figure 2 illustrates a sample performance of the estimator considered here. By construction of the data generation, the three versions of estimates for the density function of valuations should be almost identical (one is based on $(2^{nd}, 4^{th})$ order statistics pair, the others are on $(3^{rd}, 4^{th})$ and $(2^{nd}, 3^{rd})$.)

6.2 Estimation Results

In the first stage regression obtaining the approximated function of the observed heterogeneity part, $\hat{l}(x)$. We used a linear specification for reducing the computational burden of our SNP estimation.

We consider the following covariates: X_1 is the vector of dummy variables indicating the make such as Hyundai, Daewoo, Kia or Others; X_2 is the age; X_3 is the mileage; X_4 is the engine size; X_6 is the remaining time of the current title; X_7 is the dummy variable for the transmission type; X_8 is the dummy variable for the fuel type; X_9 is the dummy variable for the colors and X_{10} is the dummy variable for the options. Table 3 provides the first-stage estimation results with the linear specification. The signs for the coefficient of age, transmission, engine size, and rating variables look reasonable, while the signs for the fuel type, mileage, colors, options, title remaining are not so clear.

We estimate $l^*(x)$ separately for each car make and obtain the estimate of pseudo valuations as residuals imposing the restriction $l^*(0) = 0$ for identification. Thus, we actually use the basis functional form $l_H(X_2, X_3, ..., X_{10})$ for Hyundai and use $c_m + l_m(X_2, X_3, ..., X_{10})$ for others, $m \in \{\text{Daewoo, Kia and Others}\}$.

Based on the estimated pseudo valuations, in the second step, we estimate the distribution of valuations using three different pairs of order statistics $(2^{nd}, 4^{th})$, $(3^{rd}, 4^{th})$, and $(2^{nd}, 3^{rd})$. Figure 3 illustrates the estimated density function of valuations using the linear estimation in the first stage.[27]

With these three nonparametric estimates of the distribution, we conduct our nonparametric test of IPV. Our test statistic indicates that the null hypothesis of IPV is not rejected with 1% significance level when we compare all three possible combinations of the densities that are estimated from three different pairs of order statistics. Values for the test statistic is provided in Table 4.

[27] Following the cross-validation strategy explained in Appendix B, we can pick the optimal lengths of series k_1^* for the first stage regression and K^* for the SNP estimator.

7 Implications and Discussions

7.1 Optimal Reserve Price

The key policy issue for the seller is the reserve price. The seller wants to maximize the expected profit by setting a minimum acceptance price so that only bidders have higher valuations than the reserve price attend the auction. The optimum depends on the distribution of valuations, which is our primary interests and derived in previous sections. We are willing to assume the following, which is implied by the standard regularity condition of Myerson (1981) and assumed in Haile and Tamer (2003).

Assumption 7.1 $(p - vc_0)[1 - F_v(p)]$ *is strictly pseudo-concave in* p *on* $(\underline{v}, \bar{v})$*,*

where vc_0 is the cost associated with the auction. The pseudo-optimal reserve price (without the observed heterogeneity) is characterized by

$$p^* = \arg\max_p (p - vc_0)[1 - F_v(p)],$$

which becomes, under Assumption 7.1

$$p^* = vc_0 + \frac{1 - F(p^*)}{f(p^*)}$$

One nice feature of the additively separability assumed in (4) is that the equilibrium bidding is preserved under the observed heterogeneity as $B(V(x)) = l^*(x) + B(v)$, where $B(v)$ is the bidding function under being absence of the observed auction heterogeneity. Thus, the optimal reserve price also has the simple additive form of the observed heterogeneity part and the pseudo-optimal reserve price:

$$p^*(X) = l^*(X) + p^* = l^*(X) + vc_0 + \frac{1 - F_v(p^*)}{f_v(p^*)}$$

Thus, we can estimate $p^*(x)$ using the previous estimates of $\hat{l}(\cdot)$, $\hat{f}_v(\cdot)$ and $\hat{F}_v(\cdot)$ as

$$\hat{p}(x) = \hat{l}(x) + \hat{p},$$

where $\hat{p}$ solves $p = \hat{vc} + \frac{1-\hat{F}_v(p)}{\hat{f}_v(p)}$ and $\hat{vc}$ is a consistent estimator of vc_0. It will be very interesting to compare these implied optimal reserve prices from the distribution of valuations and the actual reserve prices recorded in each auction of WUCA, since the actual reserve price data is readily available in our data set. If significant difference emerges between these two and a particular pattern is found in there difference , then it may shed lights on the seller's strategic behavior in WUCA, if any.

7.2 Bounds Estimation

Until now, we have disregarded the minimum increment of around 30 dollars in WUCA. In this section, we discuss how to obtain the bounds of the distribution of valuations incorporating the fact that there exists the minimum increment in WUCA. The bounds considered here is much simpler than those considered in Haile and Tamer (2003), since in WUCA, by construction, a high order statistic of valuations other than the first highest one is bounded as

$$b_{(i:n)} \leq v_{(i:n)} \leq b_{(i:n)} + \Delta, \text{ for all } i = 1, \ldots, n-1, \tag{15}$$

where $(i:n)$ denotes the i^{th} order statistic out of the n sample. By the first-order stochastic dominance, noting $G_{b_{(i:n)+\Delta}}(v) = G_{b_{(i:n)}}(v - \Delta)$, (15) implies

$$G_{b_{(i:n)}}(v) \geq G_{v_{(i:n)}}(v) \geq G_{b_{(i:n)}}(v - \Delta),$$

, where $G.(\cdot)$ is the distribution of the order statistics. Then, using the identification method discussed in previous sections, we have

$$F_b(v) \geq F_v(v) \geq F_b(v - \Delta) \cong F_b(v) - f_b(v)\Delta,$$

where $F_b(\cdot)$, $f_b(\cdot)$ are the CDF and PDF, respectively, of valuations based on bids and the last weak equality comes from the first-order Taylor series expansion. Therefore, we can estimate the bounds of $F_v(v)$ as

$$\hat{F}_b(v) \geq F_v(v) \geq \hat{F}_b(v) - \hat{f}_b(v)\Delta,$$

where $\hat{f}_b(\cdot)$ the SNP estimator based on the certain observed order statistics of bids, $\hat{F}_b(x) = \int_{\min(b)}^{x} \hat{f}_b(v)dv$ and $\min(b)$ is the minimum among the observed bids considered.

7.3 Discussions

In this section, we discuss our IPV test. With our estimates, we do not reject the null hypothesis of IPV in our auction. We can interpret this result as an evidence of no informational dependency among bidders about the valuations of observed characteristics of a used-car and no effect of any unobserved (to an econometrician) characteristics on valuations.

Our conjecture is that this situation may arise because each dealer operates in her own local market and there is no interdependency among those markets, or when a dealer has a specific demand (order) from a final buyer on hand.[28]

[28] This can happen when a dealer gets an order for a specific used-car but does not have one in stock or when a consumer looks up the item list of an auction and asks a dealer to buy a specific car for her.

If the assumption of IPV were rejected, it could have been due to a violation of the independence assumption or a violation of the private value assumption. This could happen when a dealer does not have a specific demand on hand, but she anticipates some demand in near future from the analysis of the overall performance of the national market since every used-car won in the auction is to be resold to a final consumer. In this case, a dealer may have some incentives to find out other dealers' opinions about the prospect of the national market. With our data, we find no evidence to support this hypothesis.

8 Conclusions

In this paper, we conducted a structural analysis of ascending-price auctions using a new data set on a wholesale used-car auction, in which there is no jump bidding. Exploiting the data, we estimated the distribution of bidders' valuations nonparametrically within the IPV paradigm.

Using the estimates of the distribution, we developed a new nonparametric test of IPV for the case where the number of potential bidders is unknown by extending the work of Athey and Haile (2002). We could implement our test because the data enabled us to exploit information from observed losing bids. For our identification and estimation, we utilized and extended the work of Song (2005) for identifying and estimating the distribution of valuations when the number of potential bidders of ascending auctions is unknown in an IPV setting.

We find that the null hypothesis of IPV is not rejected with our data after controlling for observed auction heterogeneity, and therefore our estimation result remains a valid approximation of the distribution of dealers' valuations.

The richness of our data has allowed us to conduct a structural analysis that bridges the gap between theoretical models based on Milgrom and Weber (1982) and real-world ascending-price auctions.

In this paper, we have considered the auction as a collection of isolated single-object auctions. In future work, we will look at the data more closely in the alternative environments. For example, we will examine intra-day dynamics of auctions with daily budget constraints for bidders, or possible complementarity and substitutability in a multi-object auction environment.

References

[1] An, M. Y. and Liu, M. (2000), "Using Indirect Inference to Solve the Initial-Conditions Problem," *Review of Economics and Statistics* 82, 656-667.

[2] Arnold, B., Balakrishnan, N., and Nagaraja, H. (1992), *A First Course in Order Statistics.* (New York: John Wiley & Sons.)

[3] Athey, S. and Haile, P. (2002), "Identification of Standard Auction Models," *Econometrica*, 70(6), 2107-2140.

[4] Athey, S. and Haile, P. (2005), "Nonparametric Approaches to Auctions," forthcoming in J. Heckman and E. Leamer, eds., Handbook of Econometrics, Vol. 6, Elsevier.

[5] Athey, S. and Levin, J. (2001), "Information and Competition in U.S. Forest Service Timber Auctions," *Journal of Political Economy*, 109, 375-417.

[6] Bikhchandani, S., Haile, P. and Riley, J. (2002), "Symmetric Separating Equilibria in English Auctions," *Games and Economic Behavior*, 38, 19-27.

[7] Bikhchandani, S. and Riley, J. (1991), "Equilibria in Open Common Value Auctions," *Journal of Economic Theory*, 53, 101-130.

[8] Bikhchandani, S. and Riley, J. (1993), "Equilibria in Open Auctions," mimeo, UCLA.

[9] Campo, S., Perrigne, I., and Vuong, Q. (2003) "Asymmetry in First-Price Auctions With Affiliated Private Values," *Journal of Applied Econometrics*, 18, 179-207.

[10] Coppejans, M. and Gallant, A. R. (2002), "Cross-validated SNP density estimates," *Journal of Econometrics*, 110, 27-65.

[11] David, H. A. (1981), *Order Statistics.* Second Edition. (New York: John Wiley & Sons.)

[12] Dridi, R., Guay, A., and Renault E. (2003), "Indirect Inference and Calibrations of Dynamic Stochastic General Equilibrium Models," working paper, Université di Montréal.

[13] Fenton V. M. and Gallant, A. R. (1996a), "Convergence Rates of SNP Density Estimators," *Econometrica*, 64, 719-727.

[14] Fenton V. M. and Gallant, A. R. (1996b), "Qualitative and Asymptotic Performance of SNP Density Estimators," *Journal of Econometrics*, 74, 77-118.

[15] Gallant, A. R. and Nychka, D. (1987), "Semi-nonparametric Maximum Likelihood Estimation," *Econometrica*, 55, 363-390.

[16] Gallant, A. R. and G. Souza (1991), "On the Asymptotic Normality of Fourier Flexible Form Estimates," *Journal of Econometrics*, 50, 329-353.

[17] Gourieroux, C., Monfort, A. and Renault, E. (1993), "Indirect Inference," *Journal of Applied Econometrics*, 8, 85-118.

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

[19] Haile, P., Hong, H., and Shum, M. (2003), "Nonparametric Tests for Common Values in First-Price Sealed-Bid Auctions," NBER Working Paper No. 10105.

[20] Haile, P. and Tamer, E. (2003), "Inference with an Incomplete Model of English Auctions," *Journal of Political Economy*, 111(1), 1-51.

[21] Hansen, B. (2004), "Nonparametric Conditional Density Estimation," Working paper, University of Wisconsin-Madison.

[22] Harstad, R. and Rothkopf, M. (2000), "An "Alternating Recognition" Model of English Auctions," *Management Science*, 46, 1-12.

[23] Hendricks, K. and Paarsch, H. (1995), "A Survey of Recent Empirical Work concerning Auctions," *Canadian Journal of Economics*, 28, 403-426.

[24] Hendricks, K., Pinkse, J., and Porter, R. (2003), "Empirical Implications of Equilibrium Bidding in First-Price, Symmetric, Common Value Auctions," *Review of Economic Studies*, 70(1), 115-146.

[25] Hoeffding, W. (1963), "Probability Inequalities for Sums of Bounded Random Variables," *Journal of the American Statistical Association*, 58, 13-30.

[26] Hong, H., and Shum, M. (2003), "Econometric Models of Asymmetric Ascending Auctions," *Journal of Econometrics*, 112, 327-358.

[27] Hong, Y.-M. and White, H. (2005), "Asymptotic Distribution Theory for an Entropy-Based Measure of Serial Dependence," *Econometrica*, 73, 837-902.

[28] Izmalkov, S. (2003), "English Auctions with Reentry," mimeo, MIT.

[29] Krasnokutskaya, E. (2004), "Identification and Estimation in Highway Procurement Auctions Under Unobserved Auction Heterogeneity," mimeo, University of Pennsylvania.

[30] Keane, M. and A. Smith, Jr. (2003), "Generalized Indirect Inference for Discrete Choice Models," working paper, Department of Economics, Yale University.

[31] Kim, K. (2005), "Uniform Convergence Rate of the SNP Density Estimator and Testing for Similarity of Two Unknown Densities," mimeo, UCLA.

[32] Kullback, S. and Leibler, R.A. (1951), "On Information and Sufficiency," *Ann. Math. Stat.*, 22, 79–86.

[33] Laffont, J.-J. (1997), "Game Theory and Empirical Economics: The Case of Auction Data," *European Economic Review*, 41, 1-35.

[34] Li, Q. (1996), "Nonparametric Testing of Closeness between Two Unknown Distribution Functions," *Econometric Reviews*, 15, 261-274.

[35] Li, T., Perrigne, I., and Vuong, Q. (2000), "Conditionally Independent Private Information in OCS Wildcat Auctions," *Journal of Econometrics*, 98, 129-161.

[36] Li, T., Perrigne, I., and Vuong, Q. (2002), "Structural Estimation of the Affiliated Private Value Auction Model," *RAND Journal of Economics*, 33, 171-193.

[37] Lorentz, G. (1986), *Approximation of Functions*, New York: Chelsea Publishing Company.

[38] Milgrom, P. and Weber, R. (1982), "A Theory of Auctions and Competitive Bidding," *Econometrica*, 50, 1089-1122.

[39] Myerson, R. (1981), "Optimal Auction Design," *Math. Operations Res.*, 6, 58-73.

[40] Newey, W. (1997), "Convergence Rates and Asymptotic Normality for Series Estimators," *Journal of Econometrics*, 79, 147-168.

[41] Newey, W., J. Powell and F. Vella (1999), "Nonparametric Estimation of Triangular Simultaneous Equations Models," *Econometrica*, 67, 565-603.

[42] Paarsch, H. (1992), "Deciding Between the Common and Private Value Paradigms in Empirical Models of Auctions," *Journal of Econometrics*, 51, 191-215.

[43] Riley, J. (1988), "Ex Post Information in Auctions," *Review of Economic Studies*, 55, 409-429.

[44] Robinson, P. M. (1991), "Consistent Nonparametric Entropy-Based Testing," *Review of Economic Studies*, 58, 437-453.

[45] Song, U. (2005), "Nonparametric Estimation of an eBay Auction Model with an Unknown Number of Bidders," Economics Department Discussion Paper 05-14, University of British Columbia.

[46] Ullah, A. and Singh, R.S. (1989), "Estimation of a Probability Density Function with Applications to Nonparametric Inference in Econometrics," in B. Raj (ed.) *Advances in Econometrics and Modelling,* Kluwer Academic Press, Holland.

Appendices

A Normalized Hermite Polynomials

Here are examples of the normalized Hermite polynomials that we use for approximating the density of valuations. For $j = 0, \ldots, 9$

$$
\begin{aligned}
H_1 &= \frac{1}{\sqrt[4]{2\pi}} e^{-\frac{1}{4}x^2} \\
H_2 &= \frac{x}{\sqrt[4]{2\pi}} e^{-\frac{1}{4}x^2} \\
H_3 &= \frac{1}{2\sqrt[4]{2\pi}} \left(x^2\sqrt{2} - \sqrt{2}\right) e^{-\frac{1}{4}x^2} \\
H_4 &= \frac{1}{6\sqrt[4]{2\pi}} \left(x^3\sqrt{6} - 3x\sqrt{6}\right) e^{-\frac{1}{4}x^2} \\
H_5 &= \frac{1}{12\sqrt[4]{2\pi}} \left(3\sqrt{6} - 6x^2\sqrt{6} + x^4\sqrt{6}\right) e^{-\frac{1}{4}x^2} \\
H_6 &= \frac{1}{60\sqrt[4]{2\pi}} \left(15x\sqrt{30} - 10x^3\sqrt{30} + x^5\sqrt{30}\right) e^{-\frac{1}{4}x^2} \\
H_7 &= \frac{1}{60\sqrt[4]{2\pi}} \left(45x^2\sqrt{5} - 15\sqrt{5} - 15x^4\sqrt{5} + x^6\sqrt{5}\right) e^{-\frac{1}{4}x^2} \\
H_8 &= \frac{1}{420\sqrt[4]{2\pi}} \left(105x^3\sqrt{35} - 105x\sqrt{35} - 21x^5\sqrt{35} + x^7\sqrt{35}\right) e^{-\frac{1}{4}x^2} \\
H_9 &= \frac{1}{1680\sqrt[4]{2\pi}} \left(105\sqrt{70} - 420x^2\sqrt{70} + 210x^4\sqrt{70} - 28x^6\sqrt{70} + x^8\sqrt{70}\right) e^{-\frac{1}{4}x^2} \\
H_{10} &= \frac{1}{5040\sqrt[4]{2\pi}} \left(945x\sqrt{70} - 1260x^3\sqrt{70} + 378x^5\sqrt{70} - 36x^7\sqrt{70} + x^9\sqrt{70}\right) e^{-\frac{1}{4}x^2}
\end{aligned}
$$

Note that $\int_{-\infty}^{\infty} H_j^2 dx = 1$ and $\int_{-\infty}^{\infty} H_j H_k dx = 0$, $j \neq k$.

B Choosing the optimal smoothing parameters

Instead of using the Leave-one-out method, we will partition the data into P groups, making the size of each group as equal as possible and use the *Leave-one partition-out* method. This is because it will be computationally too expensive to use the Leave-one-out method, since the data size is so large. We let T_p denote the set of the data indices that belongs to the p^{th} group such that $T_p \cap T_{p'} = \phi$ for $p \neq p'$ and $\cup_{p=1}^{P} T_p = \{1, 2, \ldots, T\}$.

