

Collegio Carlo Alberto



On the Asymptotic Distribution of the Transaction Price in a Clock Model of a Multi-Unit, Oral, Ascending-Price Auction within the Common-Value Paradigm

Han Hong

Harry J. Paarsch

Pai Xu

No. 186

December 2010

Carlo Alberto Notebooks

www.carloalberto.org/working_papers

On the Asymptotic Distribution of the Transaction Price in a Clock Model of a Multi-Unit, Oral, Ascending-Price Auction within the Common-Value Paradigm

Han Hong

Department of Economics, Stanford University

Harry J. Paarsch*

Department of Economics, University of Melbourne

Pai Xu

School of Economics and Finance, University of Hong Kong

20 December 2010

Abstract

Using a clock model of a multi-unit, oral, ascending-price auction, within the common-value paradigm, we analyse the asymptotic behaviour of the transaction price as the number of bidders gets large. We find that even though the transaction price is determined by a (potentially small) fraction of losing drop-out bids, that price converges in probability to the *ex ante* unknown, true value. Subsequently, we derive the asymptotic distribution of the transaction price. Finally, we apply our methods to data from an auction of taxi license plates held in Shenzhen, China.

Key words: common value; information aggregation; multi-unit auctions; taxis.

JEL Classification: C20, D44, L1.

*Corresponding author: Harry J Paarsch, Level 5, Economics & Commerce Building, Parkville, 3010, Victoria, Australia. Telephone: +61 3 8344 4655; Facsimile: +61 3 8344 6899. E-mail: hjp@paarsch.unimelb.edu.au.

1 Motivation and Introduction

During the past half century, economists have made considerable progress in understanding the theoretical structure of equilibrium strategic behaviour under market mechanisms, such as auctions; see Krishna [2010] for a comprehensive presentation and evaluation of progress.

One analytic device, commonly used to describe bidder motivation at single-object auctions, is a continuous random variable that represents individual-specific signals concerning the object's true, but unknown, value. This true, but unknown, value will be revealed only *after* the auction has ended, when the winner has been determined and the transaction price paid. Regardless of the winner, however, the value of the object is the same to all.

The conceptual experiment involves each potential bidder's receiving a draw from a signal distribution. Conditional on his draw, a bidder is then assumed to act purposefully, using the information in his signal along with Bayes' rule to maximize either the expected profit or the expected utility of profit from winning the auction. Another frequently-made assumption is that the signal draws of bidders are independent and that the bidders are *ex ante* symmetric—their draws coming from the same distribution of signals. This framework is often referred to as the *symmetric common-value paradigm* (symmetric CVP).

Under these assumptions, a researcher can then focus on a representative agent's decision rule when characterizing equilibrium behaviour. Wilson [1977] invented this framework to illustrate that the winner's curse could not obtain, in equilibrium, among rational bidders. He also demonstrated that, when the number of bidders n is large (tends to infinity), the winning bid at first-price, sealed-bid auctions converges almost surely to the true, but unknown, value of the object. In other words, the auction format and pricing rule play an important role in aggregating the disparate, individual pieces of information held by the bidders. Milgrom [1979] subsequently provided a precise characterization of the structure the signal distribution must possess in order for this convergence property to hold; Pesendorfer and Swinkels [1997] have referred to this as *full information aggregation*.

When several, say k , units of a good are simultaneously for sale, at least two important ques-

tions arise: specifically, who will be the winning bidders and what price(s) will those winners pay? Weber [1983] has described a number of different multi-unit auction formats as well as pricing rules under those formats. For example, Milgrom [1981] developed a natural generalization of the Wilson [1977] model. In Milgrom's model, each bidder submits a price and the auctioneer then aggregates these demands, allocating the units to those bidders with the highest k submitted bids. The winners then pay a uniform price—specifically, the highest rejected bid.

Pesendorfer and Swinkels [1997] have built on this research by investigating a sequence of auctions $\{A_r\}$ in which both n_r and k_r increase. They demonstrated that a necessary and sufficient condition for full information aggregation is that $k_r \rightarrow \infty$ and $(n_r - k_r) \rightarrow \infty$, a condition they referred to as *double largeness*. Under this condition, non-negligible supply can be a substitute for the strong signal structure required in Wilson [1977] as well as Milgrom [1979, 1981]. Kremer [2002] has investigated this further.

While it is heartening to know there are conditions under which transaction prices will converge in probability to the true, but unknown, values of objects for sale, the rate at which these prices converge is also of interest. In particular, Hong and Shum [2004] asked the question “How large must n be to be large enough?” and then investigated the rates of information aggregation in common-value environments. Knowing the conditions under which the transaction price provides a potentially useful estimate of the object's unknown value is important to understanding the process some refer to as *price discovery* because, in practice, neither the number of bidders nor the number of units for sale at an auction ever really gets to infinity.

Of course, the pricing rule investigated in Wilson [1977] and Milgrom [1979, 1981] as well as Pesendorfer and Swinkels [1997, 2000] is not the only pricing rule that could be used under a sealed-bid format. For example, another pricing rule would involve allocating the k units to those bidders who tendered the highest k bids, but each winner would then pay what he bid for the unit(s) he won. In general, at multi-unit auctions, different auction formats and different pricing rules induce different equilibrium behaviour and, thus, translate into different transaction prices as well as potentially different expected revenues for sellers. Hence, as Jackson and Kremer [2004, 2006]

have emphasized, understanding the effects of auction formats and pricing rules has important practical relevance. Even small changes can have effects, as has been illustrated by Mezzetti and Tsetlin [2008, 2009].

In a companion paper to Milgrom and Weber [1982], which was published nearly two decades later, Milgrom and Weber [2000] proposed an interesting pricing rule for multi-unit, oral, ascending-price auctions. The model considered by Milgrom and Weber is the multi-unit variant of the clock model introduced by Milgrom and Weber [1982] in order to investigate behaviour at single-object, oral, ascending-price (often referred to as *English*) auctions. In the multi-unit model, bidders are assumed to demand at most one unit of the good for sale; Milgrom [2004] has referred to this as *singleton demand*. The current price for all units on sale rises continuously according to some device, such as a clock. As the price rises, the drop-out prices of losing participants are recorded when they exit the auction. The transaction price is the drop-out price of the last participant to exit the auction. Each of the remaining k participants is then allocated one unit at the transaction price.

One attractive feature of oral, ascending-price auctions *vis-à-vis* sealed-bid ones, regardless of the pricing rule, is the scope for information release at oral, ascending-price auctions. This is particularly important in informational environments with substantial common-value components. In such environments, by observing the actions of his competitors, a bidder can augment the information contained in his signal and, thus, may be able to reduce the uncertainty concerning the unknown value of the object for sale. Other things being equal, this reduction in uncertainty can induce participants to bid more aggressively than under sealed-bid formats, which means the revenues the seller can expect to garner can increase. The greater is the linkage between a bidder's information and what he perceives others will bid, the higher the bidding; Milgrom and Weber [1982] have referred to this as the *linkage principle*. In models of single-object auctions, they used it to rank the revenues a seller can expect to garner under the different auction formats and pricing rules. Specifically, in a theoretical model with one object for sale as well as risk-neutral potential buyers who have affiliated signals from the same marginal distribution, Milgrom and Weber [1982] demonstrated that the English auction format yields, on average, more revenue than first-price auc-

tions, such as the oral, descending-price (Dutch) format or the first-price, sealed-bid format. One can deduce from the structure of the proof in Milgrom and Weber [1982] that the same linkage principle applies to the generalized Milgrom–Weber auction we study below.

Our paper is in six additional sections. In the next, we use the Milgrom–Weber clock model to develop a theoretical framework within which to investigate the stochastic behaviour of the transaction price at a multi-unit, oral, ascending-price auction within the common-value paradigm, while in section 3, we demonstrate that the transaction price converges in probability to the *ex ante* unknown, true value as the number of bidders n and the number of units k get large in the Pesendorfer–Swinkels sense. In section 4, we characterize the asymptotic distribution of the transaction price when both the number of bidders and the number of units get large. We also demonstrate that the asymptotic variance of the transaction price under the Milgrom–Weber pricing rule is less than that under the pricing rule used by Pesendorfer and Swinkels. Thus, if the transaction prices under different auction formats and pricing rules are viewed as statistical estimators of the true, but unknown, value of the units for sale, then the Milgrom–Weber rule is a more efficient estimator of the unknown value than the first-price, sealed-bid rule because more information is released under the Milgrom–Weber rule than under sealed-bid ones. Note, however, that in our model, when the number of bidders is large, no difference exists between the average transaction prices under the two auction formats and pricing rules because they converge to the true value. Moreover, affiliation in signals can only reduce the variance of the transaction price under the Milgrom–Weber pricing rule. Thus, by working within the symmetric CVP, we provide an upper bound concerning the improvement in variance reduction that can obtain under the Milgrom–Weber pricing rule. In section 5, we derive the likelihood function of observed drop-out prices, while in section 6, we apply our methods to data from an auction of taxi license plates held in Shenzhen, China; our empirical work provides an estimate of the upper bound concerning the decrease in the variances of transaction prices when using the Milgrom–Weber pricing rule. In the final section, we summarize and conclude. Any details too cumbersome to be included in the text of the paper have been collected in the appendix at the end of the paper.

