Papers to be covered:

- Laffont, Ossard, and Vuong (1995)
- Guerre, Perrigne, and Vuong (2000)
- Haile, Hong, and Shum (2003)

In this lecture, focus on empirical of auction models. Auctions are models of asymmetric information which have generated the most interest empirically. We begin by summarizing some relevant theory.

1 Theoretical background

An auction is a game of incomplete information. Assume that there are N players, or bidders, indexed by $i = 1, \ldots, N$. There are two fundamental random elements in any auction model.

- Bidders' private signals $X_1, \ldots, X_N$. We assume that the signals are scalar random variables, although there has been recent interest in models where each signal is multi-dimensional.
- Bidders' utilities: $u_i(X_i, X_{-i})$, where $X_{-i} \equiv \{X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_N\}$, the vector of signals excluding bidder i's signal. Since signals are private, $V_i \equiv u_i(X_i, X_{-i})$ is a random variable from *all* bidders' point of view. In what follows, we will also refer to bidder i's (random) utility as her *valuation*.

Differing assumptions on the form of bidders' utility function lead to the important distinction between **common value** and **private value** models. In the private value case, $V_i = X_i$, $\forall i$: each bidder knows his own valuation, but not that of his rivals.[1] In the (pure) common value case, $V_i = V$, $\forall i$, where V is in turn a random variable from all bidders' point of view, and bidders' signals are to be interpreted as their noisy estimates of the true but known common value V. Therefore, signals will generally not be independent when common values are involved. More generally a common value model arises when $u_i(X_i, X_{-i})$ is functionally dependent on X_{-i}.

[1] More generally, in a private value model, $u_i(X_i, X_{-i})$ is restricted to be a function *only* of X_i.

Before proceeding, we give some examples to illustrate the auction formats discussed above.

- Symmetric independent private values (IPV) model: $X_i \sim F$, *i.i.d.* across all bidders i, and $V_i = X_i$. Therefore, $F(X_1, \ldots, X_N) = F(X_1) * F(X_2) \cdots * F(X_N)$, and $F(V_1, X_1, \ldots, V_N, X_N) = \prod_i [F(X_i)]^2$.
- Conditional independent model: signals are independent, conditional on a common component V. $V_i = V, \forall i$, but $F(V, X_1, \ldots, X_N) = F(V) \prod_i F(X_i|V)$.

Models also differ depending on the auction rules. In a **first-price** auction, the object is awarded to the highest bidder, at her bid. A **second-price** auction also awards the object to the highest bidder, but she pays a price equal to the bid of the second-highest bidder. (Sometimes second-price auctions are also called "Vickrey" auctions, after the late Nobel laureate William Vickrey.) In an **English** or **ascending** auction, the price the raised continuously by the auctioneer, and the winner is the last bidder to remain, and he pays an amount equal to the price at which all of his rivals have dropped out of the auction. In a **Dutch** auction, the price is *lowered* continuously by the auctioneer, and the winner is the first bidder to agree to pay any price.

There is a large amount of theory and empirical work. In this lecture, we focus on first-price auction models. We also discuss a few theoretical concepts that will come up in the empirical papers discussed later.

1.1 Equilibrium bidding

In discussing equilibrium bidding in the different auction models, we will focus on the general symmetric affiliated model, used in the seminal paper of Milgrom and Weber (1982). The assumptions made in this model are:

- $V_i = u_i(X_i, X_{-i})$
- Symmetry: $F(V_1, X_1, \ldots, V_N, X_N)$ is symmetric (i.e., exchangeable) in the indices i so that, for example, $F(V_N, X_N, \ldots, V_1, X_1) = F(V_1, X_1, \ldots, V_N, X_N)$.
- The random variables $V_1, \ldots, V_N, X_1, \ldots, X_N$ are affiliated. Let $Z_1, \ldots, Z_M$ and $Z_1^*, \ldots, Z_M^*$ denote two realizations of a random vector process, and $\bar{Z}$ and $\underline{Z}$ denote, respectively, the component-wise maximum and minimum. Then we say that

$Z_1, \ldots, Z_M$ are affiliated if $F(\bar{Z})F(\underline{Z}) \geq F(Z_1, \ldots, Z_M)F(Z_1^*, \ldots, Z_M^*)$. In other words, large values for some of the variables make large values for the other variables most likely. Affiliation implies useful stochastic orderings on the conditional distributions of bidders' signals and valuations, which is necessary in deriving monotonic equilibrium bidding strategies.

Let $Y_i \equiv \max_{j \neq i} X_j$, the highest of the signals observed by bidder i's rivals. Given affiliation, the conditional expectation $E[V_i|X_i, Y_i]$ is increasing in both X_i and Y_i.

Winner's curse Another consequence of affiliation is the *winner's curse*, which is just the fact that

$$E[V_i|X_i] \geq E[V_i|X_i > Y_i]$$

where the conditioning event in the second expectation $(X_i > Y_i)$ is the event of winning the auction.