B.1 Choosing K^*

Coppejans and Gallant (2002) employ a cross-validation method based on the ISE (Integrated Squared Error) criteria. The ISE is defined for $\hat{h}(x)$, a density estimate of $h(x)$

$$\begin{aligned} \text{ISE}(\hat{h}) &= \int \hat{h}^2(x)dx - 2\int \hat{h}(x)h(x)dx + \int h(x)^2 dx \\ &= M_{(1)} - 2M_{(2)} + M_3. \end{aligned}$$

To approximate the ISE in terms of $p^*(y|x)$, we again use the cross-validation strategy with the data partitioned into P groups. We first approximate $M_{(1)}$ with

$$\begin{aligned} \widehat{M}_{(1)}(K) &= \int (\hat{p^*}_K(y|x))^2 \\ &= \int \left(\frac{6(\hat{F}_K(y) - \hat{F}_K(x))(1-\hat{F}_K(y))\hat{f}_K(y)}{(1-\hat{F}_K(x))^3} \right)^2, \end{aligned}$$

where $\hat{f}_K(\cdot)$ denotes the SNP estimate with the length of the series equal to K and $\hat{F}_K(z) = \int_c^z \hat{f}_K(t)dt$. For $M_{(2)}$, we consider

$$\begin{aligned} \widehat{M}_{(2)}(K) &= \frac{1}{T}\sum_{p=1}^{P}\sum_{t\in T_p} \hat{p^*}_{p,K}(y_t|x_t) \\ &= \frac{1}{T}\sum_{p=1}^{P}\sum_{t\in T_p} \frac{6(\hat{F}_{p,K}(y_t) - \hat{F}_{p,K}(x_t))(1-\hat{F}_{p,K}(y_t))\hat{f}_{p,K}(y_t)}{(1-\hat{F}_{p,K}(x_t))^3}, \end{aligned}$$

where $\hat{f}_{p,K}(\cdot)$ denotes the SNP estimate obtained from the sample excluding p^{th} group with the length of the series, K and $\hat{F}_{p,K}(z) = \int_c^z \hat{f}_{p,K}(t)dt$. Noting $M_{(3)}$ is not a function of K, we pick K^* such that

$$K^* = \text{argmin}_K\, CVK(K) = \widehat{M}_{(1)}(K) - 2\widehat{M}_{(2)}(K)$$

C Nonparametric Extension: Control on the Observed Heterogeneity

Until now, we assume that $l^*(\cdot)$ belongs to a parametric family. Here we consider a nonparametric specification of $l^*(\cdot)$. Assuming the independence of the idiosyncratic factor, v_i, on the observables, X_i, we can approximate the unknown function $l(X_i)$ in (5) using a sieve estimation such as power series or splines. We first approximate the function space $\mathcal{L}$ containing $l^*(X_i)$ with the following power series sieve space $\mathcal{L}_T$

$$\mathcal{L}_T = \{l(X)|l(X) = R^{k_1(T)}(X)'\pi \text{ for all } \pi \text{ satisfying } \|l\|_{\Lambda^{\gamma_1}} \leq c_1\},$$

where $R^{k_1}(X)$ is a triangular array of some basis polynomials with the length of k_1. Here $\|\cdot\|_{\Lambda^\gamma}$ denotes the Hölder norm:

$$||g||_{\Lambda^\gamma} = \sup_x |g(x)| + \max_{a_1+a_2+\ldots a_{d_x}=\gamma} \sup_{x\neq x'} \frac{|\nabla^a g(x) - \nabla^a g(x')|}{(||x-x'||_E)^{\gamma-\underline{\gamma}}} < \infty,$$

where $\nabla^a g(x) \equiv \frac{\partial^{a_1+a_2+\ldots a_{d_x}}}{\partial x_1^{a_1}\ldots\partial x_{d_x}^{a_{d_x}}} g(x)$ with $\underline{\gamma}$ the largest integer smaller than γ. The Hölder ball (with radius c) $\Lambda_c^\gamma(\mathcal{X})$ is defined accordingly as

$$\Lambda_c^\gamma(\mathcal{X}) \equiv \{g \in \Lambda^\gamma(\mathcal{X}) : ||g||_{\Lambda^\gamma} \leq c < \infty\}.$$

Thus, we have $l(X) \in \Lambda_{c_1}^{\gamma_1}(\mathcal{X})$. Functions in $\Lambda_c^\gamma(\mathcal{X})$ can be approximated well by various sieves such as power series, Fourier series, splines, and wavelets.

The functions in $\mathcal{L}_T$ is getting dense as $T \to \infty$ but not that fast, i.e. $k_1 \to \infty$ as $T \to \infty$ but $k_1/T \to 0$. Then, according to Theorem 8, p.90 in Lorentz (1986), there exists a π_{k_1}[29] such that for $R^{k_1}(x)$ on the compact set $\mathcal{X}$ (the support of X)

$$\sup_{x\in\mathcal{X}} |l^*(x) - R^{k_1}(x)'\pi_{k_1}| < c_1 {k_1}^{-\frac{[\gamma_1]}{d_x}}, \tag{16}$$

where $[s]$ is the largest integer less than s and d_x is the dimension of X. Thus, in what follows, we approximate the pseudo-value v_i in (5) as

$$V_i^{k_1} = \ln \mathbb{V}_i - l_{k_1}(X), \tag{17}$$

where $l_{k_1}(x) = R^{k_1}(x)'\pi_{k_1}$.

Specifically, we consider the following polynomial basis considered by Newey, Powell, and Vella (1999). First let $\mu = (\mu_1, \ldots, \mu_{d_x})'$ denote a vector of nonnegative integers with the norm $|\mu| = \sum_{j=1}^{d_x} \mu_j$, and let $x^\mu \equiv \prod_{j=1}^{d_x}(x_j)^{\mu_j}$. For a sequence $\{\mu(k)\}_{k=1}^\infty$ of distinct such vectors, we construct a tensor-product power series sieve as

$$R^{k_1}(x) = (x^{\mu(1)}, \ldots, x^{\mu(k_1)})'$$

Then, replacing each power x^μ by the product of orthonormal univariate polynomials of the same order, we may reduce collinearity.

[29] We will suppress the argument T in $k_1(T)$, unless otherwise noted from now on

D Distribution of Valuations

In this section, to simplify the discussion, we assume that there is no observed heterogeneity in the auction. In other words, we impose $\ln \mathbb{V} = V_i$, i.e. $l^*(\cdot) \equiv 0$. First, consider the estimation of the distribution of valuations using the second- and third-highest bids in each auction. Let $(\widetilde{V}_t, V_t)$ denote the second- and fourth-highest pseudo-bids for each auction t and let $(\widetilde{v}_t, v_t)$ denote their realizations, respectively. Also let $v_m = \min_t v_t$ and $v_M = \max_t \widetilde{v}_t$[30] Noting $F(v)$ for $v < v_m$ nor $v > v_M$ can be recovered form the data, we treat $F^*(\cdot) = F(\cdot|v_m, v_M)$ as the model primitive of interest, where $F(\cdot|v_m, v_M)$ denotes the truncated distribution of $F(\cdot)$ from below at v_m and from above at v_M as

$$F^*(v) \equiv F(v|v_m, v_M) = \frac{F(v) - F(v_m)}{F(v_M) - F(v_m)}$$

and hence

$$f^*(v) \equiv f(v|v_m, v_M) = \frac{f(v)}{F(v_M) - F(v_m)} \tag{18}$$

Then, we obtain the density of $\widetilde{V}_i$ conditional on V_i denoted by $p_{2|4}(\widetilde{v}_i|V_i = v_i)$ from (3) as

$$p_{2|4}(\widetilde{v}|V = v) = \frac{6(F(\widetilde{v}) - F(v))(1 - F(\widetilde{v}))f(\widetilde{v})}{(1 - F(v))^3} \quad \text{for } v_M \geq \widetilde{v} > v \geq v_m. \tag{19}$$

To estimate the unknown function $f^*(z)$ (hence, $F^*(z) = \int_{v_m}^{z} f^*(t)dt$), we first approximate $f^*(z)$ with $f^K(z)$ as a member of $\mathcal{F}_T$ up to the order $K(T)$:[31]

$$\mathcal{F}_T = \{f^K : \int_{v_m}^{v_M} f^K(z)dz = 1 \text{ and } 0 < f^K(\cdot) < \infty\}$$

One possible specification of $f^K(\cdot)$ is

$$f^K(z) = \frac{\left(1 + \sum_{j=1}^{K} a_j \left(\frac{z-\mu}{\sigma}\right)^{j-1}\right)^2 \phi(z; \mu, \sigma, c)}{\int_{v_m}^{v_M} \left(1 + \sum_{j=1}^{K} a_j \left(\frac{t-\mu}{\sigma}\right)^{j-1}\right)^2 \phi(t; \mu, \sigma, c)dt}, \tag{20}$$

which is proposed by Song (2005) when $v_M = \infty$, where $\phi(\cdot; \mu, \sigma, v_m, v_M)$ is the density of $N(\mu, \sigma)$ truncated below at v_m and above at v_M. Note this is an extension of the SNP density specification of Gallant and Nychka (1987) (in the univariate case) to the truncated distribution.

Then, we construct the sample likelihood based on $f^K(\cdot)$ instead of the true $f(\cdot)$ using (19):

$$L(f^K; \widetilde{v}_t, v_t) = \frac{6(F^K(\widetilde{v}_t) - F^K(v_t))(1 - F^K(\widetilde{v}_t))f^K(\widetilde{v}_t)}{(1 - F^K(v_t))^3}, \tag{21}$$

[30] Note that v_m is a consistent estimator of $\underline{v}$ under no binding reserve price and a consistent estimator of the reserve price under the binding case. Similarly v_M is a consistent estimator of $\overline{v}$.

[31] We will suppress the argument T in $K(T)$ unless noted otherwise

where $F^K(z) = \int_{v_m}^{z} f^K(t)dt$. Noting that (21) is a parametric estimation problem for a given value of K, one could approximate $f^K(\cdot)$ with $\hat{f}(\cdot)$ as the maximum likelihood estimator:

$$\hat{f}(z) = \frac{\left(1 + \sum_{j=1}^{K} \hat{a}_j \left(\frac{z-\hat{\mu}}{\hat{\sigma}}\right)^{j-1}\right)^2 \phi(z; \hat{\mu}, \hat{\sigma}, c)}{\int_{v_m}^{v_M} \left(1 + \sum_{j=1}^{K} \hat{a}_j \left(\frac{t-\hat{\mu}}{\hat{\sigma}}\right)^{j-1}\right)^2 \phi(t; \hat{\mu}, \hat{\sigma}, c)dt},$$

where

$$(\hat{a}_1, \ldots, \hat{a}_K, \hat{\mu}, \hat{\sigma}) = \arg \max_{a_1,\ldots,a_K,\mu,\sigma>0} \frac{1}{T} \sum_{t=1}^{T} \ln L(f^K; y_t, x_t) \tag{22}$$

Now note that actually a pseudo-bid z is defined as the residual in (5) and is approximated as the residual in (17)[32]. Thus, we have another set of parameter (π_{k_1}) to estimate in (22) as

$$\hat{f}(\hat{z}) = \frac{\left(1 + \sum_{j=1}^{K} \hat{a}_j \left(\frac{\hat{z}-\hat{\mu}}{\hat{\sigma}}\right)^{j-1}\right)^2 \phi(\hat{z}; \hat{\mu}, \hat{\sigma}, c)}{\int_{v_m}^{v_M} \left(1 + \sum_{j=1}^{K} \hat{a}_j \left(\frac{t-\hat{\mu}}{\hat{\sigma}}\right)^{j-1}\right)^2 \phi(t; \hat{\mu}, \hat{\sigma}, c)dt},$$

where

$$(\hat{\pi}, \hat{a}_1, \ldots, \hat{a}_K, \hat{\mu}, \hat{\sigma}) = \arg \max_{\pi,a_1,\ldots,a_K,\mu,\sigma>0} L(f^K). \tag{23}$$

Note that our estimator requires a rich data set, since we estimate two nonparametric functions at the same time. The approximation precision depends on the choice of smoothing parameters k_1 and K. Here, we pick the optimal length of series (the dimension of the sieve space $\mathcal{H}_T$), K^*, following the Coppejans and Gallant (2002)'s method, which is a cross-validation strategy as used in Kernel density estimations. Using a similar idea[33], we also can pick the optimal k_1^*.

One obvious problem of using the usual polynomials in (20) is that the estimation may be unstable due to underflows and overflows in $\sum_{j=1}^{K} a_j \left(\frac{z-\mu}{\sigma}\right)^j$. Moreover the estimated conditional

[32] We develop our discussion based on the nonparametric sepcification of $l^*(\cdot)$, which nests the parametric case.

[33] We can use a sample version of the Mean Squared Error criterion for the cross-validation as

$$SMSE(\hat{l}) = \frac{1}{T} \sum_{i=1}^{T} [\hat{l}(X_i) - l^*(X_i)]^2,$$

where $\hat{l}(\cdot) = R^K(\cdot)'\hat{\pi}$. We again use the *Leave-one partition-out* method to reduce the computational burden. Namely, we estimate the function $l^*(\cdot)$ from the sample after deleting the p^{th} group with the length of the series equal to k_1 and denote this as $\hat{l}_{p,k_1}(\cdot)$. As a next step, we choose k_1^* such that

$$\text{argmin}_{k_1} CV(k_1) = \frac{1}{T} \sum_{p=1}^{P} \sum_{t \in T_p} [\ln V_t - D_t'\hat{\gamma} - \hat{l}_{p,k_1}(X_t)]^2,$$

where T_p denotes the set of the data indices belonging to the p^{th} group such that $T_p \cap T_{p'} = \phi$ for $p \neq p'$ and $\cup_{p=1}^{P} T_p = \{1, 2, \ldots, T\}$.

density $\hat{f}(\cdot)$ itself is a complicated function of truncated parameters and thus it is difficult to establish the convergence or the convergence rate of $\hat{f}(\cdot)$ to $f^*(z)$. On the other hand this approach is absurd in the sense that we implicitly assume that the support of the values is $(-\infty, \infty)$ where the true support is $(\underline{v}, \overline{v})$.

The first problem of instability can be resolved by using a truncated density version of the specification experimented in Fenton and Gallant (1996) as

$$f^{FG}(z;\theta) = \frac{f_{FG}(z;\theta)}{\int_{v_m}^{v_M} f_{FG}(z;\theta)dz} = \frac{\left(\sum_{j=1}^{K} \theta_j H_j(\frac{z-\mu}{\sigma})\right)^2}{\int_{v_m}^{v_M} \left(\sum_{j=1}^{K} \theta_j H_j(\frac{z-\mu}{\sigma})\right)^2 dz}, \tag{24}$$

where

$$f_{FG}(z;\theta) = \frac{\left(\sum_{j=1}^{K} \theta_j H_j(\frac{z-\mu}{\sigma})\right)^2 + \epsilon_0 \phi(\frac{z-\mu}{\sigma})}{\sum_{j=1}^{K} \theta_j^2 + \epsilon_0} \tag{25}$$

for a small positive constant ϵ_0. $H_j(t)$ is defined recursively as

$$\begin{aligned} H_1(t) &= (\sqrt{2\pi})^{-1/2} e^{-t^2/4}, \\ H_2(t) &= (\sqrt{2\pi})^{-1/2} t e^{-t^2/4}, \\ H_j(t) &= [tH_{j-1}(t) - \sqrt{j-1} H_{j-2}(t)]/\sqrt{j}, \text{ for } j \geq 3, \end{aligned} \tag{26}$$

which is based on the normalized Hermite polynomials. Replacing $f^K(z;\theta)$ in (21) with its alternative (24) and solving (23), one can obtain estimator denoted by $\hat{f}_{FG}(\cdot)$. Fenton and Gallant (1996) argue that the estimation based on (25) is much more efficient computationally and is more stable. Moreover, it is less sensitive to the choice of starting values.

However, the second and the third problem still remain. Here we entertain the specification proposed in Kim (2005) where a truncated version of the SNP density estimator with a compact support is developed. Hereafter, to simplify the notation, we assume $(\mu, \sigma) = (0, 1)$ without loss of generality since we can always standardize the data before the estimation and also the SNP estimator is the scale- and location- invariant. Instead of defining the true density as (18), we follow Kim (2005)'s specification by denoting the true density as

$$f_0(z) = h_{f_0}^2(z) e^{-z^2/2} + \epsilon_0 \frac{\phi(z)}{\int_{\mathcal{V}} \phi(z)dz}$$

where $\mathcal{V}$ denotes the support of z. Kim (2005) uses a truncated version of Hermite polynomials to approximate a density function f_0 with a compact support as

$$\underline{\mathcal{F}}_T = \left\{ f : f(z,\theta) = \left(\sum_{j=1}^{K} \vartheta_j w_{jK}(z) \right)^2 + \epsilon_0 \frac{\phi(z)}{\int_{\mathcal{V}} \phi(z)dz}, \theta \in \Theta_T \right\}, \tag{27}$$

where $\Theta_T = \left\{\theta = (\vartheta_1, \ldots, \vartheta_{K(T)}) : \sum_{j=1}^{K(T)} \vartheta_j^2 + \epsilon_0 = 1\right\}$ and $\{w_{jK}(z)\}$ are defined following Kim (2005) as follows. First we define

$$\overline{w}_{jK}(z) = \frac{H_j(z)}{\sqrt{\int_{\mathcal{V}} H_j^2(z)dx}}$$

that is bounded by

$$\sup_{z\in\mathcal{V}, j\leq K} |\overline{w}_{jK}(z)| \leq \frac{1}{\sqrt{\min\limits_{j\leq K} \int_{\mathcal{Z}} H_j^2(v)dv}} \sup_{z\in\mathcal{V}} |H_j(z)| < C\widetilde{H}$$

for some constant $C < \infty$, since $\int_{\mathcal{V}} H_j^2(z)dz$ is bounded away from zero for all j and $|H_j(z)| < \widetilde{H}$ uniformly over z and j. Denoting $\overline{W}^K(z) = (\overline{w}_{1K}(z), \ldots, \overline{w}_{KK}(z))'$, further define $Q_{\overline{W}} = \int_{\mathcal{V}} \overline{W}^K(z)\overline{W}^K(z)'dz$ and its symmetric matrix square root as $Q_{\overline{W}}^{-1/2}$. Now let

$$W^K(z) \equiv (w_{1K}(z), \ldots, w_{KK}(z))' \equiv Q_{\overline{W}}^{-1/2}\overline{W}^K(z) \tag{28}$$

then by construction, we have $\int_{\mathcal{V}} W^K(z)W^K(z)' = I_K$. Then these truncated and transformed Hermite polynomials are orthonormal

$$\int_{\mathcal{V}} w_{jK}^2(z)dz = 1, \int_{\mathcal{V}} w_{jK}(z)w_{kK}(z)dz = 0, j \neq k$$

from which the condition $\sum_{j=1}^{K(n)} \vartheta_j^2 + \epsilon_0 = 1$ follows since for any f in $\underline{\mathcal{F}}_T$, we have $\int_{\mathcal{V}} f dz = 1$. Now define $\zeta(K) = \sup_{z\in\mathcal{V}} \left\|W^K(z)\right\|$ using a matrix norm $\|A\| = \sqrt{\mathrm{tr}(A'A)}$ for a matrix A, which is the Euclidian norm for a vector. Then, we have $\zeta(K) = O(\sqrt{K})$ as shown in Lemma I.1. If the range of $\mathcal{V}$ are sufficiently large, then $\int_{\mathcal{V}} H_j^2(z)dz \approx 1$ and $Q_{\overline{W}} \approx I_K$ and hence $w_{jK} \approx H_j$ which implies immediately

$$\sup_{z\in\mathcal{V}} \left\|W^K(x)\right\| \approx \sup_{x\in\mathcal{X}} \sqrt{\sum_{j=1}^{K} H_j^2} \leq \sqrt{K\widetilde{H}^2} = O\left(\sqrt{K}\right).$$

Here we need to introduce a trimming device $\tau(\cdot)$ that trims out those observations $\widetilde{v}_t, v_t > \overline{v} - \varepsilon$ or $\widetilde{v}_t, v_t < \underline{v} - \varepsilon$.

Now the SNP estimator is obtained by solving

$$\begin{aligned}\widehat{f}_1 &= \operatorname*{argmax}_{f\in\underline{\mathcal{F}}_T} \frac{1}{T}\sum_{t=1}^{T} \tau(\widetilde{v}_t, v_t) \ln L(f; \widetilde{v}_t, v_t) \qquad (29)\\ &\equiv \operatorname*{argmax}_{f\in\underline{\mathcal{F}}_T} \frac{1}{T}\sum_{t=1}^{T} \tau(\widetilde{v}_t, v_t) \ln \frac{6(F(\widetilde{v}_t) - F(v_t))(1 - F(\widetilde{v}_t))f(\widetilde{v}_t)}{(1 - F(v_t))^3}\end{aligned}$$

where we redefine $F(z)=\int_{\underline{v}}^{z} f(z)dz$ or equivalently

$$\widehat{f}_1 = f(\cdot,\widehat{\theta}_K),\ \widehat{\theta}_K = \underset{\theta\in\Theta_n}{\operatorname{argmax}} \frac{1}{n}\sum_{i=1}^{n} \tau(\widetilde{v}_t, v_t)\ln L(f(\cdot,\widehat{\theta}_K);\widetilde{v}_t, v_t).$$

In the Technical Appendix F, we establish the consistency and the convergence rate of this SNP estimator. Similarly we can also identify and estimate $f_0(\cdot)$ using the pair of third- and fourth-highest bids. First, denote $(\widetilde{\widetilde{V}}_t, V_t)$ to be the third- and fourth-highest pseudo-bids for each auction t and let $(\widetilde{\widetilde{v}}_t, v_t)$ denote their realizations, respectively. Then, we have the conditional density

$$p_{3|4}(\widetilde{\widetilde{v}}|V=v) = \frac{3(1-F_0(\widetilde{\widetilde{v}}))^2 f_0(\widetilde{\widetilde{v}})}{(1-F_0(v))^3} \text{ for } \overline{v} > \widetilde{\widetilde{v}} \geq v \geq \underline{v} \tag{30}$$

from (3). Denote the estimate of $f_0(\cdot)$ based on (30) as $\hat{f}_2(\cdot)$:

$$\widehat{f}_2 = \underset{f\in\underline{\mathcal{F}}_T}{\operatorname{argmax}} \frac{1}{T}\sum_{t=1}^{T} \tau(\widetilde{\widetilde{v}}_t, v_t)\ln \frac{3(1-F(\widetilde{\widetilde{v}}_t))^2 f(\widetilde{\widetilde{v}}_t)}{(1-F(v_t))^3} \tag{31}$$

for $f\in\underline{\mathcal{F}}_T$ defined in (27).