2 Theoretical Model

Consider an oral, ascending-price auction at which k units are for sale to a total of n bidders, each of whom wants at most one unit. Focus on the Milgrom and Weber [2000] pricing rule described in the introduction. Assume that each bidder draws an independently- and identically-distributed signal X , conditional on the true, but unknown, value v^0 . Denote the cumulative distribution and probability density functions of X , conditional on v , by $F_{X|V}(x|v)$ and $f_{X|V}(x|v)$, respectively. Denote by $f_V(v)$ the prior distribution of V , the unknown value.

Consider the vector of signals (X_1, X_2, \dots, X_n) , a random sample of n draws from $F_{X|V}(x|v^0)$. Because this environment is symmetric, without loss of generality, focus below on bidder 1. Denote by Y_i the i^{th} ordered signal of the opponents of bidder 1, so

$$Y_1 \geq Y_2 \geq \dots \geq Y_{n-1}.$$

Denote by Z_i the i^{th} order statistic for all of the X_i s, so

$$Z_1 \geq Z_2 \geq \dots \geq Z_n.$$

The auction proceeds in rounds $m = n, n - 1, \dots, k + 1$. In round m , m bidders continue to participate in the auction. The auction ends in round $(k + 1)$ when the $(k + 1)^{\text{st}}$ bidder exits the auction. With no loss of generality, suppose that bidders are ordered in the reverse order of exit from the auction.

Let Ω_m denote the information that has already been revealed in round m by all the bidders who have already left the auction. Hence, Ω_m equals $\{z_n, z_{n-1}, \dots, z_{m+1}\}$, where Ω_n is the empty set \emptyset . According to Milgrom and Weber [2000], the symmetric equilibrium bidding rule in round m can be written as

$$\beta_m(x) = \mathcal{E}[V|X_1 = Y_k = \dots = Y_{m-1} = x, \Omega_m] \quad (1)$$

where \mathcal{E} denotes the expectation operator. Here, Y_k, \dots, Y_{m-1} denote the k^{th} through $(m - 1)^{\text{st}}$ order

statistics among the bidders who remain competing with bidder 1. On the other hand, the order statistics in the event Ω_m denote the order statistics for *all* the bidders who have exited the auction. For completeness, we describe below our reasoning behind a characterization of the equilibrium; in their paper, Milgrom and Weber [2000] presumably omitted an argument like this because they found it obvious.

At price p , bidder 1 is concerned with the event that Y_k, \dots, Y_{m-1} all drop-out simultaneously at $\beta_m^{-1}(p)$, where $\beta_m^{-1}(p)$ is the inverse bid function. In this event, bidder 1 will be one of the winners of the auction, together with his remaining $(k - 1)$ competitors. Bidder 1 should stay in the auction at price level p if and only if

$$\mathcal{E} \left[V | X_1 = x, Y_k = \dots = Y_{m-1} = \beta_m^{-1}(p), \Omega_m \right] > p.$$

In equilibrium, $p = \beta_m(x)$, so $x = \beta_m^{-1}(p)$ —the price level at which bidder 1 should exit—should satisfy the relation that

$$p = \mathcal{E} \left[V | X_1 = x, Y_k = \dots = Y_{m-1} = \beta_m^{-1}(p), \Omega_m \right]. \quad (2)$$

Hence, the functional form of the bid function.

The winning price corresponds to the bid submitted by the bidder with the $(k+1)^{\text{st}}$ order statistic of the signals during round $(k + 1)$. Hence,

$$\hat{p} = \mathcal{E} [V | Z_k = Z_{k+1} = z_{k+1}, \Omega_{k+1}]. \quad (3)$$

3 Limiting Information in the Transaction Price

In this section, and the next, we have two goals: first, to study the convergence rate of the transaction price \hat{p} to the unknown true common value v^0 ; and, second, to characterize the limiting distribution of the transaction price \hat{p} . In both of these endeavours, we assume that n gets large,

tends to infinity.

In this regard, we make the following assumption concerning k , the number of units for sale relative to n , the number of bidders at the auction.

Assumption 1 $[(n - k)/n] \rightarrow \tau$, where τ is strictly between 0 and 1.

In words, the proportion of demand met has a stable limit as the numbers of bidders gets large. Were this not the case, then the transaction price would not have a stable limit, without some potentially unrealistic assumptions concerning $f_V(v)$, as was pointed out by Pesendorfer and Swinkels [1997].

With regard to our goals, we proceed in two steps. In the first, we define \hat{v} , the maximum-likelihood estimator (MLE) of v^0 , based on the unobserved (to the researcher, but known to the participants) order statistics z_{k+1}, \dots, z_n , and then we investigate the rate at which \hat{v} converges to v^0 . In the second, we investigate the rate at which \hat{p} converges to \hat{v} . In the next section, we demonstrate formally that the rate of convergence of the price \hat{p} to the true common value v^0 will be driven (dominated) by the convergence rate of \hat{v} to v^0 . In other words, $(\hat{p} - v^0)$ is $o_p(\hat{v} - v^0)$. Therefore, to understand the rate of information aggregation, it suffices to focus on how \hat{v} approaches v^0 as the “sample size” n gets large.

Under our assumptions, the MLE \hat{v} is defined as

$$\hat{v} = \underset{v}{\operatorname{argmax}} \log \left[\binom{n}{k} \mathcal{L}_n(z_{k+1}, \dots, z_n | v) \right]$$

where the joint likelihood function of all the signals revealed under the Milgrom–Weber auction is proportional to

$$\mathcal{L}_n(z_{k+1}, \dots, z_n | v) = [1 - F_{X|V}(z_{k+1} | v)]^k f_{X|V}(z_{k+1} | v) f_{X|V}(z_{k+2} | v) \cdots f_{X|V}(z_n | v). \quad (4)$$

Here, the term $[1 - F_{X|V}(z_{k+1} | v)]^k$ captures the fact that only limited information is known concerning the signal values of the k winners—specifically, their signals are greater than z_{k+1} . Also, $\binom{n}{k}$

captures the fact that there are many ways in which the k highest order statistics of signals could exceed z_{k+1} . Equation (4) is the joint likelihood of the lowest $(n - k)$ order statistics—those from z_{k+1} to z_n .

3.1 Convergence of Price to the True Value

Given equation (4), the average log-likelihood will be a function of the lowest $(n - k)$ order statistics. A general function of order statistics can be difficult to analyse because of the potentially complex correlation structure among order statistics. One possibility is to appeal to the theory of L -statistics to investigate the convergence properties of functions of order statistics. Fortunately, this particular average log-likelihood function is more tractable than an L -statistic because it can be rewritten as a function of the entire sample and the sample τ^{th} quantile. Specifically,

$$\begin{aligned}\hat{Q}_n(v) &= \frac{1}{n} \log \mathcal{L}_n(z_{k+1}, \dots, z_n | v) \\ &= \frac{k}{n} \log \left(1 - F_{X|V} \left[\hat{F}_n^{-1}(\tau) | v \right] \right) + \frac{1}{n} \sum_{i=1}^n \log f_{X|V}(X_i | v) \mathbf{1} \left[X_i \leq \hat{F}_n^{-1}(\tau) \right]\end{aligned}$$

where $\hat{F}_n(\cdot)$ denotes the empirical distribution function

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(X_i \leq x).$$

Now, under the assumptions made formal below, the sample percentile $\hat{F}_n^{-1}(\tau)$ converges in probability to the true population quantile $F_{X|V}^{-1}(\tau | v^0)$, by a uniform law of large numbers, so $\hat{Q}_n(v)$ converges uniformly in the parameter space of v to a deterministic function

$$Q^0(v) = (1 - \tau) \log \left(1 - F_{X|V} \left[F_{X|V}^{-1}(\tau | v^0) | v \right] \right) + \int_{-\infty}^{F_{X|V}^{-1}(\tau | v^0)} f_{X|V}(x | v^0) \log f_{X|V}(x | v) dx$$

where the superscript 0 is used to denote the dependence on v^0 : $Q^0(v) \equiv Q(v^0, v)$.

In order for information to aggregate fully, $Q^0(v)$, as a function of v , must be uniquely maxi-

mized at v equal v^0 . As in the case of full-sample likelihood function, this can be verified using Jensen's inequality. Thus, for any v not equal to v^0 , $Q^0(v) \leq Q^0(v^0)$. This can be shown by taking the sum of the following two inequalities. First, by Jensen's inequality,

$$\begin{aligned} & \int_{-\infty}^{F_{X|V}^{-1}(\tau|v^0)} f_{X|V}(x|v^0) \log f_{X|V}(x|v) dx - \int_{-\infty}^{F_{X|V}^{-1}(\tau|v^0)} f_{X|V}(x|v^0) \log f_{X|V}(x|v^0) dx \\ & \leq \tau \left[\log \int_{-\infty}^{F_{X|V}^{-1}(\tau|v^0)} f_{X|V}(x|v) dx - \log \tau \right]. \end{aligned}$$

Second, it is easy to see that

$$(1 - \tau) \log \left(1 - F_{X|V} \left[F_{X|V}^{-1}(\tau|v^0) | v \right] \right) + \tau \log F_{X|V} \left[F_{X|V}^{-1}(\tau|v^0) | v \right] \leq (1 - \tau) \log (1 - \tau) + \tau \log \tau$$

because the left-hand side, considered as a function of $F_{X|V} \left[F_{X|V}^{-1}(\tau|v^0) | v \right]$, is maximized at τ .