To see this, note that

$$\begin{aligned} E[V_i|X_i] = E_{X_{-i}} E\left[V_i|X_i; X_{-i}\right] &= \underbrace{\int \cdots \int}_{N-1} E\left[V_i|X_i; X_{-i}\right] F\left(dX_1, \ldots, dX_N\right) \\ &\geq \underbrace{\int^{X_i} \cdots \int^{X_i}}_{N-1} E\left[V_i|X_i; X_{-i}\right] F\left(dX_1, \ldots, dX_N\right) \\ &= E\left[V_i|X_i > X_j, j \neq i\right] = E\left[V_i|X_i > Y_i\right]. \end{aligned}$$

In other words, if bidder i "naively" bids $E\left[V_i|X_i\right]$, her expected payoff from a first-price auction is negative for every X_i. In equilibrium, therefore, rational bidders should "shade down" their bids by a factor to account for the winner's curse.

This winner's curse intuition arises in many non-auction settings also. For example, in two-sided markets where traders have private signals about unknown fundamental value of the asset, the ability to consummate a trade is "good news" for sellers, but "bad news" for buyers, implying that, without ex-ante gains from trade, traders may not be able to settle on a market-clearing price. The result is the famous "lemons" result by Akerlof (1970), as well as a version of the "no-trade" theorem in Milgrom and Stokey (1982). Glosten and Milgrom (1985) apply the same intuition to explain bid-ask spreads in financial markets.

Next, we cover some specific auction results in some detail, in order to understand methodology in the empirical papers.

1.2 First-price auctions

We derive the symmetric monotonic equilibrium bidding strategy $b^*(\cdot)$ for first-price auctions. If bidder i wins the auction, he pays his bid $b^*(X_i)$. His expected profit is

$$\begin{aligned}
&= E\left[(V_i - b)\mathbf{1}\left(b^*(Y_i) < b\right)|X_i = x\right] \\
&= E_{Y_i} E\left[(V_i - b)\mathbf{1}\left(Y_i < b^{*-1}(b)\right)|X_i = x, Y_i\right] \\
&= E_{Y_i}\left[(V(x, Y_i) - b)\mathbf{1}\left(Y_i < b^{*-1}(b)\right)|X_i = x\right] \\
&= \int_{-\infty}^{b^{*-1}(b)} (V(x, Y_i) - b) f(Y_i|x) dY_i.
\end{aligned}$$

The first-order conditions are

$$\begin{aligned}
0 &= -\int_{-\infty}^{b^{*-1}(b)} f(Y_i|x) dY_i + \frac{1}{b^{*\prime}(x)}\left[(V(x, x) - b) * f_{Y_i|X_i}(x|x)\right] \Leftrightarrow \\
0 &= -F_{Y_i|x}(x|x) + \frac{1}{b^{*\prime}(x)}\left[(V(x, x) - b) * f_{Y_i|X_i}(x|x)\right] \Leftrightarrow \\
b^{*\prime}(x) &= (V(x, x) - b^*(x))\left[\frac{f(x|x)}{F(x|x)}\right] \Rightarrow \\
b^*(x) &= \exp\left(-\int_{\underline{x}}^{x} \frac{f(s|s)}{F(s|s)} ds\right) b(\underline{x}) + \int_{\underline{x}}^{x} V(\alpha, \alpha) dL(\alpha|x)
\end{aligned}$$

where

$$L(\alpha|x) = \exp\left(-\int_{\alpha}^{x} \frac{f(s|s)}{F(s|s)}\right).$$

Initial condition: $b(\underline{x}) = V(\underline{x}, \underline{x})$.

For the IPV case:

$$\begin{aligned}
V(\alpha, \alpha) &= \alpha \\
F(s|s) &= F(s)^{N-1} \\
f(s|s) &= (n-1)F(s)^{N-2} f(s)
\end{aligned}$$

An example $X_i \sim \mathcal{U}[0,1]$, *i.i.d.* across bidders i. Then $F(s)=s$, $f(s)=1$. Then

$$\begin{aligned}
b^*(x) = 0 + &\int_0^x \alpha \exp\left(-\int_\alpha^x \frac{(n-1)f(s)}{F(s)}ds\right)\frac{(n-1)f(\alpha)}{F(\alpha)}d\alpha \\
= &\int_0^x \exp\left(-(n-1)(\log\frac{x}{\alpha})\right)(n-1)d\alpha \\
= &\int_0^x \left(\frac{\alpha}{x}\right)^{N-1}(N-1)d\alpha \\
= &\ \alpha\left(\frac{N-1}{N}\right)\left(\frac{\alpha}{x}\right)^N\Big]_0^x \\
= &\left(\frac{N-1}{N}\right)x.
\end{aligned}$$

1.2.1 Reserve prices

A reserve price just changes the initial condition of the equilibrium bid function. With reserve price r, initial condition is now $b(x^*(r)) = r$. Here $x^*(r)$ denotes the screening value, defined as

$$x^*(r) \equiv \inf\left\{x:\ E\left[V_i|X_i=x, Y_i<x\right] \geq r\right\}. \tag{1}$$

Conditional expectation in brackets is value of winning to bidder i, who has signal x. Screening value is lowest signal such that bidder i is willing to pay at least the reserve price r.

(Note: in PV case, $x^*(r) = r$. In CV case, with affiliation, generally $x^*(r) > r$, due to winners curse.)