E Comparison of Densities Using the Pseudo Kullback-Leibler Divergence Measure

The idea of using (11) as a valid measure of comparing two densities is from the spirit of indirect inference literature (see Gourieroux, Monfort, and Renault (1993), An and Liu (2000), Dridi, Guay, and Renault (2003), and Keane and Smith (2003) among others). It is argued that the parameters of interest can be inferred from some instrumental models which are possibly misspecified. It is noted that (11) is not a distance measure since it can have negative values and is not symmetric but still can serve as a divergence measure. Now define g_2, g_3, and g_4 to be the density of the second-, third-, and fourth- highest order statistics among n, respectively and hence $g_2(\cdot)=g^{(n-1:n)}$, $g_3(\cdot)=g^{(n-2:n)}$, and $g_4(\cdot)=g^{(n-3:n)}$. We consider a modification of (11) as a divergence measure of two densities, which is defined by

$$I^{g_2}(f_{o1}, f_{o2}) = \left(A\int_{\mathcal{V}} (\ln f_{o1}(z) - \ln f_{o2}(z))\, g^{(n-1:n)}(z)dz\right)^2 \tag{32}$$

for some positive constant A.

Using (32), for $\widehat{A}\to A$, we propose a test statistic of the form using the second-highest order statistics in particular[34]

$$\widehat{I}^{g_2}(\widehat{f}_1,\widehat{f}_2) = \left(\widehat{A}\frac{1}{T}\sum_{t\in\mathcal{T}_2}\left(\ln\widehat{f}_1(\widetilde{v}_t) - \ln\widehat{f}_2(\widetilde{v}_{t+1})\right)\right)^2 \tag{33}$$

[34] Instead, we can use other order statistics. In that case, simply replace g_2 with g_3 or g_4.

where $\mathcal{T}_2$ is a subset of $\{1,2,,\ldots,T-1\}$, which trims out those observations of $\widehat{f}_1(\cdot)<\delta_1(T)$ or $\widehat{f}_2(\cdot)<\delta_2(T)$ for chosen positive values of $\delta_1(T)$ and $\delta_2(T)$ that tend to zero as $T\to\infty$. To be precise, $\mathcal{T}_2$ is defined by

$$\mathcal{T}_2=\left\{t:1\le t\le T-1 \text{ such that } \widehat{f}_1(\widetilde{v}_t)>\delta_1(T) \text{ and } \widehat{f}_2(\widetilde{v}_{t+1})>\delta_2(T)\right\}.$$

This trimming is a usual device in a inference procedure for nonparametric estimations. Even though the SNP density estimator that we are interested in is always positive by construction differently from higher order Kernel estimators, we still introduce this trimming device to avoid the excess influence of one or several summands when $\widehat{f}_1(\cdot)$ or $\widehat{f}_2(\cdot)$ are arbitrary small. (33) will converge to

$$I^{g_2}(f_{o1},f_{o2})=\left(A\int_{\mathcal{V}}\left(\ln f_{o1}(z)-\ln f_{o2}(z)\right)g^{(n-1:n)}(z)dz\right)^2$$

under certain conditions that will be discussed later. However, unfortunately, both $T\widehat{I}^{g_2}(\widehat{f}_1,\widehat{f}_2)$ will have degenerate distributions under the null similarly as discussed in Robinson (1991) and cannot be used as reasonable statistics. To resolve this problem, we entertain modification of (33) in the spirit of Robinson (1991) as

$$\widehat{I}_{\gamma}^{g_2}(\widehat{f}_1,\widehat{f}_2)=\left(\widehat{A}\frac{1}{T_\gamma-1}\sum_{t\in\mathcal{T}_2}c_t(\gamma)\left(\ln\widehat{f}_1(\widetilde{v}_t)-\ln\widehat{f}_2(\widetilde{v}_{t+1})\right)\right)^2$$

where for a nonnegative constant γ,

$$\begin{aligned} c_t(\gamma) &= 1+\gamma \ \text{ if } t \text{ is } \textit{odd} \\ &= 1-\gamma \ \text{ if } t \text{ is } \textit{even} \end{aligned}$$

and n_γ is defined as[35]

$$T_\gamma=T+\gamma \ \text{ if } T \text{ is } \textit{odd} \text{ and } T_\gamma=T \ \text{ if } T \text{ is } \textit{even.} \tag{34}$$

We let

$$I^{g_2}=I^{g_2}(f_{01},f_{02}),\ \widetilde{I}_\gamma^{g_2}=\widehat{I}_\gamma^{g_2}(f_{01},f_{02}), \text{ and } \widehat{I}_\gamma^{g_2}=\widehat{I}_\gamma^{g2}(\widehat{f}_1,\widehat{f}_2)$$

for notational simplicity. Now note, for any increasing sequence $d(T)$, any positive $C<\infty$, and γ,

$$\begin{aligned} \Pr(d(T)\widehat{I}_\gamma^{g_2}<C) &\le \Pr(d(T)\left|\widehat{I}_\gamma^{g_2}-I^{g_2}\right|>d(T)I^{g_2}-C) \\ &\le \Pr(\left|\widehat{I}_\gamma^{g_2}-I^{g_2}\right|>I^{g_2}/2) \end{aligned}$$

[35] Consider $s(n)\equiv\frac{1}{n_\gamma}\sum_{i=1}^{n}c_i(\gamma)$ when $n=2m$ and $n=2m+1$, respectively. It follows that $s(2m)=\frac{((1+\gamma)m+(1-\gamma)m)}{n_\gamma}=\frac{2m}{n_\gamma}=\frac{n}{n_\gamma}$ and $s(2m+1)=\frac{1}{n_\gamma}\left((1+\gamma)m+(1-\gamma)m+(1+\gamma)\right)=\frac{2m+1+\gamma}{n_\gamma}=\frac{n+\gamma}{n_\gamma}$. Thus, by constructing n_r as (34), we have $s(n)=1$.

holds when T is sufficiently large and (7) is not true (i.e. $I^{g_2} > 0$). Since the probability $\Pr(\left|\widehat{I}_{\gamma}^{g_2} - I^{g_2}\right| > I^{g_2}/2)$ goes to zero under the alternative as long as $\widehat{I}_{\gamma}^{g_2} \underset{p}{\longrightarrow} I^{g_2}$, one could construct a test statistic of the form

$$\text{Reject (7) when } d(T)\widehat{I}_{\gamma}^{g_2} > C. \tag{35}$$

Therefore, as long as $\widehat{I}_{\gamma}^{g_2} \underset{p}{\longrightarrow} I^{g_2}$, (35) is a valid test consistent against all departures from (7). We call a test statistic consistent against one direction of departure from the null hypothesis if the rejection probability approaches one as the sample size gets large regardless of the size of that departure.

We present the main result of this section in the following theorem. See the Technical Appendix G for the proof.

Theorem E.1 *Suppose Assumption F.3 in the Technical Appendix holds. Provided that Conditions 1-3 in the Technical Appendix hold under (7), we have*

$$\widehat{\tau}_{\gamma} \equiv T\widehat{I}_{\gamma}^{g_2} \underset{d}{\longrightarrow} \chi^2(1)$$

for any $\gamma > 0$ with $\widehat{A} = 1/\sqrt{2}\gamma\widehat{\Sigma}^{\frac{1}{2}}$ and $\widehat{\Sigma} \underset{p}{\longrightarrow} \Sigma$. Thus we reject (7) if $\widehat{\tau}_{\gamma} > C_{\alpha}$ where C_{α} is the size α critical value of the $\chi^2(1)$ distribution.

F Large Sample Theory of the SNP Density Estimator

F.1 Consistency of the SNP Estimator

Here we impose the following regularity conditions. We again develop our discussion when $l^*(\cdot)$ is nonparametrically specified, which nests the parametric case.

Assumption F.1 $(\mathbb{V}_{tj}, X_t), \ldots (V_{Tj}, X_T)$ *are i.i.d. for all j and* $\mathrm{Var}(\mathbb{V}_j|X)$ *is bounded for all j.*

Assumption F.2 *(i) the smallest and the largest eigenvalue of* $\mathrm{E}[R^{k_1}(X)R^{k_1}(X)']$ *is bounded away from zero uniformly in k_1 and; (ii) there is a sequence of constants $\zeta_0(k_1)$ satisfying* $\sup_{x\in\mathcal{X}}\left\|R^{k_1}(x)\right\| \leq \zeta_0(k_1)$ *and $k_1 = k_1(T)$ such that $\zeta_0(k_1)^2 k_1/T \to 0$ as $T \to \infty$, where the matrix norm $\|A\| = \sqrt{trace(A'A)}$.*

Under Assumption F.1 and F.2 (which are essentially borrowed from Assumption 1 and 2 in Newey (1997)), we obtain the convergence rate of $\hat{l}(\cdot) = R^{k_1}(\cdot)'\hat{\pi}$ to $l^*(\cdot)$ in the sup-norm by Theorem 1 of Newey (1997), since (16) implies Assumption 3 in Newey (1997) for the polynomial series approximation. Theorem 1 in Newey (1997) states that

$$\sup_{x\in\mathcal{X}}\left|\hat{l}(x) - l^*(x)\right| = O_p(\zeta_0(k_1))[\sqrt{k_1}/\sqrt{T} + k_1^{-\frac{[\gamma_1]}{d_x}}].$$

Noting $\zeta_0(k_1) \le O(k_1)$ for power series sieves, we have

$$\begin{aligned} \sup_{x \in \mathcal{X}} \left| \hat{l}(x) - l^*(x) \right| &= \max\left\{ O_p(k_1^{3/2}/\sqrt{T}), O_p(k_1^{1-\frac{[\gamma_1]}{d_x}}) \right\} \\ &= O_p\left(T^{\max\left\{ \frac{3}{2}\varphi - \frac{1}{2}, \left(1 - \frac{[\gamma_1]}{d_x}\right)\varphi \right\}} \right) \end{aligned} \tag{36}$$

with $k_1 = O(T^{\vartheta})$ and $0 < \vartheta < \frac{1}{3}$.

The convergence rate result of (36) implies that by choosing a proper ϑ, we can insure that the convergence rate of the second step density estimator will not be affected by the first step estimation[36]. Moreover, the asymptotic distributions of the second step density estimator or its functional may not be affected by the first estimation step, which makes our discussion a lot easier. This is noted also in Hansen (2004) when he derives the asymptotic distribution of the two-step density estimator which contains the estimates of the conditional mean in the first step. This result will be immediately obtained if we consider a parametric specification of observed auction heterogeneities since we will achieve the $\sqrt{n}$ consistency for those finite parameters.

Now following Kim (2005), we consider the convergence rate of the SNP density estimator obtained from any pair of order statistics. Denote (q_t, r_t) to be a pair of k_1-th and k_2- th highest order statistics. We first derive the convergence rate of the SNP density estimator when the data $\{(q_t, r_t)\}_{t=1}^T$ are from the true values (after removing the observed heterogeneity part), not derived ones from the first step estimation described in the previous section. First, we construct

$$L(f; q_t, r_t) = \frac{(k_2 - 1)!}{(k_2 - k_1 - 1)!(k_1 - 1)!} \frac{(F(q_t) - F(r_t))^{k_2 - k_1 - 1}(1 - F(q_t))^{k_1 - 1} f(q_t)}{(1 - F(r_t))^{k_2 - 1}} \tag{37}$$

where $f \in \underline{\mathcal{F}}_T$. Then the SNP density estimator $\widehat{f}$ is obtained by solving

$$\widehat{f} = \underset{f \in \underline{\mathcal{F}}_T}{\operatorname{argmax}} \frac{1}{T} \sum_{t=1}^{T} \ln L(f(\cdot, \theta); q_t, r_t) \tag{38}$$

or equivalently

$$\widehat{f} = f(z, \widehat{\theta}_K),\ \widehat{\theta}_K = \underset{\theta \in \Theta_T}{\operatorname{argmax}} \frac{1}{T} \sum_{t=1}^{T} \ln L(f(\cdot, \theta); q_t, r_t).$$

To establish the convergence rate, we need the following conditions

Assumption F.3 *The observed data of a pair of order statistics $\{(q_t, r_t)\}$ are randomly drawn from the continuous density of the parent distribution $f_0(\cdot)$.*

[36] We have not derived this proper rate of convergence in this study. It is part of our future research. However, comparing (36) and (39), we conjecture that we will need, for any small $\delta > 0$, $\frac{\max\left\{ T^{\frac{3}{2}\varphi - \frac{1}{2}}, T^{1 - \frac{[\gamma_1]}{d_x}} \right\}}{\max\left\{ T^{-1/2 + \alpha/2 + \delta}, K(T)^{-s/2} \right\}} \to 0$ to ensure that the convergence rate of the second step density estimator is not affected by the first step estimation.

Assumption F.4 *(i) $f_0(z)$ is s-times continuously differentiable with $s \geq 3$, (ii) uniformly bounded from above and bounded away from zero on its compact support $\mathcal{V}$, (iii) $f_0(z)$ has the form of $f_0(z) = h_{f_0}^2(z)e^{-v^2/2} + \epsilon_0 \frac{\phi(z)}{\int_{\mathcal{V}} \phi(z)dz}$ for arbitrary small positive number ϵ_0*

Note that differently from Fenton and Gallant (1996) and Coppejans and Gallant (2002), we do not require a tail condition since we impose the compact truncated support. Under Assumption F.4, we obtain

Theorem F.1 *Suppose Assumption F.3 and F.4 holds and $\frac{\zeta(K)^2 K}{T} \to 0$. Then, for $K = O(T^{\alpha})$ with $\alpha < \frac{1}{3}$, we have*

$$\sup_{z\in\mathcal{V}} \left| \widehat{f}(z) - f_0(z) \right| = O\left(\zeta(K)^2\right) o_p\left(T^{-1/2+\alpha/2+\delta}\right) + O\left(\zeta(K)^2 K^{-s/2}\right) \tag{39}$$

for arbitrary small positive constant δ.

We prove this theorem in the following section. Now we consider the convergence rate of the SNP estimators where a pair of order statistics $\{(\widehat{q}_t, \widehat{r}_t)\}$ is obtained as residuals of the first step estimation. The SNP estimator is given by

$$\widehat{\widehat{f}} = \underset{f\in\underline{\mathcal{F}}_T}{\operatorname{argmax}} \frac{1}{T} \sum_{t=1}^{T} \ln L(f; \widehat{q}_t, \widehat{r}_t, \theta) \tag{40}$$

We note that under a suitable choice on the degree of approximation for the unknown function of observed heterogeneity, the effect of the first step estimation is negligible. Particularly, this is true when $l^*(\cdot)$ is parametrically specified.

F.2 Convergence Rate of the SNP Estimator (Theorem F.1)

Here we derive the convergence rate of the SNP density estimator given in (38) following Kim (2005).

Recall that we denote (q_t, r_t) to be a pair of k_1-th and k_2- th highest order statistics ($k_1 < k_2$). We first derive the convergence rate of the SNP density estimator when the data $\{(q_t, r_t)\}_{t=1}^T$ are from the true residual values (after removing the observed heterogeneity part). Though we use a particular sieve here, we derive the convergence rate results for a general sieve that satisfies some conditions. According to Theorem 8, p.90, in Lorentz (1986), we can approximate a v-times continuously differentiable function h such that there exists a K-vector γ_K that satisfies

$$\sup_{z\in\mathcal{Z}} \left| h(z) - R^K(z)'\gamma_K \right| = O(K^{-\frac{v}{\dim(z)}}) \tag{41}$$

where $\mathcal{Z} \subset \mathcal{R}^{\dim(z)}$ is the compact support of z and $R^K(z)$ is a triangular array of polynomials. Now let

$$f_0(z) = h_{f_0}^2(z)e^{-z^2/2} + \epsilon_0 \frac{\phi(z)}{\int_{\mathcal{V}} \phi(z)dz}$$

and assume $h_{f_0}(z)$ (and hence $f_0(z)$) is s-times continuously differentiable. Denote a K-vector $\theta_K = (\vartheta_{1K}, \ldots, \vartheta_{KK})'$. Then, there exists a θ_K such that

$$\sup_{z\in\mathcal{V}} \left| h_{f_0}(z) - e^{z^2/4} W^K(z)'\theta_K \right| = O(K^{-s}) \tag{42}$$

by (41) noting $h_{f_0}(z)$ is s-times continuously differentiable over $z \in \mathcal{V}$, $\mathcal{V}$ is compact, and $\left\{ e^{z^2/4} w_{jK}(z) \right\}$ are linear combinations of power series. (42) implies that

$$\sup_{z\in\mathcal{V}} \left| h_{f_0}(z)e^{-z^2/4} - W^K(z)'\theta_K \right| \leq \sup_{z\in\mathcal{V}} e^{-z^2/4} \sup_{z\in\mathcal{V}} \left| h_{f_0}(z) - e^{z^2/4} W^K(z)'\theta_K \right| = O(K^{-s}) \tag{43}$$

from $\sup_{z\in\mathcal{R}} e^{-z^2/4} \leq 1$. From this result, now it is shown below that for $f(z, \theta_K) \in \underline{\mathcal{F}}_T$,

$$\sup_{z\in\mathcal{V}} |f_0(z) - f(z, \theta_K)| = O\left(\zeta(K)K^{-s}\right)$$

First, note (43) implies

$$W^K(z)'\theta_K - O(K^{-s}) \leq h_{f_0}(z)e^{-z^2/4} \leq W^K(z)'\theta_K + O(K^{-s})$$

from which it follows that

$$\begin{aligned} \left(W^K(z)'\theta_K - O(K^{-s})\right)^2 - \left(W^K(z)'\theta_K\right)^2 &\leq h_{f_0}^2(z)e^{-x^2/2} - \left(W^K(z)'\theta_K\right)^2 \\ &\leq \left(W^K(z)'\theta_K + O(K^{-s})\right)^2 - \left(W^K(z)'\theta_K\right)^2 \end{aligned} \tag{44}$$

assuming $W^K(z)'\theta_K$ is positive without loss of generality. Now, note that

$$\sup_{z\in\mathcal{V}} \left| W^K(z)'\theta \right| \leq \sup_{z\in\mathcal{V}} \left\| W^K(z) \right\| \|\theta\| = O\left(\zeta(K)\right) \tag{45}$$

by the Cauchy-Schwarz inequality and from $\|\theta\|^2 < 1$ for any $\theta \in \Theta_n$ by construction. Now applying the mean value theorem to the upper bound of (44), we have

$$\begin{aligned} & \sup_{z\in\mathcal{V}} \left| \left(W^K(z)'\theta_K + O(K^{-s})\right)^2 - \left(W^K(z)'\theta_K\right)^2 \right| \\ = \; & \sup_{z\in\mathcal{V}} \left| \left(2W^K(z)'\theta_K + O(K^{-s})\right) O(K^{-s}) \right| \\ \leq \; & \sup_{z\in\mathcal{V}} \left| 2W^K(z)'\theta_K \right| O(K^{-s}) + O(K^{-2s}) \\ = \; & O(\zeta(K)K^{-s}) \end{aligned}$$

where the last result is from (45). Similarly for the lower bound, we have

$$\sup_{z\in\mathcal{V}}\left|\left(W^K(z)'\theta_K - O(K^{-s})\right)^2 - \left(W^K(z)'\theta_K\right)^2\right| = O(\zeta(K)K^{-s}).$$

$>$From (44), it follows that

$$\sup_{z\in\mathcal{V}}\left|h_{f_0}^2(z)e^{-z^2/2} - \left(W^K(z)'\theta_K\right)^2\right| = O(\zeta(K)K^{-s})$$

and hence

$$\sup_{z\in\mathcal{V}}|f_0(z) - f(z,\theta_K)| = O\left(\zeta(K)K^{-s}\right). \tag{46}$$

Now to establish the convergence rate of the SNP estimator, a pseudo true density function is introduced where the pseudo true density is given by

$$f_K^*(z) = \left(W^K(z)'\theta_K^*\right)^2 + \epsilon_0\frac{\phi(z)}{\int_{\mathcal{V}}\phi(z)dz}$$

such that for $L(f(\cdot,\theta);q_t,r_t)$ defined in (37)

$$\theta_K^* = \operatorname*{argmax}_{\theta} E\left[\ln L(f(\cdot,\theta);q_t,r_t)\right].$$

To simplify notation, we only consider such case that $k_2 - k_1 - 1 = 0$ and let

$$\widetilde{L}(f(\cdot,\theta);q_t,r_t) = \frac{(1-F(q_t,\theta))^{k_1-1}f(q_t,\theta)}{(1-F(r_t,\theta))^{k_2-1}}.$$

Then we have

$$\theta_K^* = \operatorname*{argmax}_{\theta} E\left[\ln\widetilde{L}(f(\cdot,\theta);q_t,r_t)\right] \tag{47}$$

and

$$\widehat{\theta}_K = \operatorname*{argmax}_{\theta}\frac{1}{T}\sum_{t=1}^{T}\ln\widetilde{L}(f(\cdot,\theta);q_t,r_t)$$

from (38). We first note

Lemma F.1 *Suppose Assumption F.4 holds. Then for the θ_K in (42),*

$$\|\theta_K - \theta_K^*\| = O\left(K^{-s/2}\right)$$

and

$$\sup_{v\in\mathcal{V}}|f_0(z) - f_K^*(z)| = O\left(\zeta(K)^2K^{-s/2}\right). \tag{48}$$

Proofs of lemmas and technical derivations are in Section I. Lemma F.1 establishes the distance between the true density and the pseudo true density.