Assumption 2 For $v \neq v^0$, either $F_{X|V}^{-1}(\tau|v) \neq F_{X|V}^{-1}(\tau|v^0)$ or, with positive probability, $X \leq F_{X|V}^{-1}(\tau|v^0)$ under v^0 , $f_{X|V}(x|v) \neq f_{X|V}(x|v^0)$.

This assumption mirrors a standard full-sample identification condition for likelihood analysis. While the monotone likelihood-ratio condition used by Milgrom and Weber [1982] is required to derive the equilibrium bidding strategy, conditional on the form of the equilibrium bidding strategy, it is not strictly necessary for full information aggregation to hold.

The first inequality will be strict under the first condition in Assumption 2. Likewise for the second inequality under the second condition in Assumption 2. Thus, we have demonstrated that $Q^0(v)$ is globally and uniquely maximized at v^0 provided the value v identifies the signal distribution $f_{X|V}(x|v)$ in the sense of Assumption 2, which is stronger than the usual full-sample identification condition whenever $\tau < 1$. The usual Jensen's inequality argument for full-sample likelihood function is just a special case of the above when τ is one.

Now, examine the following first-order condition at v^0 :

$$\begin{aligned}
\left. \frac{\partial Q^0(v)}{\partial v} \right|_{v=v^0} &= - \left[\frac{(1-\tau)}{1 - F_{X|V}[F_{X|V}^{-1}(\tau|v^0)|v]} \right] \left\{ \frac{\partial F_{X|V}[F_{X|V}^{-1}(\tau|v^0)|v]}{\partial v} \right\} \bigg|_{v=v^0} + \int_{-\infty}^{F_{X|V}^{-1}(\tau|v^0)} \frac{\partial f_{X|V}(x|v^0)}{\partial v} dx \\
&= - \left\{ \frac{\partial F_{X|V}[F_{X|V}^{-1}(\tau|v^0)|v]}{\partial v} \right\} \bigg|_{v=v^0} + \int_{-\infty}^{F_{X|V}^{-1}(\tau|v^0)} \frac{\partial f_{X|V}(x|v^0)}{\partial v} dx \\
&= 0.
\end{aligned}$$

Therefore, subject to the regularity conditions, which are outlined completely in the next section, \hat{v} is a consistent estimator of v^0 . To wit, $(\hat{v} - v^0) \xrightarrow{P} 0$.

Given that $Q^0(v)$ is a properly-defined averaged log-likelihood function that depends linearly on the observed sample up to a given sample quantile and that the central sample quantiles are \sqrt{n} -consistent as well as distributed asymptotically normal, the information equality then holds for v , and is related to the asymptotic variance of \hat{v} . Given the form of $Q^0(v)$, the expected Hessian is $\left. \frac{\partial^2}{\partial v^2} Q^0(v) \right|_{v=v^0}$ where

$$\begin{aligned}
\left. \frac{\partial^2}{\partial v^2} Q^0(v) \right|_{v=v^0} &= - \frac{\partial^2}{\partial v^2} F_{X|V}[F_{X|V}^{-1}(\tau|v^0)|v] - \frac{1}{1-\tau} \left(\frac{\partial}{\partial v} F_{X|V}[F_{X|V}^{-1}(\tau|v^0)|v] \right)^2 + \int_0^\tau \frac{\partial^2}{\partial v^2} \log f_{X|V}[F_{X|V}^{-1}(u|v^0)|v] du.
\end{aligned} \tag{5}$$

3.2 Information-Matrix Equality

In full-sample likelihood models, the asymptotic variance of the maximum-likelihood estimator is usually calculated using an information-matrix equality. Here, we show that an analogous information-matrix equality also holds for the *partial sample* information model that we consider, which we shall use to characterize the amount of limiting information contained in the price as an estimate of the true value.

One approach to calculating the information matrix equality is to view the limiting first-order

condition at v^0 as an identity, and then totally differentiate it with respect to v^0 . Specifically, because

$$\left. \frac{\partial}{\partial v} Q^0(v) \right|_{v=v^0} = 0,$$

then

$$\left. \frac{\partial}{\partial v^0} \left[\frac{\partial}{\partial v} Q^0(v) \right] \right|_{v=v^0} = 0.$$

This can be written as

$$\left. \frac{\partial^2}{\partial v^2} Q^0(v) \right|_{v=v^0} + \left. \frac{\partial}{\partial u} \left[\frac{\partial}{\partial v} Q^u(v) \right] \right|_{u=v=v^0} = 0. \quad (6)$$

In the next section, the second term on the left-hand side, which is the negative of the Hessian given in equation (5), will be shown to equal the asymptotic variance of the score function. The following provides a direct calculation of the second term in equation (6), which independently verifies equation (6) and facilitates the comparison with the variance of the score function in the next section.

To compute this term, we need to calculate

$$\frac{\partial}{\partial v} F_{X|V}^{-1}(\tau|v)$$

as well as

$$\left. \frac{\partial}{\partial v} F_{X|V} \left[F_{X|V}^{-1}(\tau|v) | v^0 \right] \right|_{v=v^0} = \frac{\partial}{\partial v} F_{X|V}^{-1}(\tau|v) f_{X|V} \left[F_{X|V}^{-1}(\tau|v^0) | v \right].$$

Both can be found by totally differentiating the identity

$$\int_{-\infty}^{F_{X|V}^{-1}(\tau|v)} f_{X|V}(x|v) \, dx = \tau,$$

which leads to

$$\frac{\partial}{\partial v} F_{X|V}^{-1}(\tau|v) = -\frac{\frac{\partial}{\partial v} F_{X|V} [F_{X|V}^{-1}(\tau|v^0)|v]}{f_{X|V} [F_{X|V}^{-1}(\tau|v^0)|v]}, \quad \frac{\partial}{\partial v} F_{X|V} [F_{X|V}^{-1}(\tau|v)|v^0] \Big|_{v=v^0} = -\frac{\partial}{\partial v} F_{X|V} [F_{X|V}^{-1}(\tau|v^0)|v] \Big|_{v=v^0}.$$

Using these relations,

$$\frac{\partial}{\partial u} \left[\frac{\partial}{\partial v} Q^u(v) \right] \Big|_{u=v=v^0} = \frac{1}{1-\tau} \left(\frac{\partial}{\partial v} F_{X|V} (F_{X|V}^{-1}(\tau|v^0)|v) \right)^2 + \int_{-\infty}^{F_{X|V}^{-1}(\tau|v^0)} \left[\frac{\partial}{\partial v} \log f_{X|V}(x|v) \right]^2 f_{X|V}(x|v^0) dx \Big|_{v=v^0}.$$

The next section formally demonstrates that the log-likelihood function of the partially observed sample in our model has a similar statistical behaviour to the usual full-sample log-likelihood function, so $\sqrt{n}(\hat{v} - v^0)$ will converge in distribution to a normal random variable whose asymptotic variance is the inverse of either $\frac{\partial}{\partial u} \left[\frac{\partial}{\partial v} Q^u(v) \right] \Big|_{u=v=v^0}$ or equivalently $\frac{\partial^2}{\partial v^2} Q^0(v) \Big|_{v=v^0}$. We now need to show that $\sqrt{n}(\hat{p} - v^0)$ is $o_p(1)$ because, then, these will also represent the asymptotic variance of $\sqrt{n}(\hat{p} - v^0)$.

For this purpose, we employ Bayesian asymptotic analysis. First, note that

$$\hat{p} = \beta_{k+1}(z_{k+1}) = \int v \frac{f_{Z|V}(Z_k = Z_{k+1} = z_{k+1}, \Omega_{k+1}|v) f_V(v)}{\int f_{Z|V}(Z_k = Z_{k+1} = z_{k+1}, \Omega_{k+1}|u) f_V(u) du} dv$$

where the likelihood of the conditioning event in the bid function is proportional to

$$f_{Z|V}(Z_k = Z_{k+1} = z_{k+1}, \Omega_{k+1}|v) = [1 - F_{X|V}(z_{k+1}|v)]^{k-1} f_{X|V}(z_{k+1}|v)^2 f_{X|V}(z_{k+2}|v) \cdots f_{X|V}(z_n|v). \quad (7)$$

Recall the definition in equation (4)

$$\begin{aligned} f_{V|Z}(Z_k = Z_{k+1} = z_{k+1}, \Omega_{k+1}|v) &= \mathcal{L}_n(z_{k+1}, \dots, z_n|v) \frac{f_{X|V}(z_{k+1}|v)}{[1 - F_{X|V}(z_{k+1}|v)]} \\ &= \exp[n\hat{Q}_n(v)] \frac{f_{X|V}(z_{k+1}|v)}{[1 - F_{X|V}(z_{k+1}|v)]}, \end{aligned}$$

which we can write, using a change of variables,

$$\sqrt{n}(\hat{p} - \hat{v}) = \int h \frac{\exp[n(\hat{Q}_n(\hat{v} + h/\sqrt{n}) - \hat{Q}_n(\hat{v}))] f_V(\hat{v} + h/\sqrt{n}) \frac{f_{X|V}(z_{k+1}|\hat{v}+h/\sqrt{n})}{[1-F_{X|V}(z_{k+1}|\hat{v}+h/\sqrt{n})]}}{\int \exp[n(\hat{Q}_n(\hat{v} + u/\sqrt{n}) - \hat{Q}_n(\hat{v}))] f_V(\hat{v} + u/\sqrt{n}) \frac{f_{X|V}(z_{k+1}|\hat{v}+u/\sqrt{n})}{[1-F_{X|V}(z_{k+1}|\hat{v}+u/\sqrt{n})]}} du dh.$$

It is demonstrated in the next section that the above renormalized posterior distribution is asymptotically normal. Intuitively, $\sqrt{n}(\hat{p} - \hat{v}) \xrightarrow{p} 0$ obtains because the mean of the above renormalized posterior distribution is asymptotically zero. It is also clear that the single-unit English-auction model investigated by Milgrom and Weber [1982] is a special case of this result—when τ is one, which corresponds to the conventional full-sample maximum-likelihood analysis and Bayesian posterior distribution. At a typical English auction, where τ is one, the only difference from full-sample maximum-likelihood analysis is that the maximum order statistic is unobserved. However, a single order statistic is asymptotically negligible. Likewise, the conditioning event in the bid function in equation (7) differs from the corresponding partial-sample likelihood in equation (4) only by a single order statistic and the difference is asymptotically negligible.