Equilibrium bidding strategy is now:

$$b^*(x)\begin{cases} = \exp\left(-\int_{x^*(r)}^x \frac{f(s|s)}{F(s|s)}ds\right)r + \int_{x^*(r)}^x V(\alpha,\alpha)dL(\alpha|x) & \text{for } x \geq x^*(r) \\ < r & \text{for } x < x^*(r)\end{cases}$$

For IPV, uniform example above:

$$b^*(x) = \left(\frac{N-1}{N}\right)x + \frac{1}{N}\left(\frac{r}{x}\right)^{N-1} r.$$

1.3 Second-price auctions

Assume the existence of a monotonic equilibrium bidding strategy $b^*(x)$. Next we derive the functional form of this equilibrium strategy.

Given monotonicity, the price that bidder i will pay (if he wins) is $b^*(Y_i)$: the bid submitted by his closest rivals. He only wins when his bid $b < b^*(Y_i)$. Therefore, his expected profit from participating in the auction with a bid b and a signal $X_i = x$ is:

$$\begin{aligned}
&E_{Y_i}\left[(V_i - b^*(Y_i))\,\mathbf{1}\,(b^*(Y_i) < b)\,|X_i = x\right]\\
=&E_{Y_i}\left[(V_i - b^*(Y_i))\,\mathbf{1}\,(Y_i < X_i)\,|X_i = x\right]\\
=&E_{Y_i|X_i}E\left[(V_i - b^*(Y_i))\,\mathbf{1}\,(Y_i < X_i)\,|X_i = x, Y_i\right]\\
=&E_{Y_i|X_i}\left[(E(V_i|X_i, Y_i) - b^*(Y_i))\,\mathbf{1}\,(Y_i < X_i)\right]\\
\equiv&E_{Y_i|X_i}\left[(v(X_i, Y_i) - b^*(Y_i))\,\mathbf{1}\,(Y_i < X_i)\right]\\
=&\int_{-\infty}^{(b^*)^{-1}(b)}(v(x, Y_i) - b^*(Y_i))\,f\,(Y_i|X_i = x)\,.
\end{aligned} \tag{2}$$

Bidder i chooses his bid b to maximize his profits. The first-order conditions are (using Leibniz' rule):

$$\begin{aligned}
0 &= b^{*-1\prime}(b) * \left[v(x, b^{*-1}(b)) - b^*(b^{*-1}(b))\right] * f(b^{*-1}(b)|X_i) \Leftrightarrow\\
0 &= \frac{1}{b^{*\prime}(b)}\left[v(x, x) - b^*(x)\right] * f(b^{*-1}(b)|X_i) \Leftrightarrow\\
b^*(x) &= v(x, x) = E\left[V_i|X_i = x, Y_i = x\right].
\end{aligned}$$

In the PV case, the equilibrium bidding strategy simplifies to

$$b^*(x) = v(x, x) = x.$$

With reserve price, equilibrium strategy remains the same, except that bidders with signals less than the screening value $x^*(r)$ (defined in Eq. (1) above) do not bid.

2 Laffont-Ossard-Vuong (1995): "Econometrics of First-Price Auctions"

- Structural estimation of 1PA model, in IPV context.
- Example of a parametric approach to estimation.
- Goal of empirical work:
  - We observe bids $b_1, \ldots, b_n$, and we want to recover valuations $v_1, \ldots, v_n$.

- Why? Analogously to demand estimation, we can evaluate the "market power" of bidders, as measured by the margin $v - p$.

  Could be interesting to examine: how fast does margin decrease as n (number of bidders) increases?
- Useful for the optimal design of auctions:
  1. What is auction format which would maximize seller revenue?
  2. What value for reserve price would maximize seller revenue?

- Another exercise in simulation estimation

■■■

MODEL

- I bidders

- Information structure is IPV: valuations $v^i,\ i = 1, \ldots, I$ are *i.i.d.* from $F(\cdot|z_l, \theta)$ where l indexes auctions, and z_l are characteristics of l-th auctions
- θ is parameter vector of interest, and goal of estimation
- p^0 denotes "reserve price": bid is rejected if $< p^0$.
- Dutch auction: strategically identical to first-price sealed bid auction.

Equilibrium bidding strategy is:

$$b^i = e\left(v^i, I, p^0, F\right) = \begin{cases} v^i - \frac{\int_{p^0}^{v^i} F(x)^{I-1} dx}{F(v^i)^{I-1}} & \text{if } v^i > p^0 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

Note: (1) $b^i(v^i = p^0) = p^0$; (2) strictly increasing in v^i.

■■■

Dataset: only observe *winning bid* b_l^w for each auction l. Because bidders with lower bids never have a chance to bid in Dutch auction.

Given monotonicity, the winning bid $b^w = e\left(v_{(I)}, I, p^0, F\right)$, where $v_{(I)} \equiv \max_i v^i$ (the highest order statistic out of the I valuations).

Furthermore, the CDF of $v_{(I)}$ is $F(\cdot|z_l,\theta)^I$, with corresponding density $I \cdot F^{I-1} f$.

■■■

Goal is to estimate θ by (roughly speaking) matching the winning bid in each auction l to its expectation.