Now the stochastic order of $\left\|\widehat{\theta}_K - \theta_K^*\right\|$ is derived. Define $\widehat{Q}_T(\theta) = \frac{1}{T}\sum_{t=1}^T \tau(q_t, r_t) \ln \widetilde{L}(f(\cdot, \theta); q_t, r_t)$ and $Q(\theta) = E[\tau(q_t, r_t) \ln \widetilde{L}(f(\cdot, \theta); q_t, r_t)]$. Then we have

$$\sup_{\theta \in \Theta_T} \left|\widehat{Q}_T(\theta) - Q(\theta)\right| = o_p\left(T^{-1/2+\alpha/2+\delta}\right) \tag{49}$$

for all sufficiently small $\delta > 0$ from Lemma I.6. Now for $\|\theta - \theta_K^*\| \leq o(\eta_T)$, we also have

$$\sup_{\left\|\theta - \theta_K^*\right\| \leq o(\eta_T), \theta \in \Theta_T} \left|\widehat{Q}_T(\theta) - \widehat{Q}_T(\theta_K^*) - (Q(\theta) - Q(\theta_K^*))\right| = o_p\left(\eta_T T^{-1/2+\alpha/2+\delta}\right) \tag{50}$$

as shown in Lemma I.7. From (49) and Lemma I.7, it follows that

Lemma F.2 *Suppose Assumption F.4 holds and* $\frac{\zeta(K)^2 K}{T} \to 0$*. Then for* $K(T) = O\left(T^{\alpha}\right)$*,*

$$\left\|\widehat{\theta}_K - \theta_K^*\right\| = o_p\left(T^{-1/2+\alpha/2+\delta}\right).$$

Thus we obtain

$$\begin{aligned} & \sup_{z \in \mathcal{V}} \left|\widehat{f}(z) - f_K^*(z)\right| \\ = \; & \sup_{z \in \mathcal{V}} \left|\left(W^K(z)'(\widehat{\theta}_K - \theta_K^*)\right)\left(W^K(z)'(\widehat{\theta}_K + \theta_K^*)\right)\right| \\ \leq \; & C_1 \left(\sup_{z \in \mathcal{V}} \left\|W^K(z)\right\|\right)^2 \left\|\widehat{\theta}_K - \theta_K^*\right\| \\ = \; & O\left(\zeta(K)^2\right) o_p\left(T^{-1/2+\alpha/2+\delta}\right) \end{aligned} \tag{51}$$

since $\|\theta\|^2 < 1$ for any $\theta \in \Theta_T$. Thus finally, we obtain

$$\begin{aligned} & \sup_{z \in \mathcal{V}} \left|\widehat{f}(z) - f_0(z)\right| \\ \leq \; & \sup_{z \in \mathcal{V}} \left|\widehat{f}(z) - f_K^*(z)\right| + \sup_{z \in \mathcal{V}} \left|f_K^*(z) - f_0(z)\right| \\ = \; & O\left(\zeta(K)^2\right) o_p\left(T^{-1/2+\alpha/2+\delta}\right) + O\left(\zeta(K)^2 K^{-s/2}\right). \end{aligned} \tag{52}$$

Thus we have proved Theorem F.1.

G Asymptotics for Test Statistic (Proof of Theorem E.1)

To prove Theorem E.1, we need the following lemma that establishes the conditions for $\widehat{I}_{\gamma}^{g_2} \underset{p}{\longrightarrow} I^{g_2}$. First, we let $E_h[\cdot]$ denote an expectation operator that takes expectation with respect to a density h.

Lemma G.1 *Suppose Assumption F.3 holds. Suppose (i) f_{01} and f_{02} are continuous, (ii) $E_{g_2}[|\ln f_{01}|] < \infty$ and $E_{g_2}[|\ln f_{02}|] < \infty$, (iii)$\frac{1}{T-1}\sum_{t=1}^{T-1}\Pr(t \notin \mathcal{T}_2) = o(1)$.Further suppose (iv)*

$$\frac{1}{T_\gamma - 1}\sum_{t\in\mathcal{T}_2} c_t(\gamma)\ln\left(\widehat{f}_1(\widetilde{v}_t)/f_{01}(\widetilde{v}_t)\right) \underset{p}{\longrightarrow} 0 \text{ and } \frac{1}{T_\gamma - 1}\sum_{t\in\mathcal{T}_2} c_t(\gamma)\ln\left(\widehat{f}_2(\widetilde{v}_{t+1})/f_{02}(\widetilde{v}_{t+1})\right) \underset{p}{\longrightarrow} 0,$$

then we have

$$\widehat{I}_{\gamma}^{g_2} \underset{p}{\longrightarrow} I^{g_2}.$$

Proof. Note

$$\begin{aligned}
& \frac{1}{T_\gamma - 1}\sum_{t\in\mathcal{T}_2} c_t(\gamma)\ln f_{01}(\widetilde{v}_t) \\
= & \frac{1}{T_\gamma - 1}\sum_{t \text{ odd}} (1+\gamma)\ln f_{01}(\widetilde{v}_t) + \frac{1}{T_\gamma - 1}\sum_{t \text{ even}} (1-\gamma)\ln f_{01}(\widetilde{v}_t) \\
& + O_p\left(\frac{1}{T-1}\sum_{t=1}^{T-1}\Pr(t\notin\mathcal{T}_2)\right) \\
= & \frac{1}{2}(1+\gamma)E_{g_2}[\ln f_{01}(\widetilde{v}_t)] + o_p(1) + \frac{1}{2}(1-\gamma)E_{g_2}[\ln f_{01}(\widetilde{v}_t)] + o_p(1) + o_p(1) \\
= & E_{g_2}[\ln f_{01}(\widetilde{v}_t)] + o_p(1)
\end{aligned}$$

by the law of large numbers under Condition (i) and by Condition (ii) and since $\{\widetilde{v}_t\}_{t=1}^T$ are iid. Similarly we have

$$\frac{1}{T_\gamma - 1}\sum_{t\in\mathcal{T}_2} c_t(\gamma)\ln f_{02}(\widetilde{v}_{t+1}) = E_{g_2}[\ln f_{02}(\widetilde{v}_{t+1})] + o_p(1)$$

and thus noting $\widehat{A}_2 \underset{p}{\longrightarrow} A_2$, by the Slutsky theorem

$$\widetilde{I}_{\gamma}^{g_2} \underset{p}{\longrightarrow} I^{g_2} \tag{53}$$

since $I^{g_2}(f_{01}, f_{02})$ can be expressed as $I^{g_2}(f_{01}, f_{02}) = (A_2 E_{g_2}[\ln f_{01}(\widetilde{v}_t) - \ln f_{02}(\widetilde{v}_t)])^2$. Condition (iv) implies

$$\widehat{I}_{\gamma}^{g_2} - \widetilde{I}_{\gamma}^{g_2} \underset{p}{\longrightarrow} 0 \tag{54}$$

applying the Slutsky theorem. From (53) and (54), we conclude

$$\widehat{I}_{\gamma}^{g_2} \underset{p}{\longrightarrow} I^{g_2}.$$

■

Now we prove Theorem E.1. We impose the following three conditions.

Condition 1 $\sum_{t\in\mathcal{T}_2} c_t(\gamma)\ln\left(\widehat{f}_1(\widetilde{v}_t)/f_{01}(\widetilde{v}_t)\right) = o_p\left(\sqrt{T}\right)$ *and* $\sum_{t\in\mathcal{T}_2} c_t(\gamma)\ln\left(\widehat{f}_2(\widetilde{v}_{t+1})/f_{02}(\widetilde{v}_{t+1})\right) = o_p\left(\sqrt{T}\right)$.

Condition 2 $\frac{1}{T-1}\sum_{t=1}^{T-1}\Pr(t\notin\mathcal{T}_2) = o\left(\frac{1}{\sqrt{T}}\right)$.

Condition 3 $E_{g_2}\left[|\ln f_{01}(\widetilde{v}_t)|^2\right] < \infty$ *and* $E_{g_2}\left[|\ln f_{02}(\widetilde{v}_t)|^2\right] < \infty$.

Proof. Considering the powers of the proposed test, we may want to achieve the largest possible order of $d(T)$ while letting $d(T)\widehat{I}_{\gamma}^{g_2}$ preserve the some limiting distributions under the null. In what follows, we show that we can achieve this with $d(T) = O(T)$. Suppose Condition 1 holds, then it follows immediately that

$$\widehat{I}_{\gamma}^{g_2} - \widetilde{I}_{\gamma}^{g_2} = o_p\left(\frac{1}{\sqrt{T}}\right)$$

for all $\gamma \geq 0$. This implies that the asymptotic distribution of $T\widehat{I}_{\gamma}^{g_2}$ will be identical to that of $T\widetilde{I}_{\gamma}^{g_2}$ under the null, which means the effect of nonparametric estimation is negligible. Now consider, under the null $f_{01} = f_{02}$

$$\begin{aligned} & \frac{1}{T_\gamma - 1}\sum_{t\in\mathcal{T}_2} c_t(\gamma)\left(\ln f_{01}(\widetilde{v}_t) - \ln f_{02}(\widetilde{v}_{t+1})\right) \qquad (55)\\ = \; & \frac{2\gamma}{T_\gamma - 1}\sum_{t\in\mathcal{Q}}\left(\ln f_{01}(\widetilde{v}_t) - \ln f_{01}(\widetilde{v}_{t+1})\right) + \frac{1+\gamma}{T_\gamma - 1}\ln f_{01}(\widetilde{v}_1) - \frac{c_{\max\mathcal{Q}+1}(\gamma)}{T_\gamma - 1}\ln f_{01}(\widetilde{v}_{\max\mathcal{Q}+1}) \\ & + O_p\left(\frac{1}{T-1}\sum_{t=1}^{T-1}\Pr(t\notin\mathcal{T}_2)\right) \end{aligned}$$

where $\mathcal{Q} = \{t : 1 \leq t \leq T-1,\, t \text{ even}\}$. First, we are willing to choose $\gamma > 0$ and look into what happens when $\gamma = 0$ afterwards. Now suppose Condition 1 holds. Further suppose Conditions 2-3 hold. Then, under the null of (7), we have

$$\frac{1}{\sqrt{T/2}}\sum_{t\in\mathcal{Q}}\left(\ln f_{01}(\widetilde{v}_{t+1}) - \ln f_{01}(\widetilde{v}_t)\right) \underset{d}{\longrightarrow} N(0,\Sigma) \text{ and } \frac{1}{\sqrt{T/2}}\sum_{t\in\mathcal{Q}}\left(\ln f_{02}(\widetilde{v}_{t+1}) - \ln f_{02}(\widetilde{v}_t)\right) \underset{d}{\longrightarrow} N(0,\Sigma) \qquad (56)$$

where $\Sigma = E_{g_2}\left[(\ln f_{01}(\widetilde{v}_{t+1}) - \ln f_{01}(\widetilde{v}_t))^2\right]$ or $E_{g_2}\left[(\ln f_{02}(\widetilde{v}_{t+1}) - \ln f_{02}(\widetilde{v}_t))^2\right]$ that equals to $2\mathrm{Var}_{g_2}[\ln f_{01}(\cdot)]$ or $2\mathrm{Var}_{g_2}[\ln f_{02}(\cdot)]$ by the Lindberg-Levy central limit theorem under the null. Therefore from (55) and (56), we conclude that under Conditions 1-3,

$$\begin{aligned} & \sqrt{T}\frac{1}{T_\gamma - 1}\sum_{t\in\mathcal{T}_2} c_t(\gamma)\left(\ln \widehat{f}_1(\widetilde{v}_t) - \ln \widehat{f}_2(\widetilde{v}_{t+1})\right) \qquad (57) \\ = \quad & \sqrt{T}\frac{1}{T_\gamma - 1}\sum_{t\in\mathcal{T}_2} c_t(\gamma)\left(\ln f_{01}(\widetilde{v}_t) - \ln f_{02}(\widetilde{v}_{t+1})\right) \underset{d}{\longrightarrow} N(0, 2\gamma^2\Sigma) \end{aligned}$$

under (7) for any $\gamma > 0$ noting $f_{01} = f_{02}$ under the null. Finally, for any $\widehat{\Sigma} = \Sigma + o_p(1)$, we conclude that

$$\frac{1}{\sqrt{2}\gamma\widehat{\Sigma}^{\frac{1}{2}}}\sqrt{T}\frac{1}{T_\gamma - 1}\sum_{t\in\mathcal{T}_2} c_t(\gamma)\left(\ln \widehat{f}_1(\widetilde{v}_t) - \ln \widehat{f}_2(\widetilde{v}_{t+1})\right) \underset{d}{\longrightarrow} N(0,1)$$

and hence

$$T\widehat{I}_\gamma^{g_2} \underset{d}{\longrightarrow} \chi^2(1)$$

with $A = \frac{1}{\sqrt{2}\gamma\Sigma^{\frac{1}{2}}}$ and $\widehat{A} = \frac{1}{\sqrt{2}\gamma\widehat{\Sigma}^{\frac{1}{2}}}$. Possible candidates of $\widehat{\Sigma}$ will be

$$\widehat{\Sigma}_1 = 2\left(\frac{1}{T}\sum_{t=1}^{T}(\ln \widehat{f}_1(\widetilde{v}_t))^2 - \left\{\frac{1}{T}\sum_{t=1}^{T}\ln \widehat{f}_1(\widetilde{v}_t)\right\}^2\right) \text{ or } \widehat{\Sigma}_2 = 2\left(\frac{1}{T}\sum_{t=1}^{T}(\ln \widehat{f}_2(\widetilde{v}_t))^2 - \left\{\frac{1}{T}\sum_{t=1}^{T}\ln \widehat{f}_2(\widetilde{v}_t)\right\}^2\right)$$

or its average $\widehat{\Sigma}_3 = \frac{\widehat{\Sigma}_1 + \widehat{\Sigma}_2}{2}$. All of these are consistent under Condition 3 and under Condition (iii) and (iv) of Lemma G.1 and (7). ■

H Primitive Conditions for the SNP Estimators

In this section, we show that all the conditions for Lemma G.1 and Theorem E.1 are satisfied for the SNP density estimators of (29) and (31). Here we should note that Lemma G.1 holds whether or not the null ($f_{01} = f_{02}$) is true while Theorem E.1 is required to hold only under the null. We start with conditions for Lemma G.1. First, note Condition (i) is directly assumed and Condition (ii) in Lemma G.1 immediately hold since f_{01} and f_{02} are continuous and $\mathcal{V}$ is compact. Condition (iii) of Lemma G.1 is verified as follows. For $\delta_1(T)$ and $\delta_2(T)$ that are positive numbers tending to zero as $T \to \infty$, consider

$$
\begin{aligned}
& \frac{1}{T-1}\sum_{t=1}^{T-1}\Pr(t\notin\mathcal{T}_2) \qquad (58)\\
\le\ & \frac{1}{T-1}\sum_{t=1}^{T-1}\Pr\left(\widehat{f}_1(\widetilde{v}_t)\le\delta_1(T)\text{ or }\widehat{f}_2(\widetilde{v}_{t+1})\le\delta_2(T)\right)\\
\le\ & \frac{1}{T-1}\sum_{t=1}^{T-1}\Pr\left(\left|\widehat{f}_1(\widetilde{v}_t)-f_{01}(\widetilde{v}_t)\right|+\delta_1(T)\ge f_{01}(\widetilde{v}_t)\text{ or }\left|\widehat{f}_2(\widetilde{v}_{t+1})-f_{02}(\widetilde{v}_{t+1})\right|+\delta_2(T)\ge f_{02}(\widetilde{v}_{t+1})\right)\\
\le\ & \frac{1}{T-1}\sum_{t=1}^{T-1}\Pr\left(\sup_{z\in\mathcal{V}}\left|\widehat{f}_1(z)-f_{01}(z)\right|+\delta_1(T)\ge f_{01}(\widetilde{v}_t)\text{ or }\sup_{z\in\mathcal{V}}\left|\widehat{f}_2(z)-f_{02}(z)\right|+\delta_2(T)\ge f_{02}(\widetilde{v}_{t+1})\right)
\end{aligned}
$$

and hence as long as $\sup_{z\in\mathcal{V}}\left|\widehat{f}_1(z)-f_{01}(z)\right|=o_p(1)$, $\sup_{z\in\mathcal{V}}\left|\widehat{f}_2(z)-f_{02}(z)\right|=o_p(1)$, $\delta_1(T)=o(1)$, and $\delta_2(T)=o(1)$, we have $\frac{1}{T-1}\sum_{t=1}^{T-1}\Pr(t\notin\mathcal{T}_2)=o_p(1)$ since f_{01} and f_{02} are bounded away from zero. Therefore, under $\alpha\le\frac{1}{3}-\frac{2}{3}\delta$ and $s>2$, the condition (iii) of Lemma G.1 holds from Theorem F.1 with a suitable choice of ϑ.

Condition (iv) of Lemma G.1 is easily established from the uniform convergence rate result. Using $|\ln(1+t)|\le 2|t|$ in a neighborhood of $t=0$, consider

$$
\begin{aligned}
& \left|\frac{1}{T_\gamma-1}\sum_{t\in\mathcal{T}_2}c_t(\gamma)\ln\left(\widehat{f}_1(\widetilde{v}_t)/f_{01}(\widetilde{v}_t)\right)\right| \qquad (59)\\
\le\ & (1+\gamma)\sup_{z\in\mathcal{V}}\left|\ln\widehat{f}_1(z)-\ln f_{01}(z)\right|\le(1+\gamma)\sup_{z\in\mathcal{V}}2\left|\frac{f_{01}(z)-\widehat{f}_1(z)}{\widehat{f}_1(z)}\right|\\
=\ & O\left(\zeta(K)^2\right)o_p\left(T^{-1/2+\alpha/2+\delta}\right)+O\left(\zeta(K)^2K^{-s/2}\right)
\end{aligned}
$$

from Theorem F.1 (with a suitable choice of ϑ) and since $\widehat{f}_1(\cdot)$ is bounded away from zero. Therefore, $\frac{1}{T_\gamma-1}\sum_{t\in\mathcal{T}_2}c_t(\gamma)\ln\left(\widehat{f}_1(\widetilde{v}_t)/f_{01}(\widetilde{v}_t)\right)=o_p(1)$ under $\alpha\le\frac{1}{3}-\frac{2}{3}\delta$ and $s>2$ with $K=O(T^\alpha)$ noting $\zeta(K)=O\left(\sqrt{K}\right)$. Similarly we can show $\frac{1}{T_\gamma-1}\sum_{t\in\mathcal{T}_2}c_t(\gamma)\ln(\widehat{f}_2(\widetilde{v}_{t+1})/f_{02}(\widetilde{v}_{t+1}))=o_p(1)$ under $\alpha\le\frac{1}{3}-\frac{2}{3}\delta$ and $s>2$.