3.3 Simple Example

Consider the following example, which can be solved in closed-form. Suppose that the conditional distribution of X is exponential, having mean v , so

$$f_{X|V}(x|v) = \frac{1}{v} \exp(-x/v) \quad \text{for } x \geq 0, v > 0.$$

The posterior distribution needed to compute the bid function in equation (1) is proportional to

$$\begin{aligned} \left(\frac{1}{v}\right)^{n-m} \exp\left(-\frac{1}{v} \sum_{j=n}^{m+1} z_j\right) \left(\frac{1}{v}\right)^{m-k+1} \exp(-z/v)^{m-k+1} \exp(-z/v)^{k-1} f_V(v) = \\ \left(\frac{1}{v}\right)^{n-k+1} \exp\left[-\frac{1}{v} \left(\sum_{j=n}^{m+1} z_j + mz\right)\right] f_V(v) \end{aligned}$$

Suppose $f_V(v)$ is a diffuse prior.¹ In this case, the above posterior distribution is then an inverse gamma distribution having parameters $(n - k + 1)$ and $(\sum_{j=n}^{m+1} z_j + mz)$, which has mean

$$\mathcal{E}[V|X_1 = Y_k = \dots = Y_{m-1} = z, \Omega_m] = \frac{\sum_{j=n}^{m+1} z_j + mz}{n - k},$$

which is also the bid function at round m . Therefore, the transaction price is given by the bid function with m equal $(k + 1)$ and z equal z_{k+1} :

$$\hat{p} = \frac{\sum_{j=n}^{k+2} z_j + (k + 1)z_{k+1}}{n - k}.$$

To see why \hat{p} converges to the true v^0 , note that in this example, $Z_{k+1} \xrightarrow{p} F_{X|V}^{-1}(\tau)$ which equals $-v^0 \log(1 - \tau)$. Also, by invoking a law of large numbers,

$$\frac{\sum_{j=n}^{k+2} z_j}{n} \xrightarrow{p} v^0 \left[\frac{\log(1 - \tau)}{1 - \tau} + \tau \right].$$

Therefore,

$$\hat{p} \xrightarrow{p} v^0 \frac{1}{\tau} [(\log(1 - \tau)(1 - \tau) + \tau)] - v^0 \log(1 - \tau) \frac{1 - \tau}{\tau} = v^0.$$

The maximum-likelihood estimator \hat{v} , which is the mode of the posterior distribution, is

$$\hat{v} = \frac{\sum_{j=n}^{k+2} z_j + (k + 1)z_{k+1}}{n - k + 2}.$$

Hence,

$$\hat{v} = \frac{n - k}{n - k + 2} \hat{p}.$$

It can then be verified that

$$\sqrt{n}(\hat{p} - \hat{v}) \xrightarrow{p} 0.$$

¹We could also use a Pareto prior, but this would clutter the calculations considerably.

4 Asymptotic Distribution of Transaction Price

In this section, we provide formal conditions to justify the claims made in the previous section. Our analysis is broken into two parts: in the first, we derive the asymptotic distribution of $\sqrt{n}(\hat{v} - v^0)$, while in the second we show that $\sqrt{n}(\hat{p} - \hat{v})$ is $o_p(1)$. As Newey and McFadden [1994] as well as Chernozhukov and Hong [2003] have pointed out, both parts depend on the stochastic equicontinuity properties of the sample-averaged log-likelihood function $\hat{Q}_n(v)$.

To begin, we state assumptions sufficient to the task. Instead of striving for the weakest possible set of assumptions, we are content with potentially overly-strong sufficient conditions that illustrate the main results. Note, too, that in theoretical models of auctions, the monotone likelihood-ratio condition is typically imposed, which restricts how weak the conditions for equicontinuity can be.

Assumption 3 *The support of the prior distribution $f_V(v)$ is a compact closed interval, and the true common value v^0 is contained in the interior of that interval.*

Assumption 4 *The support of $f_{X|V}(x|v)$ is independent of v and bounded, while $\log f_{X|V}(x|v)$ is uniformly bounded, having bounded continuous third derivatives in both arguments on its support.*

These two assumptions are regularity conditions required to demonstrate uniform convergence and stochastic equicontinuity.

Theorem 1 *Under Assumptions 1 to 4, if $f_V(v)$ is continuous at v^0 with a finite mean, then*

$$\sqrt{n}(\hat{v} - v^0) \xrightarrow{d} \mathcal{N}\left(0, \Sigma(\tau) = \frac{\partial}{\partial u} \left[\frac{\partial}{\partial v} Q^u(v) \right] \Big|_{u=v=v^0}^{-1}\right),$$

and

$$\sqrt{n}(\hat{p} - \hat{v}) \xrightarrow{p} 0,$$

so

$$\sqrt{n}(\hat{p} - v^0) \xrightarrow{d} \mathcal{N}[0, \Sigma(\tau)].$$

Remark 1: In the Pesendorfer–Swinkels model, under the same assumption $(k/n) \rightarrow (1 - \tau)$, only the signal of a single last-losing bidder is revealed, instead of the signals of all the losing bidders. Therefore, intuitively, the transaction price in the Pesendorfer–Swinkels model should aggregate less information than that in the Milgrom–Weber model. In fact, this turns out to be true. While the prices in both the Pesendorfer–Swinkels and the Milgrom–Weber models converge to v^0 at rate \sqrt{n} , the asymptotic variance of the Pesendorfer–Swinkels price is greater than the Milgrom–Weber price. We demonstrate this result formally using the influence function representation of the asymptotic variance. We note, first, from the proof of the theorem that $\Sigma(\tau)$ equals $\text{Var}[\psi_1(X, \tau)]^{-1}$, where the *influence function* $\psi_1(X, \tau)$ is given by

$$\begin{aligned} \psi_1(X, \tau) \equiv & \quad \frac{\partial}{\partial v} \log f_{X|V}(X|v^0) \mathbf{1}[X \leq F_{X|V}^{-1}(\tau|v^0)] - \\ & \quad \mathcal{E}\left(\frac{\partial}{\partial v} \log f_{X|V}(X|v^0) \mathbf{1}[X \leq F_{X|V}^{-1}(\tau|v^0)]\right) + \\ & \quad \frac{1}{1-\tau} \left(\frac{\partial}{\partial v} F_{X|V}[F_{X|V}^{-1}(\tau|v)|v]\right) (\mathbf{1}[X \leq F_{X|V}^{-1}(\tau|v)|v] - \tau). \end{aligned}$$

Next, we characterize the average log-likelihood function as well as the score and influence functions in the Pesendorfer–Swinkels model, and show that they imply a variance larger than $\Sigma(\tau)$.

The average log-likelihood of the Pesendorfer–Swinkels model, which depends only on a single order statistic $z_{k+1} = \hat{F}_n^{-1}(\tau)$, is given by

$$\begin{aligned} \tilde{Q}_n(v) &= \quad \frac{k}{n} \log[1 - F_{X|V}(z_{k+1}|v)] + \left(1 - \frac{k}{n}\right) \log F_{X|V}(z_{k+1}|v) \\ &= \quad (1 - \tau) \log\left(1 - F_{X|V}[\hat{F}_n^{-1}(\tau)|v]\right) + \tau \log F_{X|V}[\hat{F}_n^{-1}(\tau)|v]. \end{aligned}$$

Its corresponding score function is

$$\frac{\partial}{\partial v} \tilde{Q}_n(v) = -\left(\frac{1-\tau}{1-F_{X|V}[\hat{F}_n^{-1}(\tau)|v]} - \frac{\tau}{F_{X|V}[\hat{F}_n^{-1}(\tau)|v]}\right) \frac{\partial}{\partial v} F_{X|V}[\hat{F}_n^{-1}(\tau)|v].$$

If we evaluate the first-order approximation of the score function with respect to $\hat{F}_n^{-1}(\tau)$ as it approaches $F_{X|V}^{-1}(\tau|v^0)$ at $v = v^0$, and make use of the well-established asymptotic approximation

of the sample quantile

$$\sqrt{n} \left[\hat{F}_n^{-1}(\tau) - F_{X|V}^{-1}(\tau|v^0) \right] = -\frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\mathbf{1} \left[X_i \leq F_{X|V}^{-1}(\tau|v^0) \right] - \tau}{f_{X|V} \left[F_{X|V}^{-1}(\tau|v^0) \right]} + o_p(1), \quad (8)$$

then we find the following influence function representation for the Pesendorfer–Swinkels score function:

$$\sqrt{n} \frac{\partial}{\partial v} \tilde{Q}_n(v^0) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \psi_2(X_i, \tau) + o_p(1)$$

where

$$\psi_2(X, \tau) \equiv \frac{1}{\tau(1-\tau)} \frac{\partial}{\partial v} F_{X|V} \left[F_{X|V}^{-1}(\tau|v) | v \right] \left(\mathbf{1} \left[X \leq F_{X|V}^{-1}(\tau|v^0) \right] - \tau \right).$$