Expected winning bid is (for simplicity, drop z_l and θ now)

$$\begin{aligned}
E_{v_{(I)}>p^0}(b^w) &= \int_{p^0}^{\infty} e\left(v_{(I)}, I, p^0, F\right) I \cdot F(v|\theta)^{I-1} f(v|\theta) dv \\
&= I \int_{p^0}^{\infty} \left(v - \frac{\int_{p^0}^{v} F(x)^{I-1} dx}{F(v)^{I-1}} \right) F(v|\theta)^{I-1} f(v|\theta) dv \\
&= I \int_{p^0}^{\infty} \left(v \cdot F(v)^{I-1} - \int_{p^0}^{\infty} F(x)^{I-1} dx \right) f(v) dv. \quad (*).
\end{aligned}$$

■■■

If we were to estimate by simulated nonlinear least squares, we would proceed by finding θ to minimize the sum-of-squares between the observed winning bids and the predicted winning bid, given by expression (*) above. Since (*) involves complicated integrals, we would simulate (*), for each parameter vector θ.

How would this be done:

- Draw valuations v^s, $s = 1, \ldots, S$ i.i.d. according to $f(v|\theta)$. This can be done by drawint $u_1, \ldots, u_S$ i.i.d. from the $U[0,1]$ distribution, then transform each draw:

$$v_s = F^{-1}(u_s|\theta).$$

- For each simulated valuation v_s, compute integrand $\mathcal{V}_s = v_s F(v_s|\theta)^{I-1} - \int_{p^0}^{v_s} F(x|\theta)^{I-1} dx$. (Second term can also be simulated, but one-dimensional integral is that very hard to compute.)
- Approximate the expected winning bid as $\frac{1}{S} \sum_s \mathcal{V}_s$.

However, the authors do not do this— they propose a more elegant solution. In particular, they simplify the simulation procedure for the expected winning bid by appealing to the **Revenue-Equivalence Theorem**: an important result for auctions where bidders' signals are independent, and the model is symmetric. (This was first derived explicitly in Myerson (1981), and this statement is due to Klemperer (1999).)

Theorem 1 *(Revenue Equivalence) Assume each of N risk-neutral bidders has a privately-known signal X independently drawn from a common distribution F that is strictly increasing and atomless on its support $[\underline{X}, \bar{X}]$. Any auction mechanism which is (i)* efficient *in awarding the object to the bidder with the highest signal with probability one; and (ii) leaves any bidder with the lowest signal $\underline{X}$ with zero surplus yields the same expected revenue for the seller, and results in a bidder with signal x making the same expected payment.*

From a mechanism design point of view, auctions are complicated because they are multiple-agent problems, in which a given agent's payoff can depend on the reports of all the agents. However, in the independent signal case, there is no gain (in terms of stronger incentives) in making any given agent's payoff depend on her rivals' reports, so that a symmetric auction with independent signal essentially boils down to independent contracts offered to each of the agents individually.

Furthermore, in any efficient auction, the probability that a given agent with a signal x wins is the same (and, in fact, equals $F(x)^{N-1}$). This implies that each bidder's expected surplus function (as a function of his signal) is the same, and therefore that the expected payment schedule is the same.

■■■

By RET:

- expected revenue in 1PA same as expected revenue in 2PA
- expected revenue in 2PA is $Ev^{(I-1)}$
- with reserve price, expected revenue in 2PA is $E\max(v^{(I-1)}, p^0)$. (Note: with IPV structure, reserve price r screens out same subset of valuations $v \leq r$ in both 1PA and 2PA.)

■■■

Hence, we have that

$$Eb^*(v_{(I)}) = E\left[\max\left(v_{(I-1)}, p^0\right)\right]$$

which is insanely easy to simulate:

For each parameter vector θ, and each auction l

- For each simulation draw $s = 1, \ldots, S$:
  - Draw $v_1^s, \ldots, v_{I_l}^s$: vector of simulated valuations for auction l (which had I_l participants)
  - Sort the draws in ascending order: $v_{1:I_l}^s < \cdots < v_{I_l:I_l}^s$
  - Set $b_l^{w,s} = v_{I-1_l:I_l}$ (ie. the second-highest valuation)
  - If $b_l^{w,s} < p_l^0$, set $b_l^{w,s} = p_l^0$. (ie. $b_l^{w,s} = \max\left(v_{I-1_l:I_l}^s, p_l^0\right)$)
- Approximate $E\left(b_l^w; \theta\right) = \frac{1}{S}\sum_s b_l^{w,s}$.

Estimate θ by simulated nonlinear least squares:

$$\min_\theta \frac{1}{L}\sum_{l=1}^{L}\left(b_l^w - E\left(b_l^w; \theta\right)\right)^2.$$

Results.

■■■

Remarks:

- Problem: bias when number of simulation draws S is fixed (as number of auctions $L \to \infty$). Propose bias correction estimator, which is consistent and asymptotic normal under these conditions.
- This clever methodology is useful for independent value models: works for all cases where revenue equivalence theorem holds.
- Does not work for affiliated value models (including common value models)

■■■

3 Application: internet used car auctions

- Consider Lewis (2007) paper on used cars sold on eBay (simplified exposition)
- Question: does information revealed by sellers lead to high prices? (Question about the credibiltiy of information revealed by sellers.)

- Observe transsactions price in ascending auction. Assume that transaction price is equal to
$$v(X_{n-1:n}, X_{n-1:n})$$
(as in second-price auction).