Now we establish conditions for Theorem E.1. Again Condition 3 immediately holds since f_{01} and f_{02} are assumed to be continuous and $\mathcal{V}$ is compact. Next, we show Condition 2. >From (58) and the Markov inequality, we have

$$\begin{aligned}
& \frac{1}{T-1}\sum_{t=1}^{T-1}\Pr(t\notin\mathcal{T}_2) \qquad (60)\\
\le\ & \sup_{z\in\mathcal{V}}\left(\frac{1}{f_{01}(z)^\tau}\right)\frac{1}{T-1}\sum_{t=1}^{T-1}E_{g_2}\left[\left|\widehat{f}_1(\widetilde{v}_t)-f_{01}(\widetilde{v}_t)+\delta_1(T)\right|^\tau\right]\\
& +\sup_{z\in\mathcal{V}}\left(\frac{1}{f_{02}(z)^\tau}\right)\frac{1}{T-1}\sum_{t=1}^{T-1}E_{g_2}\left[\left|\widehat{f}_2(\widetilde{v}_{t+1})-f_{02}(\widetilde{v}_{t+1})+\delta_2(T)\right|^\tau\right]
\end{aligned}$$

and hence $\frac{1}{T-1}\sum_{t=1}^{T-1}\Pr(t\notin\mathcal{T}_2)=o_p(\frac{1}{\sqrt{T}})$ as long as

$$\sup_{z\in\mathcal{V}}\left|\widehat{f}_1(z)-f_{01}(z)\right|^\tau=o_p(\frac{1}{\sqrt{T}}),\ \sup_{z\in\mathcal{V}}\left|\widehat{f}_2(z)-f_{02}(z)\right|^\tau=o_p(\frac{1}{\sqrt{T}}), \qquad (61)$$

$\delta_1(T)^\tau=o(\frac{1}{\sqrt{T}})$, and $\delta_2(T)^\tau=o(\frac{1}{\sqrt{T}})$ noting $f_{01}(\cdot)$ and $f_{02}(\cdot)$ are bounded away from zero. Note

$$\begin{aligned}
\sup_{z\in\mathcal{V}}\left|\widehat{f}_1(z)-f_{01}(z)\right|^\tau &= o_p\left(T^{(-1/2+3\alpha/2+\delta)\tau}\right)+O\left(T^{(1-s/2)\alpha\tau}\right) \text{ and}\\
\sup_{z\in\mathcal{V}}\left|\widehat{f}_2(z)-f_{02}(z)\right|^\tau &= o_p\left(T^{(-1/2+3\alpha/2+\delta)\tau}\right)+O\left(T^{(1-s/2)\alpha\tau}\right)
\end{aligned}$$

from Theorem F.1 (with a suitable choice of ϑ) and $\zeta(K)=O\left(\sqrt{K}\right)$ letting $K=O\left(T^\alpha\right)$. In particular, we choose $\tau=4$ and hence (61) holds under $\alpha\le\frac{1}{4}-\frac{2}{3}\delta$ and $s\ge 2+\frac{1}{4\alpha}$. $\delta_1(T)^\tau=o(\frac{1}{\sqrt{T}})$ and $\delta_2(T)^\tau=o(\frac{1}{\sqrt{T}})$ hold under $\delta_1(T)=o(T^{-\frac{1}{8}})$ and $\delta_2(T)=o(T^{-\frac{1}{8}})$. Therefore, under $\alpha\le\frac{1}{4}-\frac{2}{3}\delta$, $s\ge 2+\frac{1}{4\alpha}$, $\delta_1(T)=o(T^{-\frac{1}{8}})$, and $\delta_2(T)=o(T^{-\frac{1}{8}})$, we finally have $\frac{1}{T-1}\sum_{t=1}^{T-1}\Pr(t\notin\mathcal{T}_2)=o\left(\frac{1}{\sqrt{T}}\right)$. Condition 3 remains to be verified. We conjecture that we can verify this condition similarly with Kim (2005).

I Technical Lemmas and Proofs

I.1 Bound of the Truncated Hermite Series

Lemma I.1 *Suppose $W^K(v)$ is given by (28). Then,* $\sup_{v\in\mathcal{V}}\left\|W^K(v)\right\|=\zeta(K)=O(\sqrt{K})$.

Proof. Without loss of generality, we can impose $\mathcal{V}$ to be symmetric around zero such that $\underline{v}+\overline{v}=0$. Then, this lemma follows from Kim (2005). ■

I.2 Proof of Lemma F.1

Proof. In the following proof, we treat $\epsilon_0=0$ to simplify discussions since we can pick ϵ_0 arbitrary small but still we need to impose that $f(\cdot,\theta)\in\underline{\mathcal{F}}_T$ is bounded away from zero.

Now define $Q_0 = E\left[\tau(q_t,r_t)\ln\widetilde{L}(f_0;q_t,r_t)\right]$ where $\widetilde{L}(f_0;q_t,r_t) = \frac{(1-F_0(q_t))^{k_1-1}f_0(q_t)}{(1-F_0(r_t))^{k_2-1}}$ with $F_0(z) = \int_{\underline{v}}^{z} f_0(z)dz$ and $\tau(q_t,r_t)$ is a trimming device (an indicator function) that excludes those values $q_t, r_t > \overline{v}-\varepsilon$ and $q_t, r_t < \underline{v}+\varepsilon$. We denote $\mathcal{V}_\varepsilon = [\underline{v}+\varepsilon, \overline{v}-\varepsilon]$. Also define $Q(\theta) = E\left[\tau(q_t,r_t)\ln\widetilde{L}(f(\cdot,\theta);q_t,r_t)\right]$. Then, by definition, $\theta_K^* = \underset{\theta}{\operatorname{argmax}}\, Q(\theta)$. Consider

$$\underset{\theta}{\operatorname{argmin}}\, Q_0 - Q(\theta) = \underset{\theta}{\operatorname{argmax}}\, Q(\theta)$$

which implies that among the parametric family $\{f(z,\theta) : \theta = (\vartheta_1,\ldots,\vartheta_K)\}$, $Q(\theta_K^*)$ will have the minimum distance to Q_0 noting for all $\theta \in \Theta_T$, $Q(\theta) \le Q_0$ from the information inequality (see Gallant (1987, p.484)). First, we show that

$$Q_0 - Q(\theta_K) = O(K^{-s}). \tag{62}$$

Define $F(z,\theta_K) = \int_{\underline{v}}^{z} f(z,\theta_K)dz$ and consider

$$\begin{aligned}
|Q_0 - Q(\theta_K)| &\le E\left[\tau(q_t,r_t)\left|\ln\frac{\widetilde{L}(f_0;q_t,r_t)}{\widetilde{L}(f(\cdot,\theta);q_t,r_t)}\right|\right] && (63)\\
&\le E\left[\tau(q_t,r_t)\left|\ln\frac{f_0(q_t)}{f(q_t,\theta_K)}\right|\right] && (64)\\
&\quad + E\left[(k_1-1)\,\tau(q_t,r_t)\left|\ln\frac{1-F_0(q_t)}{1-F(q_t,\theta_K)}\right|\right] && (65)\\
&\quad + E\left[(k_2-1)\,\tau(q_t,r_t)\left|\ln\frac{1-F_0(r_t)}{1-F(r_t,\theta_K)}\right|\right] && (66)
\end{aligned}$$

by the triangular inequality. We bound (64), (65), and (66) in turns.

(i) Bound of (64)

Now denoting a random variable Z to follow the distribution with the density f_0 with the support $\mathcal{V}$ and using $|\ln(1+t)| \le 2|t|$ in a neighborhood of $t=0$, consider

$$\begin{aligned}
& E_{g_{k_1}}\left[\tau(q_t,r_t)\left|\ln\frac{f_0(q_t)}{f(q_t,\theta_K)}\right|\right]\\
\le\ & E_{g_{k_1}}\left[2\left|\frac{f_0(q_t)}{f(q_t,\theta_K)}-1\right|\right]\\
=\ & 2\int_{\mathcal{V}}\frac{1}{f(z,\theta_K)}\left|f_0(z)-f(z,\theta_K)\right| g^{(n-k_1+1:n)}(z)dz\\
=\ & 2\int_{\mathcal{V}}\frac{g^{(n-k_1+1:n)}(z)}{f(z,\theta_K)}\frac{\sqrt{f_0(z)}}{\sqrt{f_0(z)}}\left|h_{f_0}(z)e^{-z^2/4}+W^K(z)'\theta_K\right|\left|h_{f_0}(z)e^{-z^2/4}-W^K(z)'\theta_K\right|dz\\
\le\ & 2\sup_{z\in\mathcal{V}}\frac{g^{(n-k_1+1:n)}(z)\sqrt{f_0(z)}}{f(z,\theta_K)}\sup_{z\in\mathcal{V}}\left|h_{f_0}(z)e^{-z^2/4}-W^K(z)'\theta_K\right| E_{f_0}\left[\frac{\left|h_{f_0}(Z)e^{-Z^2/4}\right|+\left|W^K(Z)'\theta_K\right|}{\sqrt{f_0(Z)}}\right].
\end{aligned}$$

Note

$$E_{f_0}\left[\frac{\left|h_{f_0}(Z)e^{-Z^2/4}\right|}{\sqrt{f_0(Z)}}\right] \leq \sqrt{E_{f_0}\left[\frac{h_{f_0}^2(Z)e^{-Z^2/2}}{f_0(Z)}\right]} < 1 \tag{67}$$

since $0 < \frac{h_{f_0}^2(z)}{f_0(z)} < 1$ for all $z \in \mathcal{V}$ by construction and note

$$E_{f_0}\left[\frac{\left|W^K(Z)'\theta_K\right|}{\sqrt{f_0(Z)}}\right] \leq \sqrt{E\left[\frac{\theta_K' W^K(Z)W^K(Z)'\theta_K}{f_0(Z)}\right]} = \sqrt{\theta_K'\theta_K} = \|\theta_K\| < 1 \tag{68}$$

since $\|\theta_K\|^2 < 1$. Also note $\sup_{z\in\mathcal{V}} \frac{g^{(n-k_1+1:n)}(z)\sqrt{f_0(z)}}{f(z,\theta_K)} < \infty$ since $g^{(n-k_1+1:n)}(\cdot)$ and $f_0(\cdot)$ are bounded from above and since $f(z,\theta_K)$ is bounded away from zero. Thus, from (43):$\sup_{z\in\mathcal{V}}\left|h_{f_0}(z)e^{-z^2/4} - W^K(z)'\right.$ $O\left(K^{-s}\right)$, we have

$$E\left[\left|\ln\frac{f_0(q_t)}{f(q_t,\theta_K)}\right|\right] = O\left(K^{-s}\right) \text{ and similarly } E\left[\left|\ln\frac{f_0(r_t)}{f(r_t,\theta_K)}\right|\right] = O\left(K^{-s}\right).$$

(iii) Bound of (65) and (66)

For some $0 < \alpha < 1$, note

$$\begin{aligned}
& E_{g_{k_1}}\left[\left|F_0(q_t) - F(q_t,\theta_K)\right|\right] \\
= \;& E_{g_{k_1}}\left[\left|\int_{\underline{v}}^{q_t} f_0(z) - f(z,\theta_K)dz\right|\right] \\
\leq \;& E_{g_{k_1}}\left[(q_t - \underline{v})\left|f_0(\alpha q_t + (1-\alpha)\underline{v}) - f(\alpha q_t + (1-\alpha)\underline{v},\theta_K)\right|\right] \\
\leq \;& (\overline{v} - \underline{v})\, E_{g_{k_1}}\left[\left|f_0(\alpha q_t + (1-\alpha)\underline{v}) - f(\alpha q_t + (1-\alpha)\underline{v},\theta_K)\right|\right]
\end{aligned}$$

where the last inequality is from $\overline{v} > q_t > \underline{v}$. Using the change of variable $z = \alpha q_t + (1-\alpha)\underline{v}$, note

$$\begin{aligned}
& E_{g_{k_1}}\left[\left|f_0(\alpha q_t + (1-\alpha)\underline{v}) - f(\alpha q_t + (1-\alpha)\underline{v},\theta_K)\right|\right] \\
\leq \;& \sup_{z\in\mathcal{V}}\left|h_{f_0}(z)e^{-z^2/4} - W^K(z)'\theta_K\right| \int_{\underline{v}}^{\overline{v}-(1-\alpha)(\overline{v}-\underline{v})} \frac{1}{\alpha} g^{(n-k_1+1:n)}(z)\sqrt{f_0(z)}\frac{\left|h_{f_0}(z)e^{-z^2/4} + W^K(z)'\theta_K\right|}{\sqrt{f_0(z)}}dz \\
\leq \;& \sup_{z\in\mathcal{V}}\left|h_{f_0}(z)e^{-z^2/4} - W^K(z)'\theta_K\right| \int_{\underline{v}}^{\overline{v}} \frac{1}{\alpha} g^{(n-k_1+1:n)}(z)\sqrt{f_0(z)}\frac{\left|h_{f_0}(z)e^{-z^2/4} + W^K(z)'\theta_K\right|}{\sqrt{f_0(z)}}dz
\end{aligned}$$

where the second inequality is from $(1-\alpha)(\overline{v} - \underline{v}) > 0$. Note

$$\begin{aligned}
& \int_{\underline{v}}^{\overline{v}} \frac{1}{\alpha} g^{(n-k_1+1:n)}(z)\sqrt{f_0(z)}\frac{\left|h_{f_0}(z)e^{-z^2/4} + W^K(z)'\theta_K\right|}{\sqrt{f_0(z)}}dz \\
< \;& \sup_{z\in\mathcal{V}} \frac{1}{\alpha} g^{(n-k_1+1:n)}(z)\sqrt{f_0(z)} E_{f_0}\left[\frac{\left|h_{f_0}(Z)e^{-Z^2/4}\right| + \left|W^K(Z)'\theta_K\right|}{\sqrt{f_0(Z)}}\right] < \infty
\end{aligned}$$

by (67) and (68) and hence

$$E_{g_{k_1}}[|F_0(q_t) - F(q_t, \theta_K)|] = O(K^{-s}). \tag{69}$$

Similarly we can obtain

$$E_{g_{k_2}}[|F_0(r_t) - F(r_t, \theta_K)|] = O(K^{-s}) \tag{70}$$

and thus $E\left[\left|\ln \frac{F_0(q_t)-F_0(r_t)}{F(q_t,\theta_K)-F(r_t,\theta_K)}\right|\right] = O(K^{-s})$.

Now using $|\ln(1+t)| \le 2\,|t|$ in a neighborhood of $t = 0$, note

$$\begin{aligned} E\left[\tau(q_t, r_t)\left|\ln \frac{1-F_0(q_t)}{1-F(q_t,\theta_K)}\right|\right] &\le 2E\left[\tau(q_t, r_t)\left|\frac{F(q_t,\theta_K)-F_0(q_t)}{1-F(q_t,\theta_K)}\right|\right] \\ &\le CE_{g_{k_1}}[\tau(q_t, r_t)\,|F_0(q_t) - F(q_t, \theta_K)|] \end{aligned}$$

since $F(q_t, \theta_K) < 1$ and hence from (69), we conclude (65)$= O(K^{-s})$. Similarly, from (70), we can show that (66)$= O(K^{-s})$.

>From these results (i)-(iii), we conclude $|Q_0 - Q(\theta_K)| = O\left(K^{-s}\right)$. It follows that

$$0 \le Q(\theta_K^*) - Q(\theta_K) \le Q(\theta_K^*) - Q_0 + C_1 K^{-s} \le C_1 K^{-s} \tag{71}$$

where the first inequality is by definition of $\theta_K^* = \underset{\theta}{\operatorname{argmax}}\, Q(\theta)$, the second inequality is by (62), and the last inequality is since $Q(\theta_K^*) \le Q_0$ from the information inequality (see Gallant (1987, p.484)). Using the second order Taylor expansion where $\widetilde{\theta}$ lies between a given $\theta^o \in \Theta_T$ with $\dim(\theta^o) = K$ and θ_K^*, we have

$$\begin{aligned} & Q(\theta_K^*) - Q(\theta^o) \qquad (72) \\ = & -\frac{\partial Q}{\partial \theta'}(\theta_K^*)(\theta^o - \theta_K^*) - \frac{1}{2}(\theta^o - \theta_K^*)'\frac{\partial^2 Q(\widetilde{\theta})}{\partial\theta\partial\theta'}(\theta^o - \theta_K^*) \\ = & -\frac{1}{2}(\theta^o - \theta_K^*)'\frac{\partial^2 Q(\widetilde{\theta})}{\partial\theta\partial\theta'}(\theta^o - \theta_K^*) \end{aligned}$$

since $\frac{\partial Q}{\partial \theta}(\theta_K^*) = 0$ by F.O.C of (47).

$$\Gamma_3(\theta) \equiv E\left[\tau(q_t, r_t)\frac{1}{f(q_t,\theta)}\frac{\partial^2 f(q_t,\theta)}{\partial\theta\partial\theta'} - \tau(q_t, r_t)\frac{1}{f(q_t,\theta)^2}\frac{\partial f(q_t,\theta)}{\partial\theta}\frac{\partial f(q_t,\theta)}{\partial\theta'}\right].$$

Note $\frac{\partial^2 f(\cdot,\theta)}{\partial\theta\partial\theta'} = 2W^K(\cdot)W^K(\cdot)'$ and $\frac{\partial f(\cdot,\theta)}{\partial\theta}\frac{\partial f(\cdot,\theta)}{\partial\theta'} = 4\left(W^K(\cdot)'\theta\right)^2 W^K(\cdot)W^K(\cdot)' = 4f(\cdot,\theta)W^K(\cdot)W^K(\cdot)'$.