Letting \tilde{p} to denote the transaction price in the Pesendorfer–Swinkels auction model, we have

$$\sqrt{n}(\tilde{p} - v^0) \xrightarrow{d} \mathcal{N} \left(0, \text{Var} [\psi_2(X, \tau)]^{-1} \right).$$

In order to show that $\text{Var} [\psi_2(X, \tau)] \leq \text{Var} [\psi_1(X, \tau)]$, we compute

$$\begin{aligned} \psi_1(X, \tau) - \psi_2(X, \tau) &= \frac{\partial}{\partial v} \log f_{X|V}(X|v^0) \mathbf{1} \left[X \leq F_{X|V}^{-1}(\tau|v^0) \right] - \\ &\quad \mathcal{E} \left(\frac{\partial}{\partial v} \log f_{X|V}(X|v^0) \mathbf{1} \left[X \leq F_{X|V}^{-1}(\tau|v^0) \right] \right) - \\ &\quad \frac{1}{\tau} \left(\frac{\partial}{\partial v} F_{X|V} \left[F_{X|V}^{-1}(\tau|v^0) \right] \right) \left(\mathbf{1} \left[X \leq F_{X|V}^{-1}(\tau|v^0) \right] - \tau \right). \end{aligned}$$

We can then easily verify that

$$\text{Cov} [\psi_1(X, \tau) - \psi_2(X, \tau), \psi_2(X, \tau)] = 0.$$

Hence,

$$\text{Var} [\psi_1 (X, \tau)] \geq \text{Var} [\psi_2 (X, \tau)].$$

Remark 2: Above, we have indexed the asymptotic variance by τ , the proportion of losing bidders. Intuitively, the larger the fraction of losing bidders, the more information revealed at the auction. Therefore, we expect $\Sigma(\tau)$ to be a monotonically decreasing function of τ , in a matrix sense. In other words, for any $0 < \tau_1 \leq \tau_2 \leq 1$,

$$\Sigma(\tau_1) \geq \Sigma(\tau_2).$$

This also turns out to be true. To wit, for $\tau_1 > \tau_2$, $\text{Var} [\psi_1 (X, \tau_1)] \leq \text{Var} [\psi_1 (X, \tau_2)]$. This will, in turn, follow from

$$\text{Cov} [\psi_1 (X, \tau_1) - \psi_1 (X, \tau_2), \psi_1 (X, \tau_1)] = 0. \quad (9)$$

Verifying equation (9) is tedious, but straightforward: it depends on the following two key relations. First,

$$\mathcal{E} \left(\frac{\partial}{\partial v} \log f_{X|V} (X|v^0) \mathbf{1} [X \leq F_{X|V}^{-1} (\tau|v^0)] \right) = \frac{\partial}{\partial v} F_{X|V} [F_{X|V}^{-1} (\tau|v) | v^0]$$

and, second, that, for $\tau_1 > \tau_2$,

$$\mathbf{1} [X \leq F_{X|V}^{-1} (\tau_1|v) | v^0] \mathbf{1} [X \leq F_{X|V}^{-1} (\tau_2|v) | v^0] = \mathbf{1} [X \leq F_{X|V}^{-1} (\tau_2|v) | v^0].$$

Hence, under the assumptions made above, especially the common support Assumption 4, for $0 < \tau \leq 1$, the larger is τ , the more information is aggregated in the Milgrom–Weber model, in the sense of having a smaller variance despite that the rate of convergence stays the same. It can also be shown that this conclusion continues to hold without the support Assumption 4. When

the upper support is increasing in v , while the condition still holds, the rate of convergence can improve beyond \sqrt{n} when τ equals one. On the other hand, if the lower support is also increasing in v , then it is possible that the convergence rate is faster than \sqrt{n} even when τ is zero. In this case, while there will be no information loss when τ increases above zero, there may be no additional asymptotic information either until τ becomes one.

This desirable monotonicity property of information aggregation in the Milgrom–Weber model is in contrast to the Pesendorfer–Swinkels model. The amount of information aggregated asymptotically in the price of the Pesendorfer–Swinkels model is not monotonic in τ . For example, when $f_{X|V}(x|v^0)$ is uniform in X , the worst τ for information aggregation is one-half in the Pesendorfer–Swinkels model, because this involves the worst balance between the winner’s curse and the loser’s curse. In general, the optimal τ in the Pesendorfer–Swinkels model obviously depends on the shape of this conditional density. Intuitively, in the Pesendorfer–Swinkels model, a different τ selects a different information set, while in the Milgrom–Weber model, a larger τ always selects a larger information set.

5 Deriving Likelihood Function of Observed Drop-Out Prices

In section 2, we derived the bid function of a representative bidder as well as characterized the transaction price; see equations (1) and (3). In sections 3 and 4, we then demonstrated that the transaction price converged in probability to the true, unknown value v^0 and derived its asymptotic distribution. To provide a framework within which to conduct our empirical analysis in section 6, in this section, we derive the likelihood function of the bid data observed by an econometrician. We highlight the fact that the sampling variability of the econometrician’s estimate of the true, but unknown value v^0 will depend on nuisance parameters unknown to the econometrician.

We first introduce some additional notation. We denote by \hat{p}_j the j^{th} drop-out price, so $j = 1, 2, \dots, n - k$. For example, in our empirical application, we have n equal forty bidders and k equal twenty units, so there are twenty drop-out prices, the last being the transaction price, which

we denoted above by \hat{p} , but now denote as \hat{p}_{n-k} . Thus, our observables are $(\hat{p}_1, \hat{p}_2, \dots, \hat{p}_{n-k-1}, \hat{p}_{n-k})$. Now, from equation (2), we can recover the signal consistent with the first bidder's drop-out price—viz.,

$$\tilde{z}_n = \beta_n^{-1}(\hat{p}_1).$$

Likewise, for each of $j = 2, 3, \dots, n-k$, we can recursively recover \tilde{z}_j , the signals of the $(n-k-1)$ losing bidders, so

$$\tilde{z}_{n-j+1} = \beta_{n-j+1}^{-1}(\hat{p}_j; \Omega_{n-j+1}).$$

For the k bidders who win the auction, all we know is that Z_j exceeds $\beta_{n-(n-k-1)}^{-1}(\hat{p}_{n-k}; \Omega_{k+1})$.

In the general case, the bid function $\beta_m(x)$ in equation (1) takes the following form:

$$\beta_m(x; \Omega_m) = \int v \frac{f_V(v)g(X_1 = Y_k = \dots = Y_{m-1} = x, \Omega_m|v)}{\int f_V(u)g(X_1 = Y_k = \dots = Y_{m-1} = x, \Omega_m|u)} dv$$

where

$$g(X_1 = Y_k = \dots = Y_{m-1} = x, \Omega_m|v) = [1 - F_{X|V}(x|v)]^{k-1} f_{X|V}(x|v)^{m-k+1} \prod_{j=n}^{m+1} f_{X|V}(z_j|v).$$

If we assume that $f_V(v)$ is diffuse and that X given v is normal, having mean v and variance σ^2 , then we can write

$$g(X_1 = Y_k = \dots = Y_{m-1} = x, \Omega_m|v) = \left[1 - \Phi\left(\frac{x-v}{\sigma}\right)\right]^{k-1} \frac{1}{\sigma^{m-k+1}} \phi\left(\frac{x-v}{\sigma}\right)^{m-k+1} \frac{1}{\sigma^{n-m}} \prod_{j=n}^{m+1} \phi\left(\frac{z_j-v}{\sigma}\right).$$

To summarize, under the assumptions of normality as well as a diffuse prior,

$$\beta_m(x; \Omega_m) = \int v \frac{\left[1 - \Phi\left(\frac{x-v}{\sigma}\right)\right]^{k-1} \cdot \phi\left(\frac{x-v}{\sigma}\right)^{m-k+1} \prod_{j=n}^{m+1} \phi\left(\frac{z_j-v}{\sigma}\right)}{\int \left[1 - \Phi\left(\frac{x-u}{\sigma}\right)\right]^{k-1} \phi\left(\frac{x-u}{\sigma}\right)^{m-k+1} \prod_{j=n}^{m+1} \phi\left(\frac{z_j-u}{\sigma}\right)} du dv.$$

Consider $(\tilde{z}_n, \tilde{z}_{n-1}, \dots, \tilde{z}_{k+1})$, the vector of $(n-k)$ signals consistent with the observed drop-out

prices as well as the transaction price. The joint likelihood function of all the signals consistent with the drop-out prices revealed under the Milgrom–Weber auction is

$$\begin{aligned} \tilde{\mathcal{L}}_n(\tilde{z}_{k+1}, \dots, \tilde{z}_n | v, \boldsymbol{\theta}) &= [1 - F_{X|V}(\tilde{z}_{k+1} | v, \boldsymbol{\theta})]^k f_{X|V}(\tilde{z}_{k+1} | v, \boldsymbol{\theta}) \left| \frac{\partial \beta_{n-k}^{-1}(\hat{p}_{n-k})}{\partial \hat{p}_{n-k}} \right| \\ & \quad f_{X|V}(\tilde{z}_{k+2} | v, \boldsymbol{\theta}) \left| \frac{\partial \beta_{n-k-1}^{-1}(\hat{p}_{n-k-1})}{\partial \hat{p}_{n-k-1}} \right| \cdots f_{X|V}(\tilde{z}_n | v, \boldsymbol{\theta}) \left| \frac{\partial \beta_1^{-1}(\hat{p}_1)}{\partial \hat{p}_1} \right| \end{aligned} \quad (10)$$

Here, $\boldsymbol{\theta}$ denotes a vector of unknown parameters, and captures the fact the probability density and cumulative distribution functions of signals can depend on parameters known to the bidders, but unknown to the econometrician.