- Consider pure common value setup with conditionally independent signals. Log-normality is assumed:
$$\tilde{v} \equiv \log v = \mu + \sigma\epsilon_v \;\sim\; N(\mu, \sigma^2)$$
$$x_i|\tilde{v} = \tilde{v} + r\epsilon_i \;\sim\; N(\tilde{v}, r^2)$$
These are, respectively, the prior distribution of valuations, and the conditional distribution of signals.

- Allow seller information variables z to affect the mean and variance of the prior distribution:
$$\mu = \alpha' z$$
$$\sigma = \kappa(\beta' z).$$
$\kappa(\cdot)$ is just a transformation of the index $\beta' z$ to ensure that the estimate of $\sigma > 0$.

- z includes variables such as: number of photos, how much text is on the website. (Larger z denotes better information.)

- Question: is $\alpha > 0$?

- Results.

4 Guerre-Perrigne-Vuong (2000): Nonparametric Identification and Estimation in IPV First-price Auction Model

The recent emphasis in the empirical literature is on *nonparametric* identification and estimation of auction models. Motivation is to estimate bidders' unobserved valuations, while avoiding parametric assumption (as in the LOV paper).

- Recall first-order condition for equilibrium bid (general affiliated values case):
$$b'(x) = (v(x,x) - b(x)) \cdot \frac{f_{y_i|x_i}(x|x)}{F_{y_i|x_i}(x|x)}; \quad y_i \equiv \max_{j \neq i} x_i. \tag{4}$$

- In IPV case:

$$\begin{aligned} v(x,x) &= x \\ F_{y_i|x_i}(x|x) &= F(x)^{n-1} \\ f_{y_i|x_i}(x|x) &= \frac{\partial}{\partial x} F(x)^{n-1} = (n-1)F(x)^{n-2} f(x) \end{aligned}$$

so that first-order condition becomes

$$\begin{aligned} b'(x) &= (x - b(x)) \cdot (n-1) \frac{F(x)^{n-2} f(x)}{F(x)^{n-1}} \\ &= (x - b(x)) \cdot (n-1) \frac{f(x)}{F(x)}. \end{aligned} \tag{5}$$

- Now, note that because equilibrium bidding function $b(x)$ is just a monotone increasing function of the valuation x, the change of variables formulas yield that (take $b_i \equiv b(x_i)$)

  –
  $$G(b_i) = F(x_i)$$

  –
  $$g(b_i) = f(x_i) \cdot 1/b'(x_i)$$

  .

Hence, substituting the above into Eq. (5):

$$\begin{aligned} \frac{1}{g(b_i)} &= (n-1) \frac{x_i - b_i}{G(b_i)} \\ \Leftrightarrow x_i &= b_i + \frac{G(b_i)}{(n-1)g(b_i)}. \end{aligned} \tag{6}$$

Everything on the RHS of the preceding equation is observed: the equilibrium bid CDF G and density g can be estimated directly from the data *nonparametrically.* Assuming a dataset consisting of T n-bidder auctions:

$$\begin{aligned} \hat{g}(b) &\approx \frac{1}{T \cdot n} \sum_{t=1}^{T} \sum_{i=1}^{n} \frac{1}{h} \mathcal{K}\left(\frac{b - b_{it}}{h}\right) \\ \hat{G}(b) &\approx \frac{1}{T \cdot n} \sum_{t=1}^{T} \sum_{i=1}^{n} \mathbf{1}(b_{it} \leq b). \end{aligned} \tag{7}$$

The first is a *kernel density estimate* of bid density. The second is the *empirical distribution function (EDF).*

- In the above, $\mathcal{K}$ is a "kernel function". A kernel function is a function satisfying the following conditions:

  1. It is a probability density function, ie: $\int_{-\infty}^{+\infty} \mathcal{K}(d)du = 1$, and $\mathcal{K}(u) \geq 0$ for all u.
  2. It is symmetric around zero: $\mathcal{K}(u) = \mathcal{K}(-u)$.
  3. h is bandwidth: describe below
  4. Examples:
     (a) $\mathcal{K}(u) = \phi(u)$ (standard normal density function);
     (b) $\mathcal{K}(u) = \frac{1}{2}\mathbf{1}(|u| \leq 1)$ (uniform kernel);
     (c) $\mathcal{K}(u) = \frac{3}{4}(1-u^2)\mathbf{1}(|u| \leq 1)$ (Epanechnikov kernel)

- To get some intuition for the kernel estimate of $\hat{g}(b)$, consider the histogram

$$h(b) = \frac{1}{Tn} \sum_t \sum_i \mathbf{1}(b_{it} \in [b - \epsilon, b + \epsilon])$$

  for some small $\epsilon > 0$. The histogram at b, $h(b)$ is the frequency with which the observed bids land within an ϵ-neighborhood of b.

- In comparison, the kernel estimate of $\hat{g}(b)$ replaces $\mathbf{1}(b_{it} \in [b-\epsilon, b+\epsilon])$ with $\frac{1}{h}\mathcal{K}\left(\frac{b-b_{it}}{h}\right)$. This is:

  – always ≥ 0
  – takes large values for b_{it} close to b; small values (or zero) for b_{it} far from b
  – takes values in $\mathbb{R}+$ (can be much larger than 1)
  – h is bandwidth, which blows up $\frac{1}{h}\mathcal{K}\left(\frac{b-b_{it}}{h}\right)$: when it is smaller, then this quantity becomes larger.