Consider

$$-\frac{1}{2}\Gamma_3(\theta) = E\left[\tau(q_t,r_t)\left(\frac{2\left(W^K(q_t)'\widetilde{\theta}\right)^2 W^K(q_t)W^K(q_t)'}{f(q_t,\theta)^2}-\frac{W^K(q_t)W^K(q_t)'}{f(q_t,\theta)}\right)\right] \tag{73}$$
$$= E\left[\tau(q_t,r_t)\frac{W^K(q_t)W^K(q_t)'}{f(q_t,\theta)}\right]$$
$$= \Pr\left(r_t\in\mathcal{V}_\varepsilon\right)\int_{\mathcal{V}_\varepsilon}\frac{g^{(n-k_1+1:n)}(z)}{f(z,\theta)}W^K(z)W^K(z)'dz$$
$$\geq \Pr\left(r_t\in\mathcal{V}_\varepsilon\right)\inf_{z\in\mathcal{V}_\varepsilon}\frac{g^{(n-k_1+1:n)}(z)}{f_0(z)}\inf_{z\in\mathcal{V}}\frac{f_0(z)}{f(z,\theta)}\int_{\mathcal{V}}W^K(z)W^K(z)'dz$$

where the last inequality comes form assuming ε is sufficiently small. Now note

$$\begin{aligned}&\left|f_0(z)-f(z,\widetilde{\theta})\right|\\ \leq\ &|f_0(z)-f(z,\theta_K)|+\left|f(z,\theta_K)-f(z,\widetilde{\theta})\right|\\ \leq\ &O\left(\zeta(K)K^{-s}\right)+O\left(\zeta(K)^2\right)\left(\|\theta^o-\theta_K^*\|+\|\theta_K^*-\theta_K\|\right)\end{aligned}$$

from (46) and hence (noting $f_0(z)$ is bounded away from zero and bounded from above)

$$\begin{aligned}\frac{f_0(z)}{f(z,\widetilde{\theta})} &= \frac{f_0(z)}{f(z,\widetilde{\theta})-f_0(z)+f_0(z)}\geq\frac{f_0(z)}{\sup_{z\in\mathcal{V}}\left|f(z,\widetilde{\theta})-f_0(z)\right|+f_0(z)}\\ &\geq \frac{f_0(z)}{O\left(\zeta(K)K^{-s}\right)+O\left(\zeta(K)^2\right)\left(\|\theta^o-\theta_K^*\|+\|\theta_K^*-\theta_K\|\right)+f_0(z)}.\end{aligned}$$

Noting $\inf_{z\in\mathcal{V}_\varepsilon}\frac{g^{(n-k_1+1:n)}(z)}{f_0(z)}$ and $\Pr\left(r_t\in\mathcal{V}_\varepsilon\right)$ are bounded away from zero, it follows that

$$\begin{aligned}&-\frac{1}{2}\Gamma_3(\theta)\\ \geq\ &C_2\inf_{z\in\mathcal{V}}\left(\frac{f_0(z)}{O\left(\zeta(K)K^{-s}\right)+O\left(\zeta(K)^2\right)\left(\|\theta^o-\theta_K^*\|+\|\theta_K^*-\theta_K\|\right)+f_0(z)}\right)I_K.\end{aligned} \tag{74}$$

>From this, putting $\theta^o=\theta_K$, we note

$$\begin{aligned}-\frac{1}{2}\Gamma_3(\theta) &\geq C_2(1-o(1))\text{ if }\|\theta_K^*-\theta_K\|=o\left(\zeta(K)^{-2}\right)\\ -\frac{1}{2}\Gamma_3(\theta) &\geq C_3\frac{1}{O\left(\zeta(K)^2\right)\left(\|\theta_K-\theta_K^*\|\right)}\text{ otherwise.}\end{aligned} \tag{75}$$

>From (75), we conclude

$$\begin{aligned}\lambda_{\min}\left(-\frac{1}{2}\frac{\partial^2Q(\widetilde{\theta})}{\partial\theta\partial\theta'}\right) &\geq C_4(1-o(1))\text{ if }\|\theta_K^*-\theta_K\|=o\left(\zeta(K)^{-2}\right)\\ \lambda_{\min}\left(-\frac{1}{2}\frac{\partial^2Q(\widetilde{\theta})}{\partial\theta\partial\theta'}\right) &\geq C_5\frac{1}{O\left(\zeta(K)^2\right)\left(\|\theta_K-\theta_K^*\|\right)}\text{ otherwise}\end{aligned}$$

since $\lambda_{\min}(I_K) = 1$. Thus, from $Q(\theta_K^*) - Q(\theta_K) \geq \lambda_{\min}\left(-\frac{1}{2}\frac{\partial^2 Q(\widetilde{\theta})}{\partial\theta\partial\theta'}\right)\|\theta_K - \theta_K^*\|^2$, it follows that

$$\begin{aligned}
Q(\theta_K^*) - Q(\theta_K) &\geq C_4 \|\theta_K - \theta_K^*\|^2 \text{ if } \|\theta_K^* - \theta_K\| = o\left(\zeta(K)^{-2}\right) \\
Q(\theta_K^*) - Q(\theta_K) &\geq O\left(\zeta(K)^{-2}\right)\|\theta_K - \theta_K^*\| \geq O\left(\zeta(K)^{-4}\right) \text{ otherwise.}
\end{aligned} \tag{76}$$

However, the case of (76) contradicts to (71) if $s > 2$, which means (71) implies $\|\theta_K^* - \theta_K\| = o\left(\zeta(K)^{-2}\right)$ under $s > 2$ and hence

$$Q(\theta_K^*) - Q(\theta_K) \geq C_4 \|\theta_K - \theta_K^*\|^2. \tag{77}$$

Together with (71), it implies $C_1 K^{-s} \geq Q(\theta_K^*) - Q(\theta_K) \geq C_4 \|\theta_K - \theta_K^*\|^2$ and hence under $s > 2$

$$\|\theta_K - \theta_K^*\| = O\left(K^{-s/2}\right) \tag{78}$$

as claimed in Lemma F.1. Finally note $\|\theta_K - \theta_K^*\| = o\left(\zeta(K)^{-2}\right)$ as long as (78) holds under $s > 2$.

Now consider

$$\begin{aligned}
& \sup_{z\in\mathcal{V}} |f(z,\theta_K) - f(z,\theta_K^*)| \\
\leq\ & \sup_{z\in\mathcal{V}} \left\|\left(W^K(z)'\theta_K\right)^2 - \left(W^K(z)'\theta_K^*\right)^2\right\| \\
\leq\ & \sup_{z\in\mathcal{V}} \left\|W^K(z)'\theta_K - W^K(z)'\theta_K^*\right\| \sup_{z\in\mathcal{V}} \left\|W^K(z)'\theta_K + W^K(z)'\theta_K^*\right\| \\
\leq\ & \sup_{z\in\mathcal{V}} \left\|W^K(z)\right\| \|\theta_K - \theta_K^*\| \sup_{z\in\mathcal{V}} \left\|W^K(z)\right\| (\|\theta_K\| + \|\theta_K^*\|) \\
=\ & O\left(\zeta(K)^2 K^{-s/2}\right)
\end{aligned}$$

from the Cauchy-Schwarz inequality, (78), $\sup_{z\in\mathcal{V}} \left\|W^K(z)\right\| \leq \zeta(K)$, and $\|\theta\|^2 < 1$ for any $\theta \in \Theta_n$. It follows that

$$\begin{aligned}
& \sup_{z\in\mathcal{V}} |f_0(z) - f_K^*(z)| \\
\leq\ & \sup_{z\in\mathcal{V}} |f_0(z) - f(z,\theta_K)| + \sup_{z\in\mathcal{V}} |f(z,\theta_K) - f(z,\theta_K^*)| \\
\leq\ & O\left(\zeta(K)K^{-s}\right) + O\left(\zeta(K)^2 K^{-s/2}\right) = O\left(\zeta(K)^2 K^{-s/2}\right).
\end{aligned}$$

■

I.3 Uniform Law of Large Numbers

Note that we have defined

$$\widehat{Q}_T(\theta) = \frac{1}{T}\sum_{t=1}^{T} \tau(q_t, r_t) \ln \widetilde{L}(f(\cdot,\theta); q_t, r_t) \text{ and } Q(\theta) = E[\tau(q_t, r_t) \ln \widetilde{L}(f(\cdot,\theta); q_t, r_t)].$$

Here we establish a uniform convergence with rate as

$$\sup_{\theta\in\Theta_T}\left|\widehat{Q}_T(\theta)-Q(\theta)\right|=o_p\left(T^{-1/2+\alpha/2+\delta}\right)$$

following Lemma 2 in Fenton and Gallant (1996).

Lemma I.2 (Lemma 2 in Fenton and Gallant (1996)) *Let $\{\Theta_T\}$ be a sequence of compact subsets of a metric space (Θ,ρ). Let $\{s_{Tt}(\theta):\theta\in\Theta;\ t=1,\ldots,T;\ T=1,\ldots\}$ be a sequence of real valued random variables defined over a complete probability space $(\Omega,\mathcal{A},P)$. Suppose that there are sequences of positive numbers $\{d_T\}$ and $\{M_T\}$ such that for each θ^o in Θ_T and for all θ in $\eta_T(\theta^o)=\{\theta\in\Theta_T:\rho(\theta,\theta^o)<d_T\}$, we have $|s_{Tt}(\theta)-s_{Tt}(\theta^o)|\leq\frac{1}{T}M_T\rho(\theta,\theta^o)$. Let $G_T(\tau)$ be the smallest number of open balls of radius τ necessary to cover Θ_T. If $\sup\limits_{\theta\in\Theta_T}P\left\{\left|\sum_{t=1}^T\left(s_{Tt}(\theta)-E\left[s_{Tt}(\theta)\right]\right)\right|>\epsilon\right\}\leq\Gamma_T(\epsilon)$, then for all sufficiently small $\varepsilon>0$ and all sufficiently large T,*

$$P\left\{\sup_{\theta\in\Theta_T}\left|\sum_{t=1}^T\left(s_{Tt}(\theta)-E\left[s_{Tt}(\theta)\right]\right)\right|>\varepsilon M_Td_T\right\}\leq G_T\left(\frac{\varepsilon d_T}{3}\right)\Gamma_T\left(\frac{\varepsilon M_Td_T}{3}\right).$$

First denote a set τ^c that contains (q_t,r_t)'s that survive the trimming device $\tau(\cdot)$. Now define $s_{Tt}(\theta)=\frac{1}{T}\tau(q_t,r_t)\ln\widetilde{L}(f(\cdot,\theta);q_t,r_t)$. Then, we have $\widehat{Q}_T(\theta)=\sum_{t=1}^T s_{Tt}(\theta)$ and $Q(\theta)=\sum_{t=1}^T E\left[s_{Tt}(\theta)\right]$. To entertain Lemma (I.2), in what follows, three conditions are verified.

Lemma I.3 *Suppose Assumption F.4 holds. Then, $|s_{Tt}(\theta)-s_{Tt}(\theta^o)|\leq C\frac{1}{T}\zeta(K(T))^2\left\|\theta-\theta^o\right\|$.*

Proof. First note that

$$\widetilde{L}(f(\cdot,\theta);q_t,r_t)=\tau(q_t,r_t)\frac{(F(q_t,\theta)-F(r_t,\theta))^{k_2-k_1-1}(1-F(q_t,\theta))^{k_1-1}f(q_t,\theta)}{(1-F(r_t,\theta))^{k_2-1}}$$

where we denote $F(z,\theta)=\int_{\underline{v}}^z f(z,\theta)dz$.

Now note if $0<c\leq a\leq b$, then $|\ln a-\ln b|\leq|a-b|/c$. Since $f(z,\theta)$ is bounded away from zero for all $\theta\in\Theta_T$ and $z\in\mathcal{V}$, $F(r_t,\theta)$ and $F(q_t,\theta)$ are bounded away from one (since $\max\limits_t r_t<\max\limits_t q_t\leq\overline{v}-\epsilon$ for $q_t,r_t\in\tau^c$), and $q_t>r_t$ for all t, we have $0<C\leq\widetilde{L}(f(\cdot,\theta);q_t,r_t)$ for all $q_t,r_t\in\tau^c$. It follows that

$$|s_{Tt}(\theta)-s_{Tt}(\theta^o)|\leq\left|\widetilde{L}(f(\cdot,\theta);q_t,r_t)-\widetilde{L}(f(\cdot,\theta^o);q_t,r_t)\right|/nC.$$

Consider

$$\begin{aligned}
&\left|\widetilde{L}(f(\cdot,\theta);q_t,r_t)-\widetilde{L}(f(\cdot,\theta^o);q_t,r_t)\right| \\
\leq\ & C_1\left|f(q_t,\theta)-f(q_t,\theta^o)\right| \\
=\ & C_1\left|W^K(q_t)'\theta+W^K(q_t)'\theta^o\right|\left|W^K(q_t)'\theta-W^K(q_t)'\theta^o\right| \\
\leq\ & C_1\sup_{z\in\mathcal{V}}\left\|W^K(z)\right\|(\|\theta\|+\|\theta^o\|)\sup_{z\in\mathcal{V}}\left\|W^K(z)\right\|\|\theta-\theta^o\| \\
\leq\ & C_2\zeta(K)^2\|\theta-\theta^o\|
\end{aligned}$$

where the first inequality is obtained since $F(r_t,\theta)$ is bounded above from one, $0<F(q_t,\theta)-F(r_t,\theta)<1$, and $0<F(q_t,\theta)<1$ for all $t, r_t \in \tau^c$. The last inequality is obtained from $\|\theta\|<1$ for all $\theta\in\Theta_T$ and $\sup_{z\in\mathcal{V}}\left\|W^K(z)\right\|=\zeta(K)$. It follows that

$$\begin{aligned}
&|s_{Tt}(\theta)-s_{Tt}(\theta^o)| \\
\leq\ & C_2\zeta(K)^2\|\theta-\theta^o\|/T.
\end{aligned}$$

■

Lemma I.4 *Suppose Assumption F.4 holds and* $\zeta(K)=O\left(K^{\zeta}\right)$. *Then,*

$$\Pr\left\{\left|\sum_{t=1}^{T}\left(s_{Tt}(\theta)-E\left[s_{Tt}(\theta)\right]\right)\right|>\epsilon\right\}\leq 2\exp\left(\frac{-2\epsilon^2}{T\left(\frac{1}{T}2\zeta\ln K(T)+\frac{1}{T}C\right)^2}\right).$$

Proof. We have $0<C_1\leq f(z,\theta)<C_2K^{2\zeta}+\epsilon_0\frac{\phi(0)}{\int_{\mathcal{V}}\phi(z)dz}$ by construction and since $f(z,\theta)$ is bounded away from zero. Moreover $0<F(q_t,\theta)<1$, $0<F(r_t,\theta)<1$, and $0<F(q_t,\theta)-F(r_t,\theta)<1$ for all $(q_t,r_t)\in\tau^c$. Thus it follows that $\frac{1}{T}C_3<s_{Tt}(\theta)<\frac{1}{T}2\zeta\ln K+\frac{1}{T}C_4$ for sufficiently large K and for all $(q_t,r_t)\in\tau^c$. Hoeffding's (1963) inequality implies that $\Pr\left(|Y_1+\ldots+Y_T|\geq\epsilon\right)\leq 2\exp\left(-2\epsilon^2/\sum_{t=1}^T(b_t-a_t)^2\right)$ for independent random variables centered zero with ranges $a_t\leq Y_t\leq b_t$. Applying this inequality, we have

$$\Pr\left\{\left|\sum_{t=1}^{T}\left(s_{Tt}(\theta)-E\left[s_{Tt}(\theta)\right]\right)\right|>\epsilon\right\}\leq 2\exp\left(-2\epsilon^2/T\left(\frac{1}{T}2\zeta\ln K(T)+\frac{1}{T}C\right)^2\right).$$

■

Lemma I.5 (Lemma 6 in Fenton and Gallant (1996)) *The number of open balls of radius* δ *required to cover* Θ_T *is bounded by* $2K(T)(2/\delta+1)^{K(T)-1}$.

Proof. *Lemma 1 of Gallant and Souza (1991) shows that the number of radius-*δ *balls needed to cover the surface of a unit sphere in* $\mathcal{R}^p$ *is bounded by* $2p(2/\delta+1)^{p-1}$. *Noting* $\dim(\Theta_T)=K(T)$, *the result follows immediately.* ■

Applying the results of Lemma I.3-I.5, finally we obtain

Lemma I.6 *Let $K(T) = C \cdot T^{\alpha}$ with $0 < \alpha < 1$ and suppose Assumption F.4 holds. Then,*

$$\sup_{\theta \in \Theta_T} \left| \widehat{Q}_T(\theta) - Q(\theta) \right| = o_p\left(T^{-1/2+\alpha/2+\delta}\right).$$

Proof. Let $M_T = C_1 O\left(K^{2\zeta}\right) = C_2 T^{2\zeta\alpha}$, $d_n = \frac{1}{C_1} T^{-(2\zeta-1)\alpha-\beta}$, and $\rho(\theta, \theta^o) = \|\theta - \theta^o\|$. Then from Lemma I.2, we have

$$\begin{aligned}
& \Pr\left\{\sup_{\theta \in \Theta_T} \left| \sum_{t=1}^{T} \left(s_{Tt}(\theta) - E\left[s_{Tt}(\theta)\right]\right) \right| > \varepsilon T^{\alpha-\beta} \right\} \\
\leq \quad & 4C \cdot T^{\alpha} \left(\frac{6C_1}{\varepsilon} T^{(2\zeta-1)\alpha+\beta} + 1\right)^{T^{\alpha}-1} \exp\left(-2\left(\frac{\varepsilon T^{\alpha-\beta}}{3}\right)^2 / T \left(\frac{1}{T} 2\zeta \ln K(T) + \frac{1}{T} C_2\right)^2\right)
\end{aligned}$$

applying Lemma I.3, Lemma I.4, and Lemma I.5.

Note $4C \cdot T^{\alpha}\left(\frac{6C_1}{\varepsilon} T^{(2\zeta-1)\alpha+\beta} + 1\right)^{T^{\alpha}-1}$ is dominated by $C_3 T^{\alpha} T^{((2\zeta-1)\alpha+\beta)(T^{\alpha}-1)}$ for sufficiently large T and note $T\left(\frac{1}{T} 2\zeta \ln K(T) + \frac{1}{T} C_2\right)^2$ is dominated by $T\left(\frac{1}{T} 2\zeta \ln K(T)\right)^2$. Thus, we simplify

$$\begin{aligned}
& \Pr\left\{\sup_{\theta \in \Theta_T} \left| \sum_{t=1}^{T} \left(s_{Tt}(\theta) - E\left[s_{Tt}(\theta)\right]\right) \right| > \varepsilon T^{\alpha-\beta} \right\} \\
\leq \quad & C_4 \exp\left(\ln\left(T^{\alpha} T^{((2\zeta-1)\alpha+\beta)(T^{\alpha}-1)}\right) - \frac{2\varepsilon^2}{9} T^{2\alpha-2\beta+1} / \left(2\zeta \ln K(T)\right)^2\right) \\
= \quad & C_4 \exp\left(\alpha \ln T + \left((2\zeta-1)\alpha + \beta\right)\left(T^{\alpha} - 1\right) \ln T - \frac{2\varepsilon^2}{9} T^{2\alpha-2\beta+1} / \left(2\zeta \ln K(T)\right)^2\right)
\end{aligned}$$

for sufficiently large T. As long as $2\alpha - 2\beta + 1 > \alpha$, $\frac{2\varepsilon^2}{9} T^{2\alpha-2\beta+1} / \left(2\zeta \ln K(T)\right)^2$ dominates $\alpha \ln T + \left((2\zeta-1)\alpha + \beta\right)\left(T^{\alpha} - 1\right) \ln T$ and hence we conclude

$$\Pr\left\{\sup_{\theta \in \Theta_T} \left| \sum_{t=1}^{T} \left(s_{Tt}(\theta) - E\left[s_{Tt}(\theta)\right]\right) \right| > \varepsilon T^{\alpha-\beta} \right\} = o(1)$$

provided that $\frac{\alpha+1}{2} > \beta > \alpha$. By taking $\beta = \frac{1}{2} + \frac{1}{2}\alpha - \delta$ (the best possible rate), we have

$$\sup_{\theta \in \Theta_T} \left| \widehat{Q}_T(\theta) - Q(\theta) \right| = \sup_{\theta \in \Theta_T} \left| \sum_{t=1}^{T} \left(s_{Tt}(\theta) - E\left[s_{Tt}(\theta)\right]\right) \right| = o_p(T^{-\frac{1}{2}+\frac{1}{2}\alpha+\delta})$$

for all sufficiently small $\delta > 0$. ■

Lemma I.7 *Suppose (i) Assumption F.4 holds and (ii)$\frac{\zeta(K)^2 K}{T} \to 0$. Let $\eta_T = T^{-\beta_\eta}$ with $\alpha < \beta_\eta \leq 1/2 - \alpha/2 - \delta$,*

$$\sup_{\left\|\theta - \theta_K^*\right\| \leq o(\eta_T), \theta \in \Theta_T} \left| \widehat{Q}_T(\theta) - \widehat{Q}_T(\theta_K^*) - \left(Q(\theta) - Q(\theta_K^*)\right) \right| = o_p\left(\eta_T T^{-1/2+\alpha/2+\delta}\right)$$

Proof. In the following proof again, we treat $\epsilon_0 = 0$ to simplify discussions since we can pick ϵ_0 arbitrary small but still we need to impose that $f(\cdot,\theta) \in \underline{\mathcal{F}}_T$ is bounded away from zero.