The econometrician's MLE \tilde{v} is defined as

$$\tilde{v} = \underset{v}{\operatorname{argmax}} \log \left[\binom{n}{k} \tilde{\mathcal{L}}_n(\tilde{z}_{k+1}, \dots, \tilde{z}_n | v, \tilde{\boldsymbol{\theta}}) \right]$$

where $\tilde{\boldsymbol{\theta}}$ denotes the MLE of $\boldsymbol{\theta}^0$. While knowing the true nuisance parameters in $\boldsymbol{\theta}^0$ is unimportant in demonstrating that the transaction price converges in probability to the true, unknown value v^0 , because the parameters contained in $\boldsymbol{\theta}^0$ are of second-order importance, the nuisance parameters are critical when calculating an estimate of the sampling variation in \tilde{v} , the econometrician's estimate of the true, but unknown value v^0 .

6 Empirical Application

We have applied the methods described above to data from an auction of taxi license plates held in Shenzhen, China in October 2007. At this auction, the municipal transportation bureau sold 2,000 additional *red taxi* license plates. Red taxis are special in Shenzhen because they can operate both inside and outside the Special Economic Zone (SEZ), unlike *yellow taxis* which can operate only inside the SEZ, and *green taxis* which can only operate outside the SEZ.

The city of Shenzhen had not issued any new license plates for red taxis since 1993. However, rapid growth in Shenzhen's population meant that patrons were experiencing a shortage of taxis,

leading to an increase in the number of illegally-operated taxis. In 2007, the per capita number of taxis in Shenzhen was low when compared to other parts of China: only 10,305 taxis were licensed in a city of 7.5 million permanent residents, about 13.74 taxis for every 10,000 residents. The Ministry of Construction in China recommended that cities should have 21 taxis for every 10,000 residents.

Before the auction, the authorities reviewed the qualifications of all those who had applied to participate at the auction. Potential bidders could be individual taxi companies or groups formed by different companies. While fifty-one ‘firms’ apparently requested to participate, only forty potential bidders were certified to participate at the auction. Thus, n was 40.

In written documentation, potential bidders were reminded to be aware of the risks involved. For example, consider a translation of the text from one document:

Following this auction, more taxi license plates will be issued through auction or other ways over the next four years. The number of taxis in Shenzhen will reach about 20,000 by 2011. The issuance of a great number of license plates might have much impact on the taxi industry.

Despite these warnings, representatives of taxi companies in the city showed great interest in the auction, perhaps because operating a taxi has been one of the highest profit margins in the transportation industry. Also, historically, taxis have provided a stable return against investment.

Before the auction, 53 out of 73 taxi companies in Shenzhen owned between 50 to 200 taxis each. To wit, the majority of the city’s taxi companies were small- and medium-sized ones. Some incumbent taxi drivers expressed concern that entry would erode profits. One was quoted in the local newspaper (our translation) as saying that

Actually we are not earning much nowadays. If more taxis were on the road, we would have a hard time making ends meet.

In contrast, local residents supported the issuance of additional license plates. One was quoted (again our translation from the local newspaper) as saying

The sooner new taxis hit the road the better. It's too hard to hail a taxi during peak hours and holidays.

This anecdotal evidence, along with casual observation, suggests to us that the value of a red-taxi license plate in Shenzhen has a large common-value component. Before the auction, however, this common value was unknown to potential bidders. Using whatever means at their disposal, potential bidders formed estimates of the unknown common value which they then used during bidding at the auction.

The auction in Shenzhen proceeded according to the rules described in Milgrom and Weber [2000]. In written rules announced before the auction, the authorities informed potential bidders that the 2,000 license plates on sale would be distributed evenly among the final twenty highest bidders; each winner would be required to buy 100 license plates.

The auctioneer, Tian Tao, was a registered member of China's auction industry association. The reserve price was set at 150,000 yuan per license plate, but the price rose to 500,000 yuan in fourth minute of bidding. During the auction, Tian reminded bidders repeatedly to be aware of the risks involved. In fact, Tao took a break for ten minutes to allow the bidders "to cool their enthusiasm." We have translated one of his comments as "this is one of the most intensive auctions I've experienced in my career as an auctioneer." At the close of the auction, the price of a red-taxi license plate was 542,500 yuan, around US\$80,000.

In table 1, we present the prices called out during the auction along with the number of bidders who exited the auction at those prices, while in figure 1 we depict the empirical survivor function of prices. The prices in this table are in 10,000 yuan.

Zhang Hongzhi, a manager of Shenzhen Xilie Taxi Company, was reported in the newspaper to have said that he "felt very excited after we won a bid." Before his attending the auction, his company had decided on 550,000 yuan as the highest they would pay for a red-taxi license plate.

To implement equation (4), we assumed that X , conditional on v^0 , is distributed normally,

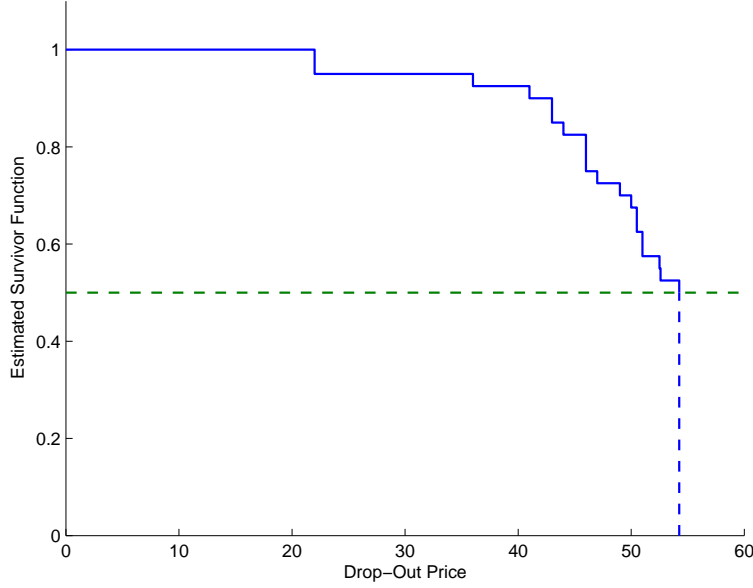


Figure 1: Estimated Survivor Function of Drop-Out Prices

having variance σ^2 , so

$$F_{X|V}(x|v^0) = \int_{-\infty}^x \frac{1}{\sigma \sqrt{2\pi}} \exp\left[-\frac{(u - v^0)^2}{\sigma^2}\right] du \equiv \Phi\left[\frac{(x - v^0)}{\sigma}\right]$$

with

$$f_{X|V}(x|v^0) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left[-\frac{(x - v^0)^2}{\sigma^2}\right] \equiv \frac{1}{\sigma} \phi\left[\frac{(x - v^0)}{\sigma}\right].$$

We also assumed that $f_V(v)$ is a diffuse prior. In table 2, we present the MLEs of v^0 and σ as well as their standard errors; the logarithm of the likelihood function for this empirical specification is -55.98 . Here, the units of the parameters estimates are in 10,000 yuan.

In order to understand the implications of these parameter estimates, we used these parameters to simulate the differences between the prices in a Milgrom–Weber auction and a Pesendorfer–Swinkels auction. A subset of these results are reported in table 3. Each entry in the table provides the difference in the expected revenue between the Milgrom–Weber auction and the Pesendorfer–Swinkels auction, measured in units of 10,000 yuan. In calculating table 3, we need three param-

eters: the prior mean and variance of the common value distribution as well as the variance of the signal distribution conditional on the common value. We used the estimate of v^0 to specify the prior mean, and the estimate of σ necessary to specify both the variance of the signal distribution and the prior variance of the value distribution. As predicted by the linkage principle of Milgrom and Weber [1982], the Milgrom–Weber auction always generates an higher expected revenue than the Pesendorfer–Swinkels auction. However, as table 3 illustrates, the difference in the expected revenues is rather small when compared to both the selling price and the estimated common value.