    Think of h as measuring the "neighborhood size" (like ϵ in the histogram). When $T \to \infty$, then we can make h smaller and smaller.

    Bias/variance tradeoff.
  – Roughly speaking, then, $\hat{g}(b)$ is a "smoothed" histogram,

- For $\hat{G}(b)$, recall definition of the CDF:

$$G(\tilde{b}) = Pr(b \leq \tilde{b}).$$

  The EDF measures these probabilities by the (within-sample) frequency of the events.

- Hence, the IPV first-price auction model is *nonparametrically identified.* For each observed bid b_i, the corresponding valuation $x_i = b^{-1}(b_i)$ can be recovered as:

$$\hat{x}_i = b_i + \frac{\hat{G}(b_i)}{(n-1)\hat{g}(b_i)}. \tag{8}$$

Hence, GPV recommend a two-step approach to estimating the valuation distribution $f(x)$:

1. In first step, estimate $G(b)$ and $g(b)$ nonparametrically, using Eqs. (7).
2. In second step, estimate $f(x)$ by using kernel density estimator of recovered valuations:

$$\hat{f}(x) \approx \frac{1}{T \cdot n} \sum_{t=1}^{T} \sum_{i=1}^{n} \frac{1}{h} \mathcal{K}\left(\frac{x - \hat{x}_{it}}{h}\right). \tag{9}$$

Athey and Haile (2002) shows many nonparametric identification results for a variety of auction models (first-price, second-price) under a variety of assumption on the information structure (symmetry, asymmetry). They focus on situations when only a subset of the bids submitted in an auction are available to a researcher.

As an example of such a result, we see that identification continues to hold, even when only the highest-bid in each auction is observed. Specifically, if only $b_{n:n}$ is observed, we can estimate $G_{n:n}$, the CDF of the maximum bid, from the data. Note that the relationship between the CDF of the maximum bid and the marginal CDF of an equilibrium bid is

$$G_{n:n}(b) = G(b)^n$$

implying that $G(b)$ can be recovered from knowledge of $G_{n:n}(b)$. Once $G(b)$ is recovered, the corresponding density $g(b)$ can also be recovered, and we could solve Eq. (8) for every b to obtain the inverse bid function.

■■■

5 Affiliated values models

Can this methodology be extended to affiliated values models (including common value models)?

However, Laffont and Vuong (1996) nonidentification result: from observation of bids in n-bidder auctions, the affiliated private value model (ie. a PV model where valuations are dependent across bidders) is indistinguishable from a CV model.

- Intuitively, all you identify from observed bid data is joint density of $b_1, \ldots, b_n$. In particular, can recover the correlation structure amongst the bids. But correlation of bids in an auction could be due to both affiliated PV, or to CV.

5.1 Affiliated private value models

Li, Perrigne, and Vuong (2002) proceed to consider nonparametric identification and estimation of the affiliated private values model. In this model, valuations $x_i, \ldots, x_n$ are drawn from some joint distribution (and there can be arbitrary correlation amongst them).

First order condition for this model is: for bidder i

$$b'(x_i) = (x_i - b(x)) \cdot \frac{f_{y_i|x_i}(x|x)}{F_{y_i|x_i}(x|x)}; \quad y_i \equiv \max_{j \neq i} x_i.$$

where $y_i \equiv \max_{j \neq i} \{x_1, \ldots, x_n\}$.

Procedure similar to GPV can be used here to recover, for each bid b_i, the corresponding valuation $x_i = b^{-1}(b_i)$. As before, exploit the following change of variable formulas:

- $$G_{b^*|b}(b|b) = F_{y|x}(x|x)$$
- $$g_{b^*|b}(b|b) = f_{y|x}(x|x) \cdot 1/b'(x)$$

where b^* denotes (for a given bidder), the highest bid submitted by this bidder's rivals: for a given bidder i, $b_i^* = \max_{j \neq i} b_j$. To prepare what follows, we introduce n subscript (so we index distributions according to the number of bidders in the auction).

Li, Perrigne, and Vuong (2000) suggest nonparametric estimates of the form

$$\begin{aligned} \hat{G}_n(b; b) &= \frac{1}{T_n \times h \times n} \sum_{t=1}^{T} \sum_{i=1}^{n} K\left(\frac{b - b_{it}}{h}\right) \mathbf{1}\left(b_{it}^* < b, n_t = n\right) \\ \hat{g}_n(b; b) &= \frac{1}{T_n \times h^2 \times n} \sum_{t=1}^{T} \sum_{i=1}^{n} \mathbf{1}(n_t = n) K\left(\frac{b - b_{it}}{h}\right) K\left(\frac{b - b_{it}^*}{h}\right). \end{aligned} \tag{10}$$