Now applying the mean value theorem for $\overline{\theta}$ that lies between θ and θ_K^*, we can rewrite

$$
\begin{aligned}
& \widehat{Q}_T(\theta) - \widehat{Q}_T(\theta_K^*) - (Q(\theta) - Q(\theta_K^*)) \qquad (79)\\
= \; & \widehat{Q}_T(\theta) - Q(\theta) - \left(\widehat{Q}_T(\theta_K^*) - Q(\theta_K^*)\right) \\
= \; & \left(\frac{\partial\left(\widehat{Q}_T(\overline{\theta}) - Q(\overline{\theta})\right)}{\partial\theta'}\right)(\theta - \theta^o).
\end{aligned}
$$

Now consider for any $\widetilde{\theta}$ such that $\left\|\widetilde{\theta} - \theta_K^*\right\| \leq o(\eta_T)$,

$$
\begin{aligned}
& \frac{\partial \widehat{Q}_T(\widetilde{\theta})}{\partial\theta} - \frac{\partial Q(\widetilde{\theta})}{\partial\theta} \qquad (80)\\
= \; & -(k_1-1)\left(\frac{1}{T}\sum_{t=1}^{T}\tau(q_t,r_t)\frac{f(q_t,\widetilde{\theta})}{1-F(q_t,\widetilde{\theta})}\frac{\partial f(q_t,\widetilde{\theta})}{\partial\theta} - E\left[\tau(q_t,r_t)\frac{f(q_t,\widetilde{\theta})}{1-F(q_t,\widetilde{\theta})}\frac{\partial f(q_t,\widetilde{\theta})}{\partial\theta}\right]\right) \qquad (81)\\
& +\left(\frac{1}{T}\sum_{t=1}^{T}\tau(q_t,r_t)\frac{1}{f(q_t,\widetilde{\theta})}\frac{\partial f(q_t,\widetilde{\theta})}{\partial\theta} - E\left[\tau(q_t,r_t)\frac{1}{f(q_t,\widetilde{\theta})}\frac{\partial f(q_t,\widetilde{\theta})}{\partial\theta}\right]\right) \qquad (82)\\
& +(k_2-1)\left(\frac{1}{T}\sum_{t=1}^{T}\tau(q_t,r_t)\frac{f(r_t,\widetilde{\theta})}{1-F(r_t,\widetilde{\theta})}\frac{\partial f(r_t,\widetilde{\theta})}{\partial\theta} - E\left[\tau(q_t,r_t)\frac{f(r_t,\widetilde{\theta})}{1-F(r_t,\widetilde{\theta})}\frac{\partial f(r_t,\widetilde{\theta})}{\partial\theta}\right]\right) \qquad (83)
\end{aligned}
$$

We first bound (81). Note $\frac{\partial f(q_t,\theta)}{\partial\theta} = 2W^K(q_t)W^K(X_i)'\theta$ and define $M_{Tt} = \tau(q_t,r_t)\frac{f(q_t,\widetilde{\theta})}{1-F(q_t,\widetilde{\theta})}W^K(q_t)W^K(q_t)'\widetilde{\theta} - E\left[\tau(q_t,r_t)\frac{f(q_t,\widetilde{\theta})}{1-F(q_t,\widetilde{\theta})}W^K(q_t)W^K(q_t)'\widetilde{\theta}\right]$. Considering that M_{Tt} is a triangular array of i.i.d random variables with mean zero, we bound (81) as follows. First consider

$$
\mathrm{Var}\left[M_{Tt}\right] = E\left[M_{Tt}M_{Tt}'\right] \leq E\left[\tau(q_t,r_t)\left(\frac{f(q_t,\widetilde{\theta})}{1-F(q_t,\widetilde{\theta})}\right)^2\left(W^K(q_t)'\widetilde{\theta}\right)^2 W^K(q_t)W^K(q_t)'\right]. \qquad (84)
$$

The right hand side of (84) is bounded by

$$
\begin{aligned}
& E\left[\tau(q_t,r_t)\frac{f(q_t,\widetilde{\theta})^4}{\left(1-F(q_t,\widetilde{\theta})\right)^2}W^K(q_t)W^K(q_t)'\right] \\
\leq \; & \sup_{z\in\mathcal{V}_\varepsilon}\left(\frac{f(z,\widetilde{\theta})^4}{\left(1-F(z,\widetilde{\theta})\right)^2}\right)\sup_{z\in\mathcal{V}_\varepsilon} g^{(n-k_1-1:n)}(z)\int_{\mathcal{V}} W^K(z)W^K(z)'dz \\
\leq \; & C_1\left(\sup_{z\in\mathcal{V}_\varepsilon}\left|f(z,\widetilde{\theta}) - f_0(z)\right|^4 + \sup_{z\in\mathcal{V}_\varepsilon} f_0(z)^4\right) I_K
\end{aligned}
$$

since $1 - F(z, \widetilde{\theta})$ is bounded away from zero uniformly over $z \in \mathcal{V}_\varepsilon$ and since $g^{(n-k_1-1:n)}(z)$ is bounded away from above uniformly over $z \in \mathcal{V}_\varepsilon$. Finally note $\sup_{z\in\mathcal{V}_\varepsilon} \left| f(z, \widetilde{\theta}) - f_0(z) \right| \leq \sup_{z\in\mathcal{V}_\varepsilon} \left| f(z, \widetilde{\theta}) - f(z, \theta_K^*) \right| + \sup_{z\in\mathcal{V}_\varepsilon} |f(z, \theta_K^*) - f_0(z)| = O\left(\zeta(K)^2\right)\left(o(\eta_T) + O(K^{-s/2})\right)$ and hence we have

$$\mathrm{Var}\left[M_{Tt}\right] \leq C_2 \zeta(K)^8 \left(o(\eta_T^4) + O\left(K^{-2s}\right)\right) I_K + C_3 I_K \leq C_4 I_K$$

under $\zeta(K)^8 \left(o(\eta_T^4) + O\left(K^{-2s}\right)\right) = o(1)$ which holds as long as $s > 2$ and $-4\beta_\eta + 4\alpha < 0$. Now note

$$E\left(\left\| \frac{1}{\sqrt{T}} \sum_{t=1}^{T} M_{Tt} \right\|\right) \leq \sqrt{\mathrm{tr}\left(\mathrm{Var}\left[L_{Tt}\right]\right)} \leq \sqrt{C_4 \mathrm{tr}\left(I_K\right)} = O(\sqrt{K})$$

and hence (81) is $O_p\left(\sqrt{\frac{K}{T}}\right)$ from the Markov inequality. Similarly we can show that (83) is also $O_p\left(\sqrt{\frac{K}{T}}\right)$.

Now we bound (82). Define $L_{Tt} = \left(\tau(q_t, r_t) \frac{W^K(q_t)W^K(q_t)'\widetilde{\theta}}{f\left(q_t, \widetilde{\theta}\right)} - E\left[\frac{W^K(q_t)W^K(q_t)'\widetilde{\theta}}{f\left(q_t, \widetilde{\theta}\right)}\right]\right)$. Then, we can rewrite (82) as $2\frac{1}{T}\sum_{t=1}^{T} L_{Tt}$. Noting again L_{Tt} is a triangular array of i.i.d random variables with mean zero, we bound (82) as follows. Consider

$$\begin{aligned}
\mathrm{Var}\left[L_{Tt}\right] &= E\left[L_{Tt} L_{Tt}'\right] \leq E\left[\tau(q_t, r_t) \left(\frac{W^K(q_t)'\widetilde{\theta}}{f\left(q_t, \widetilde{\theta}\right)}\right)^2 W^K(q_t) W^K(q_t)'\right] \\
&= E\left[\tau(q_t, r_t) \frac{1}{f\left(q_t, \widetilde{\theta}\right)} W^K(q_t) W^K(q_t)'\right] \\
&\leq \sup_{z\in\mathcal{V}_\varepsilon} \frac{g^{(n-k_1-1:n)}(z)}{f\left(z, \widetilde{\theta}\right)} \int_{\mathcal{V}} W^K(z) W^K(z)' dz \leq C_1 I_K
\end{aligned}$$

since $g^{(n-k_1-1:n)}(z)$ is bounded from above and $f\left(z, \widetilde{\theta}\right)$ is bounded from below uniformly over $z \in \mathcal{V}_\varepsilon$. It follows that

$$E\left(\left\| \frac{1}{\sqrt{T}} \sum_{t=1}^{T} L_{Tt} \right\|\right) \leq \sqrt{\mathrm{tr}\left(\mathrm{Var}\left[L_{Tt}\right]\right)} \leq \sqrt{C_1 \mathrm{tr}\left(I_K\right)} = O(\sqrt{K}).$$

Thus, we bound (82) as $O_p\left(\sqrt{\frac{K}{T}}\right)$ from the Markov inequality. We conclude

$$\left\| \frac{\partial \widehat{Q}_n(\widetilde{\theta})}{\partial \theta} - \frac{\partial Q(\widetilde{\theta})}{\partial \theta} \right\| = O_p\left(\sqrt{\frac{K}{n}}\right)$$

under $s > 2$ and $-4\beta_\eta + 4\alpha < 0$. Thus, noting $O_p\left(\sqrt{\frac{K}{n}}\right) = o_p\left(n^{-1/2+\alpha/2+\delta}\right)$ for $K = n^\alpha$ and sufficiently small δ, we have

$$\begin{aligned}
& \sup_{\left\|\theta-\theta_K^*\right\| \leq o(\eta_T), \theta \in \Theta_T} \left|\widehat{Q}_T(\theta) - \widehat{Q}_T(\theta_K^*) - (Q(\theta) - Q(\theta_K^*))\right| \\
\leq & \sup_{\left\|\theta-\theta_K^*\right\| \leq o(\eta_T), \theta \in \Theta_T} \left\| \frac{\partial\left(\widehat{Q}_T(\theta) - Q(\theta)\right)}{\partial\theta} \right\| \sup_{\left\|\theta-\theta_K^*\right\| \leq o(\eta_T), \theta \in \Theta_T} \left\|\theta - \theta_K^*\right\| \\
= & o_p\left(\eta_T T^{-1/2+\alpha/2+\delta}\right)
\end{aligned}$$

from (79) applying the Cauchy-Schwarz inequality. ■

I.4 Proof of Lemma F.2

Proof. Following Kim (2005), we can show $\left\|\widehat{\theta}_K - \theta_K^*\right\| = o_p(T^{-\kappa}) = o_p(T^{-1/2+\alpha/2+\delta})$. Here we reproduce Kim (2005)'s proof. The idea is that the convergence rate contributed iteratively by the local curvature of $\widehat{Q}_T(\theta)$ around θ_K^* can be achieved up to the convergence rate of the uniform convergence of $\widehat{Q}_T(\theta)$ to $Q(\theta)$ and hence we can obtain the convergence rate of $o_p(T^{-1/2+\alpha/2+\delta})$. A formal proof is as follows.

First, from (75) and $\|\theta_K^* - \theta_K\| = O\left(K^{-s/2}\right)$ with $s > 2$, we note that

$$\lambda_{\min}\left(-\frac{1}{2}\frac{\partial^2 Q(\widetilde{\theta})}{\partial\theta\partial\theta'}\right) \geq C_1 \inf_{v \in \mathcal{V}}\left(\frac{f_0(z)}{O\left(\zeta(K)K^{-s}\right) + O\left(\zeta(K)^2\right)\left(\|\theta - \theta_K^*\|\right) + f_0(z)}\right)$$

where $\widetilde{\theta}$ lies between θ and θ_K^* and hence from (72),

$$\begin{aligned}
Q(\theta_K^*) - Q(\theta) &\geq C_1 \left\|\theta - \theta_K^*\right\|^2 \text{ if } \left\|\theta - \theta_K^*\right\| = o\left(\zeta(K)^{-2}\right) \\
Q(\theta_K^*) - Q(\theta) &\geq C_2 \zeta(K)^{-2} \left\|\theta - \theta_K^*\right\| \text{ otherwise.}
\end{aligned} \tag{85}$$

Denote $\kappa = 1/2 - \alpha/2 - \delta$ and $\eta_{0T} = o(T^{-\kappa})$. We derive the convergence rate in two cases: one is when η_{0T} has the equal or a smaller order than $o\left(\zeta(K)^{-4}\right)$ and the other case is when η_{0T} has a larger order than $o\left(\zeta(K)^{-4}\right)$.

1) *When η_{0T} has equal or smaller order than $o\left(\zeta(K)^{-4}\right)$*, which holds under $\alpha < \frac{1}{5}$:

Now let $\delta_{0T} = \sqrt{2\eta_{0T}}$. For any c such that $C_1 c^2 > 1$, it follows

$$\begin{aligned}
& \Pr\left(\left\|\widehat{\theta}_K - \theta_K^*\right\| \geq c\delta_{0T}\right) \qquad (86)\\
\leq\ & \Pr\left(\sup_{\left\|\theta-\theta_K^*\right\| \geq c\delta_{0T}, \theta\in\Theta_T} \widehat{Q}_T(\theta) \geq \widehat{Q}_T(\theta_K^*)\right)\\
\leq\ & \Pr\left(\sup_{\theta\in\Theta_T}\left|\widehat{Q}_T(\theta) - Q(\theta)\right| > \eta_{0T}\right)\\
& + \Pr\left(\left\{\sup_{\theta\in\Theta_T}\left|\widehat{Q}_T(\theta) - Q(\theta)\right| \leq \eta_{0T}\right\} \cap \left\{\sup_{\left\|\theta-\theta_K^*\right\| \geq c\delta_{0T}, \theta\in\Theta_T} \widehat{Q}_T(\theta) \geq \widehat{Q}_T(\theta_K^*)\right\}\right)\\
\leq\ & \Pr\left(\sup_{\theta\in\Theta_T}\left|\widehat{Q}_T(\theta) - Q(\theta)\right| > \eta_{0T}\right) + \Pr\left(\sup_{\left\|\theta-\theta_K^*\right\| \geq c\delta_{0T}, \theta\in\Theta_T} Q(\theta) \geq Q(\theta_K^*) - 2\eta_{0T}\right)\\
=\ & P_1 + P_2.
\end{aligned}$$

Now note $P_1 \to 0$ from (49). Now we show $P_2 \to 0$. This holds since $Q(\theta)$ has its maximum at θ_K^*. To be precise, note

$$\begin{aligned}
Q(\theta_K^*) - Q(\theta) &\geq C_1 \left\|\theta - \theta_K^*\right\|^2 \geq 2C_1 c^2 \eta_{0T} \text{ if } \left\|\theta - \theta_K^*\right\| = o\left(\zeta(K)^{-2}\right)\\
Q(\theta_K^*) - Q(\theta) &\geq C_2 \zeta(K)^{-2} \left\|\theta - \theta_K^*\right\| \geq C_3 \zeta(K)^{-4} \text{ otherwise}
\end{aligned}$$

and hence

$$\begin{aligned}
\sup_{\left\|\theta-\theta_K^*\right\| \geq c\delta_{0T}, \theta\in\Theta_T} Q(\theta) - Q(\theta_K^*) &\leq -2C_1 c^2 \eta_{0T} \text{ if } \left\|\theta - \theta_K^*\right\| = o\left(\zeta(K)^{-2}\right)\\
\sup_{\left\|\theta-\theta_K^*\right\| \geq c\delta_{0T}, \theta\in\Theta_T} Q(\theta) - Q(\theta_K^*) &< -C_3 \zeta(K)^{-4} \text{ otherwise.}
\end{aligned}$$

Therefore, as long as $C_1 c^2 > 1$ and $\zeta(K)^4 \eta_{0T} \to 0$, we have $P_2 \to 0$. $\zeta(K)^4 \eta_{0T} \to 0$ holds under $\alpha < \frac{1}{5}$. Thus, we have proved $\left\|\widehat{\theta}_K - \theta_K^*\right\| = o_p(T^{-\kappa/2})$.

Now we refine the convergence rate by exploiting the local curvature of $\widehat{Q}_T(\theta)$ around θ_K^*. Let $\eta_{1T} = n^{-\kappa}\delta_{0T} = o\left(n^{-(\kappa+\kappa/2)}\right)$ and $\delta_{1T} = \sqrt{\eta_{1T}} = o\left(T^{-(\kappa/2+\kappa/4)}\right)$. For any c such that $C_1 c^2 > 1$, we have

$$
\begin{aligned}
& \Pr\left(\left\|\widehat{\theta}_K - \theta_K^*\right\| \geq c\delta_{1T}\right) \qquad (87)\\
\leq\ & \Pr\left(\sup_{\delta_{0T}\geq\left\|\theta-\theta_K^*\right\|\geq c\delta_{1T},\theta\in\Theta_T} \widehat{Q}_T(\theta) \geq \widehat{Q}_T(\theta_K^*)\right)\\
\leq\ & \Pr\left(\sup_{\delta_{0T}\geq\left\|\theta-\theta_K^*\right\|\geq c\delta_{1T},\theta\in\Theta_T} \left|\widehat{Q}_T(\theta) - \widehat{Q}_T(\theta_K^*) - (Q(\theta) - Q(\theta_K^*))\right| > \eta_{1T}\right)\\
& + \Pr\left(\begin{array}{c} \left\{\sup_{\delta_{0T}\geq\left\|\theta-\theta_K^*\right\|\geq c\delta_{1T},\theta\in\Theta_T} \left|\widehat{Q}_T(\theta) - \widehat{Q}_T(\theta_K^*) - (Q(\theta) - Q(\theta_K^*))\right| \leq \eta_{1T}\right\} \\ \cap \left\{\sup_{\delta_{0T}\geq\left\|\theta-\theta_K^*\right\|\geq c\delta_{1T},\theta\in\Theta_T} \widehat{Q}_T(\theta) \geq \widehat{Q}_T(\theta_K^*)\right\}\end{array}\right)\\
\leq\ & \Pr\left(\sup_{\delta_{0T}\geq\left\|\theta-\theta_K^*\right\|\geq c\delta_{1T},\theta\in\Theta_T} \left|\widehat{Q}_T(\theta) - \widehat{Q}_T(\theta_K^*) - (Q(\theta) - Q(\theta_K^*))\right| > \eta_{1T}\right)\\
& + \Pr\left(\sup_{\delta_{0T}\geq\left\|\theta-\theta_K^*\right\|\geq c\delta_{1T},\theta\in\Theta_T} Q(\theta) \geq Q(\theta_K^*) - \eta_{1T}\right)\\
=\ & P_3 + P_4
\end{aligned}
$$

where[37] $P_3 \to 0$ from Lemma I.7 and $P_4 \to 0$ similarly with P_2 noting

$$\sup_{\delta_{0T}\geq\left\|\theta-\theta_K^*\right\|\geq c\delta_{1T},\theta\in\Theta_T} Q(\theta) - Q(\theta_K^*) \leq -C_1 c^2 \eta_{1T}$$

[37] Note

$$\sup_{\delta_{0n}\geq\left\|\theta-\theta_K^*\right\|\geq c\delta_{1n},\theta\in\Theta_n} \left|\widehat{Q}_n(\theta) - \widehat{Q}_n(\theta_K^*) - (Q(\theta) - Q(\theta_K^*))\right| \leq \eta_{1n}$$

implies, for any θ such that $\delta_{0n} \geq \|\theta - \theta_K^*\| \geq c\delta_{1n}$,

$$
\begin{aligned}
-\eta_{1n} - \sup_{\delta_{0n}\geq\left\|\theta-\theta_K^*\right\|\geq c\delta_{1n},\theta\in\Theta_n} (Q(\theta) - Q(\theta_K^*)) &\leq \widehat{Q}_n(\theta) - \widehat{Q}_n(\theta_K^*)\\
&\leq \eta_{1n} + \sup_{\delta_{0n}\geq\left\|\theta-\theta_K^*\right\|\geq c\delta_{1n},\theta\in\Theta_n} (Q(\theta) - Q(\theta_K^*))
\end{aligned}
$$

and hence we obtain

$$\sup_{\delta_{0n}\geq\left\|\theta-\theta_K^*\right\|\geq c\delta_{1n},\theta\in\Theta_n} \widehat{Q}_n(\theta) - \widehat{Q}_n(\theta_K^*) \leq \eta_{1n} + \sup_{\delta_{0n}\geq\left\|\theta-\theta_K^*\right\|\geq c\delta_{1n},\theta\in\Theta_n} (Q(\theta) - Q(\theta_K^*))$$

Therefore,

$$\Pr\left(\sup_{\delta_{0n}\geq\left\|\theta-\theta_K^*\right\|\geq c\delta_{1n},\theta\in\Theta_n} \widehat{Q}_n(\theta) - \widehat{Q}_n(\theta_K^*) > 0\right) \leq \Pr\left(\sup_{\delta_{0n}\geq\left\|\theta-\theta_K^*\right\|\geq c\delta_{1n},\theta\in\Theta_n} Q(\theta) \geq Q(\theta_K^*) - \eta_{1n}\right),$$

from which we have obtained the third inequality.

by (85) and since $\|\theta - \theta_K^*\| = o\left(\zeta(K)^{-2}\right)$ for any θ such that $\delta_{0T} \geq \|\theta - \theta_K^*\| \geq c\delta_{1T}$ under $\alpha < \frac{1}{5}$. This show that $\left\|\widehat{\theta}_K - \theta_K^*\right\| = o_p(T^{-(\kappa/2+\kappa/4)})$. Repeating this refinement for infinite number of times, we obtain

$$\left\|\widehat{\theta}_K - \theta_K^*\right\| = o_p(T^{-(\kappa/2+\kappa/4+\kappa/8+\ldots)}) = o_p(T^{-\kappa}) = o_p(T^{-1/2+\alpha/2+\delta})$$

under $\alpha < \frac{1}{5}$.