7 Summary and Conclusions

Using a clock model of a multi-unit, oral, ascending-price, auction, within the common-value paradigm, under the Milgrom–Weber pricing rule, we have analysed the asymptotic behaviour of the transaction price as the number of bidders and the number of units get large. We have demonstrated that even though the transaction price is determined by a (potentially small) fraction of losing drop-out bids, that price converges almost surely to the *ex ante* unknown, true value. Subsequently, we have demonstrated that the asymptotic distribution of the transaction price is Gaussian. We also demonstrated that the asymptotic variance of the transaction price under the Milgrom–Weber pricing rule is less than that under the pricing rule used by Pesendorfer and Swinkels. Thus, if the transaction prices under different auction formats and pricing rules are viewed as statistical estimators of the true, but unknown, value of the units for sale, then the Milgrom–Weber pricing rule is a more efficient estimator of the unknown value than the first-price, sealed-bid rule because more information is released under the Milgrom–Weber rule than under sealed-bid ones. Note, however, that in our model, when the number of bidders is large, no difference exists between the average transaction prices under the two auction formats and pricing rules because they converge to the true value. Moreover, affiliation in signals can only reduce the variance of the transaction price under the Milgrom–Weber pricing rule. Thus, by working within the symmetric CVP, we

have provided an upper bound concerning the improvement in variance reduction that obtains under the Milgrom–Weber pricing rule. Finally, we applied our methods to data from an auction of taxi license plates held in Shenzhen, China, finding that our estimate of the unknown, true value was not significantly different from the transaction price at size 0.01, but is at size 0.05.

Acknowledgements

We should like to acknowledge Eugene Jeong for excellent research assistance. We also thank the municipal transportation bureau in the city of Shenzhen, China for providing helpful advice and useful information as well as the Shenzhen Auction House which provided us the data from the auction of taxi license plates. We are also grateful to Srihari Govindan, Kenneth L. Judd, and Robert B. Wilson for providing us with helpful comments and useful suggestions on an earlier draft of the paper. Paarsch gratefully acknowledges that much of the research for this paper was completed while he was a visiting research scholar at the Center for Economic Institutions in the Institute of Economic Research at Hitotsubashi University in Kunitachi, Japan and that subsequent work on the paper was also completed while he was a visiting fellow at the Collegio Carlo Alberto in Moncalieri, Italy.

Appendix

To reduce clutter in the text of the paper, in this appendix, we collect the proofs of the results claimed in the text.

Proof of Main Theorem

The proof involves verifying two high-level conditions in Newey and McFadden [1994] as well as Chernozhukov and Hong [2003]. The first condition delivers consistency, while the second delivers asymptotic normality of \hat{v} and the relation that $\sqrt{n}(\hat{p} - \hat{v})$ is $o_p(1)$. We first state these

conditions within the context of our notation.

Condition 1 For any $\delta > 0$, there exists an $\epsilon > 0$, such that

$$\liminf_{n \rightarrow \infty} P_* \left\{ \sup_{|v-v^0| \geq \delta} [\hat{Q}_n(v) - \hat{Q}_n(v^0)] \leq -\epsilon \right\} = 1.$$

Condition 2 There exists $\Delta_n(v^0)$ and J^0 such that for v in an open neighbourhood of v^0 ,

$$i. \quad n [\hat{Q}_n(v) - \hat{Q}_n(v^0)] = (v - v^0) \Delta_n(v^0) - \frac{1}{2} (v - v^0)^2 [nJ^0] + R_n(v),$$

ii. For any sequence $\delta_n \rightarrow 0$,

$$\sup_{|v-v^0| \leq \delta_n} \frac{|R_n(v)|}{1 + n|v - v^0|^2} = o_p(1).$$

iii. $\Delta_n(v^0) / \sqrt{n} \xrightarrow{d} \mathcal{N}(0, \Omega^0)$, where both J^0 and Ω^0 are positive definite.

Condition 1 is, in turn, implied by uniform convergence of $\hat{Q}_n(v)$ to $Q^0(v)$ and because $Q^0(v)$ is uniquely maximized at v^0 . The unique maximum of $Q^0(v)$ at v^0 is a direct consequence of the identification Assumption 2. To show that $\sup_{v \in \mathcal{V}} |\hat{Q}_n(v) - Q^0(v)|$ is $o_p(1)$, first note that the individual terms in the summand of the second term consist of the product of $\log f_{X|V}(X_i|v)$ and $\mathbf{1}(X_i \leq \xi)$, where ξ equals $\hat{F}_n^{-1}(\tau)$. Given Assumption 4, the first is a type II function and the second is a type I function defined in Andrews [1994]. Both satisfy Pollard's entropy condition, and are stable under multiplication. Hence,

$$\sup_{v, \xi} \left| \frac{1}{n} \sum_{i=1}^n \log f_{X|V}(X_i|v) \mathbf{1}(X_i \leq \xi) - \mathcal{E}[\log f_{X|V}(X_i|v) \mathbf{1}(X_i \leq \xi)] \right| = o_p(1).$$

Next, $\mathcal{E}[\log f_{X|V}(X_i|v) \mathbf{1}(X_i \leq \xi)]$ is a Lipschitz function in ξ and the Lipschitz constant is uniform in v . Hence, given that $\hat{F}_n^{-1}(\tau) \xrightarrow{p} F_{X|V}^{-1}(\tau|v^0)$, we also have

$$\sup_v \left| \mathcal{E}(\log f_{X|V}(X_i|v) \mathbf{1}[X_i \leq \hat{F}_n^{-1}(\tau)]) - \mathcal{E}(\log f_{X|V}(X_i|v) \mathbf{1}[X_i \leq F_{X|V}^{-1}(\tau|v^0)]) \right| = o_p(1).$$

Therefore, the second term of $\hat{Q}_n(v)$ converges uniformly in v to the second term of $Q^0(v)$. The first term of $\hat{Q}_n(v)$ is also a Lipschitz function of $\hat{F}_n^{-1}(\tau)$ with the Lipschitz constant being uniform in v . Therefore, by the same argument, the first term of $\hat{Q}_n(v)$ also converges uniformly in v to the first term of $Q^0(v)$. Hence, Condition 1 holds.

The second condition is more involved than the first. We define $\hat{\xi}$ to be $\hat{F}_n^{-1}(\tau)$ where ξ^0 denotes $F_{X|V}^{-1}(\tau|v^0)$. We rewrite $\hat{Q}_n(v)$ as $\hat{Q}_n(v, \hat{\xi})$ to emphasize its direct dependence on $\hat{\xi}$. Note that, while $\hat{Q}_n(v, \hat{\xi})$ is differentiable in v , it is not in $\hat{\xi}$, so arguments relying on stochastic continuity arguments are required. The $\Delta_n(v^0)$ and J^0 elements in Condition 2 are given by

$$\Delta_n(v^0) = n \frac{\partial}{\partial v} \hat{Q}_n(v^0, \xi^0) + n \frac{\partial^2}{\partial v \partial \xi} Q^0(v^0, \xi^0) (\hat{\xi} - \xi^0)$$

and

$$J^0 = - \frac{\partial^2}{\partial v^2} Q^0(v^0, \xi^0),$$

respectively. We decompose $R_n(v)$ into $R_n^1(v) + R_n^2(v)$ with

$$R_n^1(v) = n \left[\frac{\partial}{\partial v} \hat{Q}_n(v^0, \hat{\xi}) - \frac{\partial}{\partial v} \hat{Q}_n(v^0, \xi^0) - \frac{\partial^2}{\partial v \partial \xi} Q^0(v^0, \xi^0) (\hat{\xi} - \xi^0) \right] (v - v^0)$$

and

$$R_n^2(v) = n (v - v^0)^2 \left[\frac{\partial^2}{\partial v^2} \hat{Q}_n(v^*, \hat{\xi}) - J^0 \right]$$

where v^* is a mean value between v^0 and v . Because $\frac{\partial^2}{\partial v^2} \hat{Q}_n(v^*, \hat{\xi}) - J^0 \xrightarrow{p} 0$, it follows that

$$\sup_{|v-v^0| \leq \delta_n} \frac{|R_n^2(v)|}{1 + n|v - v^0|^2} \leq \frac{|R_n^2(v)|}{n|v - v^0|^2} = o_p(1).$$

Consider, next, $R_n^1(v)$, and define $\hat{m}(v, \xi) = \frac{\partial}{\partial v} \hat{Q}_n(v, \xi)$. By stochastic equicontinuity, it follows from the entropy property of $\hat{m}(v, \xi)$ that

$$\sqrt{n} \left[\hat{m}(v^0, \hat{\xi}) - \hat{m}(v^0, \xi^0) - m(v^0, \hat{\xi}) + m(v^0, \xi^0) \right] = o_p(1).$$

Note, too, by a second-order mean-value expansion of $m(v^0, \xi)$ in ξ , that

$$\sqrt{n} \left[m(v^0, \hat{\xi}) - m(v^0, \xi^0) - \frac{\partial^2}{\partial v \partial \xi} Q^0(v^0, \xi^0) (\hat{\xi} - \xi^0) \right] = o_p(1).$$

Therefore, we can write

$$\sqrt{n} \left[\hat{m}(v^0, \hat{\xi}) - \hat{m}(v^0, \xi^0) - \frac{\partial^2}{\partial v \partial \xi} Q^0(v^0, \xi^0) (\hat{\xi} - \xi^0) \right] = o_p(1).$$

Consequently, $R_n^1(v) = \sqrt{n}(v - v^0) o_p(1)$. Using the relation that $x/(1+x^2) \leq 1/2$, we conclude that

$$\sup_{|v-v^0| \leq \delta_n} \frac{|R_n^1(v)|}{1+n|v-v^0|^2} \leq o_p(1) \sup_{|v-v^0| \leq \delta_n} \frac{\sqrt{n}|v-v^0|}{1+n|v-v^0|^2} = o_p(1).$$