Here h and h are bandwidths and $K(\cdot)$ is a kernel. $\hat{G}_n(b; b)$ and $\hat{g}_n(b; b)$ are nonparametric estimates of

$$G_n(b; b) \equiv G_n(b|b) g_n(b) = \frac{\partial}{\partial b} \Pr(B_{it}^* \leq m, B_{it} \leq b)|_{m=b}$$

and

$$g_n(b;b) \equiv g_n(b|b)g_n(b) = \frac{\partial^2}{\partial m \partial b} \Pr(B_{it}^* \le m, B_{it} \le b)|_{m=b}$$

respectively, where $g_n(\cdot)$ is the marginal density of bids in equilibrium. Because

$$\frac{G_n(b;b)}{g_n(b;b)} = \frac{G_n(b|b)}{g_n(b|b)} \tag{11}$$

$\frac{\hat{G}_n(b;b)}{\hat{g}_n(b;b)}$ is a consistent estimator of $\frac{G_n(b|b)}{g_n(b|b)}$. Hence, by evaluating $\hat{G}_n(\cdot,\cdot)$ and $\hat{g}_n(\cdot,\cdot)$ at each observed bid, we can construct a pseudo-sample of consistent estimates of the realizations of each $x_{it} = b^{-1}(b_{it})$ using Eq. (4):

$$\hat{x}_{it} = \frac{\hat{G}_n(b_{it};b_{it})}{\hat{g}_n(b_{it};b_{it})}. \tag{12}$$

Subsequently, joint distribution of $x_1, \ldots, x_n$ can be recovered as sample joint distribution of $\hat{x}_1, \ldots, \hat{x}_n$.

5.2 Common value models: testing between CV and PV

Laffont-Vuong did not consider variation in n, the number of bidders.

In Haile, Hong, and Shum (2003), we explore how variation in n allows us to test for existence of CV.

Introduce notation:

$$v(x_i, x_i, n) = E[V_i | X_i = x_i, \max_{j \neq i} X_j = x_i, n].$$

Recall the winner's curse: it implies that $v(x, x, n)$ is invariant to n for all x in a PV model but strictly decreasing in n for all x in a CV model.

Consider the first-order condition in the common value case:

$$b'(x,n) = (v(x,x,n) - b(x,n)) \cdot \frac{f_{y_i|x_i,n}(x|x)}{F_{y_i|x_i,n}(x|x)}; \quad y_i \equiv \max_{j \neq i} x_i.$$

Hence, the Li, Perrigne, and Vuong (2002) procedure from the previous section can be used to recover the "pseudovalue" $v(x_i, x_i, n)$ corresponding to each observed bid b_i. Note that we cannot recover $x_i = b^{-1}(b_i)$ itself from the first-order condition, but can recover $v(x_i, x_i, n)$. (This insight was also articulated in Hendricks, Pinkse, and Porter (2003).)

In Haile, Hong, and Shum (2003), we use this intuition to develop a test for CV:

$$\begin{aligned} H_0 \text{ (PV)} : \ & E\left[v(X,X;\underline{n})\right] = E\left[v(X,X;\underline{n}+1)\right] = \cdots = E\left[v(X,X;\bar{n})\right] \\ H_1 \text{ (CV)} : \ & E\left[v(X,X;\underline{n})\right] > E\left[v(X,X;\underline{n}+1)\right] > \cdots > E\left[v(X,X;\bar{n})\right] \end{aligned}$$

Problem: bias at boundaries in kernel estimation of pseudo-values. The bid density $g(b,b)$ is estimated inaccurately for bids close to the boundary of the empirical support of bids.

Solution: use *quantile-trimmed means*: $\mu_{n,\tau} = E[v(X,X;n)\mathbf{1}\{x_\tau < X < x_{1-\tau}\}]$

above $\Rightarrow$

$$\begin{aligned} H_0 \text{ (PV)} \quad &: \quad \mu_{\underline{n},\tau} = \mu_{\underline{n}+1,\tau} = \cdots = \mu_{\overline{n},\tau} \\ H_1 \text{ (CV)} \quad &: \quad \mu_{\underline{n},\tau} > \mu_{\underline{n}+1,\tau} > \cdots > \mu_{\overline{n},\tau}. \end{aligned}$$

Theorem 3 Let $\hat{\mu}_{n,\tau} = \frac{1}{n \times T_n} \sum_{t=1}^{T_n} \sum_{i=1}^{n} \hat{v}_{it} \, \mathbf{1}\{b_{\tau,n} \leq b_{it} \leq b_{1-\tau,n}\}$ and assume [...conditions for kernel estimation...]. Then
(i) $\hat{\mu}_{n,\tau} \xrightarrow{p} E[v(X,X,n) \ \mathbf{1}\{x_\tau < X < x_{1-\tau}\}]$;
(ii)$\sqrt{T_n h}\,(\hat{\mu}_{n,\tau} - \mu_{n,\tau}) \xrightarrow{d} N(0,\omega_n)$, where
$\omega_n = \left[\int \left(\int K(v)K(u+v)dv\right)^2 du\right] \left[\frac{1}{n} \int_{F_b^{-1}(\tau)}^{F_b^{-1}(1-\tau)} \frac{G_n(b;b)^2}{g_n(b;b)^3} g_n(b)^2 db\right].$

Test statistic Now use standard multivariate one-sided LR test (Bartholomew, 1959) for normally distributed parameters $\hat{\mu}_{n,\tau}$