2) Now we consider *when* η_{0T} *has larger order than* $o\left(\zeta(K)^{-4}\right)$ (which holds under $\alpha \geq \frac{1}{5}$): Let $\widetilde{\delta}_{0T} = o\left(\zeta(K)^{-2}\right) T^{\beta}$ for $\beta > 0$. Then, from (86), we have

$$\begin{aligned} & \Pr\left(\left\|\widehat{\theta}_K - \theta_K^*\right\| \geq c\widetilde{\delta}_{0T}\right) \\ \leq\ & \Pr\left(\sup_{\theta \in \Theta_T} \left|\widehat{Q}_T(\theta) - Q(\theta)\right| > \eta_{0T}\right) + \Pr\left(\sup_{\|\theta-\theta_K^*\| \geq c\widetilde{\delta}_{0T}, \theta \in \Theta_T} Q(\theta) \geq Q(\theta_K^*) - 2\eta_{0T}\right) \\ =\ & P_1 + P_2. \end{aligned}$$

Again note $P_1 \to 0$ from (49). Now we show $P_2 \to 0$. Note from (85),

$$\sup_{\|\theta-\theta_K^*\| \geq c\widetilde{\delta}_{0T}, \theta \in \Theta_T} Q(\theta) - Q(\theta_K^*) \leq -C_2 \zeta(K)^{-2} \|\theta - \theta_K^*\| \leq -o\left(\zeta(K)^{-4}\right) T^{\beta}$$

since $\|\theta - \theta_K^*\| > o\left(\zeta(K)^{-2}\right)$ for any θ such that $\|\theta - \theta_K^*\| \geq c\widetilde{\delta}_{0T}$. It follows that $P_2 \to 0$ as long as $\zeta(K)^4 T^{-\beta} \eta_{0T} \to 0$, which holds under

$$\beta > \frac{5}{2}\alpha - \frac{1}{2} + \delta \tag{88}$$

and hence the convergence rate will be $o_p\left(\delta_{0T}\right) = o_p\left(T^{-\alpha+\beta}\right)$. Now we refine the convergence rate by exploiting the local curvature of $\widehat{Q}_T(\theta)$ around θ_K^* again. Let $\widetilde{\eta}_{1T} = T^{-\kappa}\widetilde{\delta}_{0T} = o\left(T^{-((\alpha-\beta)+\kappa)}\right)$ and $\widetilde{\delta}_{1T} = \sqrt{\widetilde{\eta}_{1T}} = o\left(T^{-((\alpha-\beta)/2+\kappa/2)}\right)$. Then, from (87), we have

$$\begin{aligned} & \Pr\left(\left\|\widehat{\theta}_K - \theta_K^*\right\| \geq c\widetilde{\delta}_{1T}\right) \\ \leq\ & \Pr\left(\sup_{\widetilde{\delta}_{0T} \geq \|\theta-\theta_K^*\| \geq c\widetilde{\delta}_{1T}, \theta \in \Theta_T} \left|\widehat{Q}_T(\theta) - \widehat{Q}_T(\theta_K^*) - (Q(\theta) - Q(\theta_K^*))\right| > \widetilde{\eta}_{1T}\right) \\ & + \Pr\left(\sup_{\delta_{0T} \geq \|\theta-\theta_K^*\| \geq c\delta_{1T}, \theta \in \Theta_T} Q(\theta) \geq Q(\theta_K^*) - \widetilde{\eta}_{1T}\right) \\ =\ & P_3 + P_{41} \end{aligned} \tag{89}$$

where $P_3 \to 0$ from Lemma I.7. Now we show $P_{41} \to 0$ similarly with P_2. Here again we need to consider two cases:

2-1) *When* $\widetilde{\delta}_{1T}$ *has equal or smaller order than* $o\left(\zeta(K)^{-2}\right)$, which holds under $\beta \leq \frac{1}{2} - \frac{3}{2}\alpha - \delta$ and hence from $\alpha > \beta$ and (88) it requires $1/5 \leq \alpha < 1/4$. Under this case, note

$$\begin{aligned}
\sup_{\widetilde{\delta}_{0T} \geq \left\|\theta - \theta_K^*\right\| \geq c\widetilde{\delta}_{1T}, \theta \in \Theta_T} Q(\theta) - Q(\theta_K^*) &\leq -C_1 \left\|\theta - \theta_K^*\right\|^2 \leq -C_1 c^2 \widetilde{\delta}_{1T}^2 = -C_1 c^2 \widetilde{\eta}_{1T} \text{ if } \left\|\theta - \theta_K^*\right\| = o\left(\zeta(K)^{-2}\right) \\
\sup_{\widetilde{\delta}_{0T} \geq \left\|\theta - \theta_K^*\right\| \geq c\widetilde{\delta}_{1T}, \theta \in \Theta_T} Q(\theta) - Q(\theta_K^*) &\leq -C_2 \zeta(K)^{-2} \left\|\theta - \theta_K^*\right\| \leq -C_2 \zeta(K)^{-4} \text{ otherwise}
\end{aligned}$$

by (85) and hence $P_{41} \to 0$ as long as $C_1 c^2 > 1$ and $\zeta(K)^4 \widetilde{\eta}_{1T} = \zeta(K)^4 \widetilde{\eta}_{1T} T^{-\kappa} \widetilde{\delta}_{0T} = \zeta(K)^2 o\left(T^{-\kappa} T^{\beta}\right) \to 0$, which holds under $\beta \leq \frac{1}{2} - \frac{3}{2}\alpha - \delta$. Repeating this refinement for infinite number of times (noting that for any θ such that $\widetilde{\delta}_{1T} \geq \left\|\theta - \theta_K^*\right\|$, we have $\left\|\theta - \theta_K^*\right\| = o\left(\zeta(K)^{-2}\right)$), we obtain

$$\left\|\widehat{\theta}_K - \theta_K^*\right\| = o_p(T^{-\left(\lim_{L \to \infty}\left(\sum_{l=1}^{L} \frac{\kappa}{2^l} + \frac{(\alpha-\beta)}{L}\right)\right)}) = o_p(T^{-\kappa})$$

and hence the effect of η_{0T}'s *having larger order than* $o\left(\zeta(K)^{-4}\right)$ disappear ($\frac{(\alpha-\beta)}{L}$ goes to zero as L goes to infinity)). This makes sense because the iterated convergence rate improvement using the local curvature will dominate the convergence rate from the uniform convergence.

2-2) *When* $\widetilde{\delta}_{1T}$ *has bigger order than* $o\left(\zeta(K)^{-2}\right)$, which holds under $\beta > \frac{1}{2} - \frac{3}{2}\alpha - \delta$ and $1/3 > \alpha \geq 1/4$:

In this case, we let $\widetilde{\delta}_{1T} = \widetilde{\delta}_{0T} n T^{-\gamma}$ for some $\gamma > 0$ and hence we require $\beta > \gamma$. From (89), we note

$$\begin{aligned}
& \Pr\left(\left\|\widehat{\theta}_K - \theta_K^*\right\| \geq c\widetilde{\delta}_{1T}\right) \\
\leq\ & \Pr\left(\sup_{\widetilde{\delta}_{0T} \geq \left\|\theta - \theta_K^*\right\| \geq c\widetilde{\delta}_{1T}, \theta \in \Theta_T} \left|\widehat{Q}_T(\theta) - \widehat{Q}_T(\theta_K^*) - (Q(\theta) - Q(\theta_K^*))\right| > \widetilde{\eta}_{1T}\right) \\
& + \Pr\left(\sup_{\widetilde{\delta}_{0T} \geq \left\|\theta - \theta_K^*\right\| \geq c\widetilde{\delta}_{1T}, \theta \in \Theta_T} Q(\theta) \geq Q(\theta_K^*) - \widetilde{\eta}_{1T}\right) \\
=\ & P_3 + P_{42}.
\end{aligned}$$

We have seen that P_3 goes to zero since $\widetilde{\eta}_{1T} = T^{-\kappa} \widetilde{\delta}_{0T}$ and by Lemma I.7. Now we verify P_{42} goes to zero. From (85), to have $P_{42} \to 0$, we require that $\zeta(K)^{-2} \widetilde{\delta}_{1T}$ have a bigger order than $\widetilde{\eta}_{1T}$ and hence we need $\gamma < \frac{1}{2} - \frac{3\alpha}{2} - \delta$. Now we improve the convergence rate again using the local curvature by defining $\widetilde{\eta}_{2T} = T^{-\kappa} \widetilde{\delta}_{1T} = o\left(T^{-((\alpha-\beta+\gamma)+\kappa)}\right)$ and $\widetilde{\delta}_{2T} = \sqrt{\widetilde{\eta}_{2T}} = o\left(T^{-((\alpha-\beta+\gamma)/2+\kappa/2)}\right)$. Then, similarly with before, at the end, we will obtain $\left\|\widehat{\theta}_K - \theta_K^*\right\| = o_p(T^{-\kappa})$ as long as $\widetilde{\delta}_{2T}$ has equal or smaller order than $o\left(\zeta(K)^{-2}\right)$. The tricky case is again when $\widetilde{\delta}_{2T}$ has a bigger order than $o\left(\zeta(K)^{-2}\right)$, which happens when $\beta - \gamma > \frac{1}{2} - \frac{3}{2}\alpha - \delta$ but applying the same trick, at the end, we will obtain the same convergence rate of $\left\|\widehat{\theta}_K - \theta_K^*\right\| = o_p(T^{-\kappa})$ as long as $1/3 > \alpha$. Combining these results, we conclude that under $\alpha < 1/3$, we have

$$\left\|\widehat{\theta}_K - \theta_K^*\right\| = o_p(T^{-\kappa}) = o_p(T^{-1/2+\alpha/2+\delta}).$$

This result is intuitive in the sense that ignoring δ, we obtain $o\left(\zeta\left(K\right)^{-2}\right) = o(T^{-\kappa}) = T^{-1/3}$ at $\alpha = 1/3$ and hence if $\alpha \geq 1/3$, there is no room to improve the convergence rate using the local curvature. ∎

Table 1. Summary Statistics (Sample Size: 5184)

	Age (years)	Engine Size (ℓ)	Mileage (km)	Rating (0-10)	2nd highest	3rd -	4th -	Reserve price	Opening bid
Mean	5.9278	1.7174	96888	3.1182	3812.1	3758.8	3690.5	3402.7	3102.5
S.D.	2.3689	0.3604	49414	1.2127	3004.8	2999.7	2992.1	2905.4	2866.1
Median	6.1667	1.4980	94697	3.5	3200	3150	3085	2800	2500
Max	12.75	3.4960	426970	7	29640	29580	29490	27800	27000
Min	0.1667	1.1390	69	0	100	70	70	50	0

Note: All prices are in 1000 Korean Won. (1000 Korean Won $\approx$1 US Dollar)

Table 2. Market Shares in the Sample

	Hyundai	Daewoo	Kia	Ssangyong	Others
Share (%)	44.72	30.86	21.05	2.66	0.71

(Table 3. in the next page)

Table 4. Test Statistic

Specification	A	B	C
Values	0.6824	4.8835	0.9198

Note: 1. A-Test statistic with (2nd-4th) and (3rd-4th)
B-(2nd-4th) and (2nd-3rd)
C-(3rd-4th) and (2nd-3rd)
2. Cutoff point for the chi-squared distribution with df=1 (99% cumulative probability): 6.6349

Table 3: First-Stage Estimation Results (Linear model)

Maker	Covariate	Estimate	Standard Error
Hyundai	Intercept	5.2247	0.094469
(n=1676)	Age	-0.24388	0.005780
	Engine Size (l)	0.7760	0.036711
	Mileage ($10^4 km$)	0.00043601	0.002171
	Rating	0.1241	0.008193
	Title Remaining	-0.03787	0.010661
	Automatic	0.28736	0.019231
	Gasoline	0.01551	0.038101
	Popular Colors	0.05667	0.018439
	Best Options	0.24411	0.031633
	Better Options	0.11781	0.020955
Daewoo	Intercept	5.2821	0.099089
(n=1186)	Age	-0.33693	0.006244
	Engine Size (l)	0.94696	0.047855
	Mileage ($10^4 km$)	-0.010808	0.002683
	Rating	0.05077	0.010107
	Title Remaining	-0.1131	0.012582
	Automatic	0.2430	0.021502
	Gasoline	0.3306	0.037409
	Popular Colors	-0.05506	0.019077
	Best Options	0.13444	0.029881
	Better Options	0.02440	0.021799
Kia & etc.	Intercept	5.2698	0.108150
(n=964)	Age	-0.2234	0.006959
	Engine Size (l)	0.5731	0.039005
	Mileage ($10^4 km$)	0.0054277	0.002572
	Rating	0.10036	0.011082
	Title Remaining	0.02598	0.016192
	Automatic	0.33469	0.024719
	Gasoline	-0.19433	0.041421
	Popular Colors	0.02781	0.022544
	Best Options	0.39472	0.042396
	Better Options	0.20618	0.027255

Price \ ID	025110	062101	069702	075401	076101	Total
1700					1	1
1730					1	1
1760					1	1
1790	1				1	2
1820	1				1	2
1850					1	1
1880					1	1
1910					1	1
1940			1		1	2
1970		1	1			2
2000		1			1	2
2030		1	1	1	1	4
2060	1			1	1	3
2090		1			1	2
2120	1	1				2
2150	1				1	2
2180				1	1	2
2210					1	1
Total	5	5	3	3	17	

Figure 1: Sample Bidding Log (Opening bid: 1700, Reserve price: 2000, Transaction price: 2210, Total bidders logged: 5)

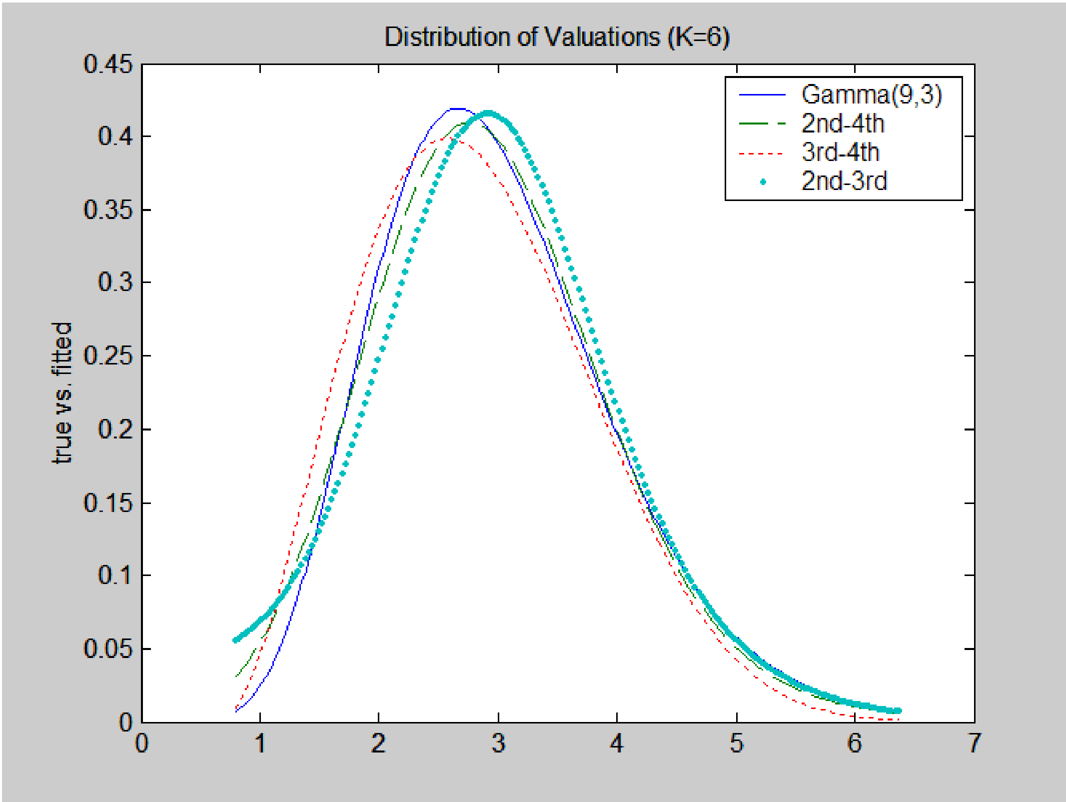

Figure 2: Estimated density function with $K = 6$ for a Monte Carlo simulated data

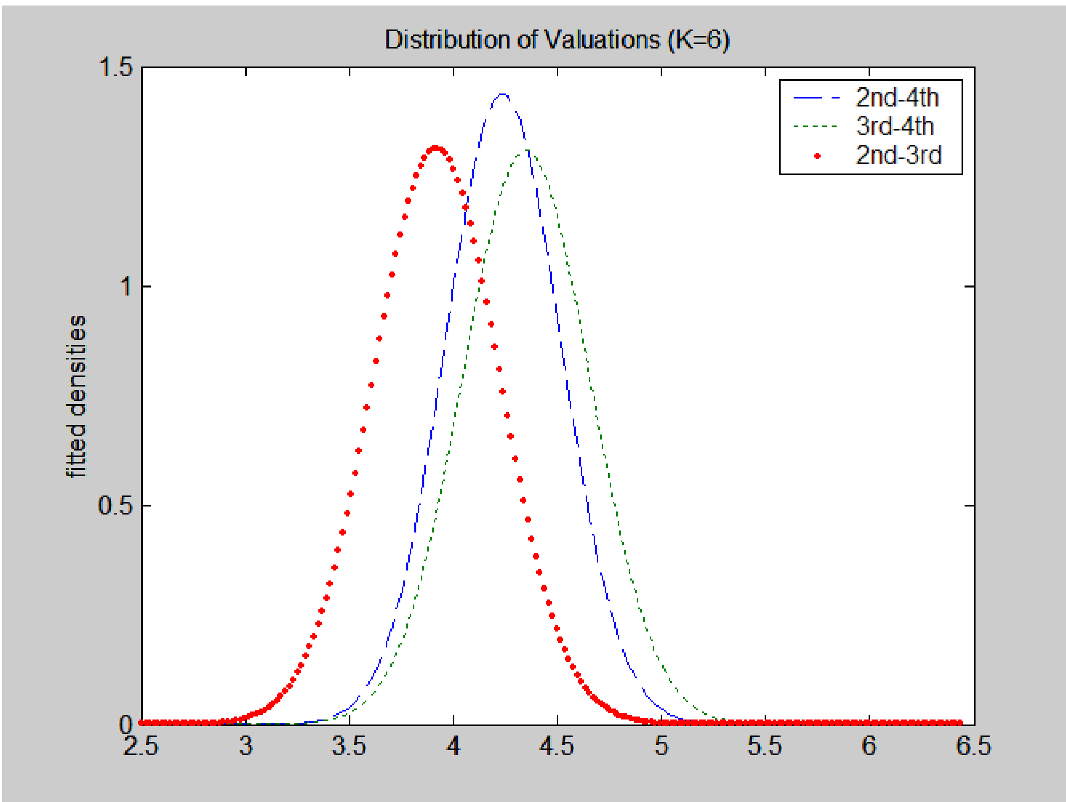

Figure 3: Estimated density function with $K = 6$ for the wholedale used-car auction data