Having verified Conditions 2.i and 2.ii, it remains to verify Condition 2.iii. The Hessian term J^0 is obviously positive definite because the limiting likelihood function is multiple-times smoothly differentiable, and because v^0 uniquely maximizes $Q^0(v^0, \xi^0)$. We note, next, that $\Delta_n(v^0)/n$ takes the form

$$\begin{aligned} & -\frac{1-\tau}{1-F_{X|V}[\hat{F}_n^{-1}(\tau)|v]} \frac{\partial}{\partial v} F_{X|V}[\hat{F}_n^{-1}(\tau)|v] + \frac{1-\tau}{1-F_{X|V}[F_{X|V}^{-1}(\tau)|v]} \frac{\partial}{\partial v} F_{X|V}[F_{X|V}^{-1}(\tau)|v] + \\ & (\hat{E} - \mathcal{E}) \left(\frac{\partial}{\partial v} \log f_{X|V}(X_i|v^0) \mathbf{1}[X_i \leq F_n^{-1}(\tau)] \right) + \hat{E} \left(\frac{\partial}{\partial v} \log f_{X|V}(X_i|v^0) \mathbf{1}[X_i \leq \hat{F}_n^{-1}(\tau)] \right) - \\ & \mathcal{E} \left(\frac{\partial}{\partial v} \log f_{X|V}(X_i|v^0) \mathbf{1}[X_i \leq F_{X|V}^{-1}(\tau|v^0)] \right) + o_p\left(\frac{1}{\sqrt{n}}\right) \end{aligned}$$

where \hat{E} denotes the empirical mean. Because we have represented the influence function of $(\hat{\xi} - \xi^0)$ as equation (8), we can compute that

$$\Delta_n(v^0) = \sum_{i=1}^n \psi(X_i) + o_p(\sqrt{n})$$

where

$$\begin{aligned} \psi(X_i) &= \frac{\partial}{\partial v} \log f_{X|V}(X_i|v^0) \mathbf{1}[X_i \leq F_{X|V}^{-1}(\tau|v)] - \mathcal{E}\left(\frac{\partial}{\partial v} \log f_{X|V}(X_i|v^0) \mathbf{1}[X_i \leq F_{X|V}^{-1}(\tau|v)]\right) - \\ &\quad \frac{1}{1-\tau} \left(\frac{\partial}{\partial v} F_{X|V}[F_{X|V}^{-1}(\tau)|v]\right) f_{X|V}[F_{X|V}^{-1}(\tau)|v] \left[\hat{F}_n^{-1}(\tau) - F_{X|V}^{-1}(\tau|v^0)\right] \\ &= \frac{\partial}{\partial v} \log f_{X|V}(X_i|v) \mathbf{1}[X_i \leq F_{X|V}^{-1}(\tau|v)] - \mathcal{E}\left(\frac{\partial}{\partial v} \log f_{X|V}(X_i|v) \mathbf{1}[X_i \leq F_{X|V}^{-1}(\tau|v)]\right) + \\ &\quad \frac{1}{1-\tau} \left(\frac{\partial}{\partial v} F_{X|V}[F_{X|V}^{-1}(\tau|v)|v]\right) \mathbf{1}[X_i \leq F_{X|V}^{-1}(\tau|v)] - \tau. \end{aligned}$$

Direct calculation of the asymptotic variance in the last line, while accounting for the covariance between the two terms, yields

$$\text{Var}[\psi(X_i)] = \mathcal{E}\left[\frac{\partial}{\partial v} \log f_{X|V}(X_i|v) \mathbf{1}[X_i \leq F_{X|V}^{-1}(\tau|v)]\right]^2 + \frac{1}{1-\tau} \left(\frac{\partial}{\partial v} F_{X|V}[F_{X|V}^{-1}(\tau|v)|v]\right)^2.$$

By inspection, we see that its inverse coincides with the asymptotic variance given in $\Sigma(\tau)$, which has been verified to equal J^0 in the information matrix equality calculation and, hence, is also positive definite. Its inverse yields the asymptotic variance of $\sqrt{n}(\hat{p} - v^0)$ and $\sqrt{n}(\hat{v} - v^0)$.

References

- Donald W.K. Andrews. Empirical process methods in econometrics. In Robert F. Engle and Daniel L. McFadden, editors, *Handbook of Econometrics, Volume 4*. New York: Elsevier, 1994.
- Victor Chernozhukov and Han Hong. A MCMC approach to classical estimation. *Journal of Econometrics*, 115:293–346, 2003.

- Han Hong and Matthew Shum. Rates of information aggregation in common value auctions. *Journal of Economic Theory*, 116:1–40, 2004.
- Matthew O. Jackson and Ilan Kremer. The relationship between the allocation of goods and a seller's revenue. *Journal of Mathematical Economics*, 40:371–392, 2004.
- Matthew O. Jackson and Ilan Kremer. The relevance of a choice of auction format in a competitive environment. *Review of Economic Studies*, 73:961–981, 2006.
- Ilan Kremer. Information aggregation in common value auctions. *Econometrica*, 70:1675–1682, 2002.
- Vijay Krishna. *Auction Theory*, Second Edition. San Diego: Academic Press, 2010.
- Claudio Mezzetti and Ilia Tsetlin. On the lowest-winning-bid and the highest-losing-bid auctions. *Journal of Mathematical Economics*, 44:1040–1048, 2008.
- Claudio Mezzetti and Ilia Tsetlin. On the lowest-winning-bid and the highest-losing-bid auctions. *Games and Economic Behavior*, 66:855–864, 2009.
- Paul R. Milgrom. A convergence theorem for competitive bidding with differential information. *Econometrica*, 47:670–688, 1979.
- Paul R. Milgrom. Rational expectations, information aggregation, and competitive bidding. *Econometrica*, 49:921–943, 1981.
- Paul R. Milgrom. *Putting Auction Theory to Work*. New York: Cambridge University Press, 2004.
- Paul R. Milgrom and Robert J. Weber. A theory of auctions and competitive bidding. *Econometrica*, 50:1089–1122, 1982.
- Paul R. Milgrom and Robert J. Weber. A theory of auctions and competitive bidding, II. In Paul Klemperer, editor, *The Theory of Auctions, II*. Cheltenham: Edward Elgar, 2000.

Whitney K. Newey and Daniel L. McFadden. Large sample estimation and hypothesis testing. In Robert F. Engle and Daniel L. McFadden, editors, *Handbook of Econometrics, Volume 4*. New York: Elsevier, 1994.

Wolfgang Pesendorfer and Jeroen M. Swinkels. The loser's curse and information aggregation in common value auctions. *Econometrica*, 65:1247–1281, 1997.

Wolfgang Pesendorfer and Jeroen M. Swinkels. Efficiency and information aggregation in auctions. *American Economic Review*, 90:499–525, 2000.

Robert J. Weber. Multi-object auctions. In Richard Engelbrecht-Wiggans Martin Shubik and Robert M. Stark, editors, *Auctions, Bidding and Contracting*. New York: New York University Press, 1983.

Robert B. Wilson. A bidding model of perfect competition. *Review of Economic Studies*, 44: 511–518, 1977.

Table 1: Announced Price, Number of Exits, and Total Exits

Price	20.00	22.00	23.00	24.00	25.00	26.00	27.00	28.00
Exits	0	2	0	0	0	0	0	0
Total Exits	0	2	2	2	2	2	2	2
Price	29.00	30.00	31.00	32.00	33.00	34.00	35.00	36.00
Exits	0	0	0	0	0	0	0	1
Total Exits	2	2	2	2	2	2	2	3
Price	37.00	38.00	39.00	40.00	41.00	42.00	43.00	44.00
Exits	0	0	0	0	1	0	2	1
Total Exits	3	3	3	3	4	4	6	7
Price	45.00	46.00	47.00	48.00	49.00	50.00	50.50	51.00
Exits	0	3	1	0	1	1	2	2
Total Exits	7	10	11	11	12	13	15	17
Price	51.20	51.40	51.50	51.60	51.70	51.80	51.90	52.00
Exits	0	0	0	0	0	0	0	0
Total Exits	17	17	17	17	17	17	17	17
Price	52.10	52.20	52.30	52.40	52.50	52.55	52.60	52.65
Exits	0	0	0	0	1	0	1	0
Total Exits	17	17	17	17	18	18	19	19
Price	52.70	52.75	52.80	52.85	52.90	52.95	53.00	53.05
Exits	0	0	0	0	0	0	0	0
Total Exits	19	19	19	19	19	19	19	19
Price	53.10	53.15	53.20	53.25	53.30	53.35	53.40	53.45
Exits	0	0	0	0	0	0	0	0
Total Exits	19	19	19	19	19	19	19	19
Price	53.50	53.55	53.60	53.65	53.70	53.75	53.80	53.85
Exits	0	0	0	0	0	0	0	0
Total Exits	19	19	19	19	19	19	19	19
Price	53.90	53.95	54.00	54.05	54.10	54.15	54.20	54.25
Exits	0	0	0	0	0	0	0	1
Total Exits	19	19	19	19	19	19	19	20

Table 2: Maximum-Likelihood Estimates of Normal Specification

Parameter	Estimate	Std.Error
v^0	56.31	0.97
σ	19.35	3.95

Table 3: Simulated Differences in Expected Revenue

	Number of Objects		
Number of Bidders	10	20	30
20	0.138		
30	0.301	0.079	
40	0.265	0.127	0.171