- $a_n = \frac{T_n h}{\omega_n}$ (inverse variance weights)
- $\overline{\mu} = \frac{\sum_{n=\underline{n}}^{\bar{n}} a_n \, \hat{\mu}_{n,\tau}}{\sum_{n=\underline{n}}^{\bar{n}} a_n}$ (MLE under null)
- $\mu^*_{\underline{n}}, \ldots, \mu^*_{\bar{n}}$ solves

$$\min_{\mu_{\underline{n}},\ldots,\mu_{\bar{n}}} \sum_{n=\underline{n}}^{\bar{n}} a_n \left(\hat{\mu}_{n,\tau} - \mu_n\right)^2 \quad s.t. \quad \mu_{\underline{n}} \geq \mu_{\underline{n}+1} \geq \cdots \geq \mu_{\bar{n}}. \tag{13}$$

- $\bar{\chi}^2 = \sum_{n=\underline{n}}^{\bar{n}} a_n \left(\mu^*_{n,\tau} - \bar{\mu}\right)^2$
  - distributed as mixture of χ^2_k rv's, $k = 0, 1, \ldots, \overline{n} - \underline{n}$
  - mixing weights: $\Pr_{H_0}$ {soln to (13) has exactly k slack constraints}
    (obtain by simulation)
- estimate ω_n using asymptotic formula or with bootstrap

5.3 Endogenous participation

The validity of this test relies crucially on the assumption that variation in n, the number of bidders, across auction is exogenous. Next, we consider how this can be relaxed.

Idea: bidder participation determined by unobservable (to us) factors, denoted W, which are also correlated with bidder valuations.

Problems:

1. valuations varying with N ($\Longrightarrow$ *second-stage test may be invalid*). Extreme case: if N is decreasing in W, then $\mu_{N=2} > \mu_{N=3}$, even under PV. "Usual" problem that endogeneity can confound results.
2. to estimate pseudo-values using the FOC, we must condition on all info (both N and W, e.g.) bidders do ($\Longrightarrow$ *first stage estimation invalid too*!). We must estimate equilibrium bid distributions g and G conditional on both N and W.

IV approach: assume there is an instrument Z which satisfies

Assumption 1 $N = \phi(Z, W)$, with ϕ nonconstant in Z and strictly increasing in W. (Implies W uniquely determined given N and Z, and discrete.)

This assumption is strong, but we will see why we need it.

Assumption 2 Z is independent of $(U_1, \ldots, U_n, X_1, \ldots, X_n, W)$.

Assumption 3 The support of $N|Z$ consists of a set of contiguous integers.

With these assumptions, it turns out there is no loss in generality from taking $\phi(\cdots)$ to be additive, and equal to:

$$\phi(Z_t, W_t) = \texttt{int}[E(N|Z_t)] + W_t.$$

Hence, the unobserved factor in auction t, is essentially "observed" after we run a first-stage nonparametric regression of N_t on Z_t:

$$\hat{W}_t = N_i - \texttt{int}[\widehat{E(N|Z_t)}].$$

This suggests that we can adapt the test in the following way:

1. Estimate bid distributions $G(b, b|n, w)$ and $g(b, b|n, w)$ conditional on both n and w.

2. For bid b_{it} in auction t, we can recover the corresponding pseudovalue as:

$$\hat{v}(x_i, x_i|n_t, w_t) = b_{it} + \frac{\hat{G}(b_{it}, b_{it}|n_t, w_t)}{\hat{g}(b_{it}, b_{it}|n_t, w_t)}.$$

3. Now the winner's curse implies that under PV, the conditional expectation $E_x v(x, x|n, w)$ conditional on (n, w) is invariant in n, for all w. However, under CV, it is decreasing in n, for all w.

References

Akerlof, G. (1970): "The Market for "Lemons": Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, 84, 488–500.

Athey, S., and P. Haile (2002): "Identification of Standard Auction Models," forthcoming in *Econometrica*.

Glosten, L., and P. Milgrom (1985): "Bid, ask, and transactions prices in a specialist market with heterogeneiously informed traders," *Journal of Financial Economics*, 14.

Guerre, E., I. Perrigne, and Q. Vuong (2000): "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 68, 525–74.

Haile, P., H. Hong, and M. Shum (2003): "Nonparametric Tests for Common Values in First-Price Auctions," NBER working paper #10105.

Hendricks, K., J. Pinkse, and R. Porter (2003): "Empirical Implications of Equilibrium Bidding in First-Price, Symmetric, Common-Value Auctions," *Review of Economic Studies*, 70, 115–145.

Klemperer, P. (1999): "Auction Theory: A Guide to the Literature," *Journal of Economic Surveys*, 13, 227–286.

Laffont, J. J., H. Ossard, and Q. Vuong (1995): "Econometrics of First-Price Auctions," *Econometrica*, 63, 953–980.

Laffont, J. J., and Q. Vuong (1996): "Structural Analysis of Auction Data," *American Economic Review, Papers and Proceedings*, 86, 414–420.

Lewis, G. (2007): "Asymmetric Information, Adverse Selection and Seller Revelation on eBay Motors," Harvard University, mimeo.

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

——— (2002): "Structural Estimation of the Affiliated Private Value Acution Model," *RAND Journal of Economics*, 33, 171–193.

Milgrom, P., and N. Stokey (1982): "Information, Trade and Common Knowledge," *Journal of Economic Theory*, 26, 17–27.

Milgrom, P., and R. Weber (1982): "A Theory of Auctions and Competitive Bidding," *Econometrica*, 50, 1089–1122.

Myerson, R. (1981): "Optimal Auction Design," *Mathematics of Operation Research*, 6, 58–73.