Unobserved Heterogeneity in Auctions

Empirical Industrial Organization

Andras Niedermayer[1]

[1]Department of Economics (THEMA), Cergy Paris Université

Winter 2024

Auctions without Heterogeneity

- Take the Guerre, Perrigne, and Vuong setting for auction econometrics with 2 bidders
- each bidder has valuation $V_i \sim F_i$
- bids are

$$A_1 = \beta_1(V_1) \tag{1}$$
$$A_2 = \beta_2(V_2) \tag{2}$$

Auctions with Observed Heterogeneity

- now assume that for each auction, there is a vector of covariates Z (e.g. for the sale of houses: the number of bedrooms, number of square meters, ZIP code)
- each bidders valuation is $Z'\gamma + V_i$ where V_i is the idiosyncratic valuation
- it can be shown that the bids B_i are additive in $Z'\beta$

$$B_1 = \beta_1(Z'\gamma + V_1) = Z'\gamma + \beta_1(V_1) = Z'\gamma + A_1 \qquad (3)$$

$$B_2 = \beta_2(Z'\gamma + V_2) = Z'\gamma + \beta_2(V_2) = Z'\gamma + A_2 \qquad (4)$$

- A_i is the "idiosyncratic bid"
- problem: B_1 and B_2 are correlated because of $Z'\gamma$, this violates Guerre, Perrigne, and Vuong's assumptions
- solution: run regression to estimate γ, use residuals $B_i - Z'\gamma$ as estimates of A_i

Auctions with Unobserved Heterogeneity

- assume that there is additionally unobserved heterogeneity Y, observed by market participants but not by the econometrician (e.g. how beautiful the view is from a house)
- Y, A_1, A_2 are independent
- again, Y is additive, observed bids B_i are

$$B_1 = Z'\gamma + Y + A_1 \tag{5}$$
$$B_2 = Z'\gamma + Y + A_2 \tag{6}$$

- problem: even after correcting for observables $B_1 - Z'\gamma$ and $B_2 - Z'\gamma$ are correlated because of Y
- solution: deconvolution method described by Krasnokutskaya (Identification and Estimation of Auction Models with Unobserved Heterogeneity, REStud, 2011)

Parametric Intuition

- before getting into the non-parametric estimation technique, we give an intuition based on parametric estimation
- assume we only want to know the variance of Y, A_1, and A_2 (e.g. because we assume that they are normally distributed, so that together with the mean, we would know everything)
- let us ignore $Z'\gamma$ (or assume we removed $Z'\gamma$ by taking residuals)

$$B_1 = Y + A_1 \tag{7}$$

$$B_2 = Y + A_2 \tag{8}$$

- then, for some parameters t_1, t_2

$$\mathrm{Var}[t_1 B_1 + t_2 B_2] = (t_1 + t_2)^2 \mathrm{Var}[Y] + t_1^2 \mathrm{Var}[A_1] + t_2^{\mathrm{Var}}[A_2]$$

Parametric Intuition

- with some algebra

$$\mathrm{Var}[t_1 B_1 + t_2 B_2] = (t_1 + t_2)^2 \mathrm{Var}[Y] + t_1^2 \mathrm{Var}[A_1] + t_2^2 \mathrm{Var}[A_2] \quad (9)$$

  can be transformed to

$$\frac{\partial}{\partial t_1} \mathrm{Var}[t_1 B_1 + t_2 B_2] \bigg|_{t_1=0} = t_2 \mathrm{Var}[Y]$$

- this gives us $\mathrm{Var}[Y]$ because the joint distribution of B_1 and B_2 is observable
- plugging $\mathrm{Var}[Y]$ into (9) and setting $t_1 = 0$, $t_2 = 1$ gives us $\mathrm{Var}[A_2]$
- setting $t_2 = 0$, $t_1 = 1$ gives us $\mathrm{Var}[A_1]$

Parametric Intuition

$$B_1 = Y + A_1$$
$$B_2 = Y + A_2$$

- to get the mean of Y, A_1, A_2, take expectations

$$E[B_1] = E[Y] + E[A_1]$$
$$E[B_2] = E[Y] + E[A_2]$$

- $\Rightarrow$ 2 equations, 3 unknowns
- $\Rightarrow$ we know the means up to a normalization (increasing $E[Y]$ by a constant and decreasing $E[A_1]$ and $E[A_2]$ by a constant is observationally equivalent)

- define the characteristic function

$$\Phi_Y(t) = E[\exp(itY)]$$

- assume $Y \sim G_Y$
- there is a one-to-one-mapping between the pdf f_Y and the characteristic function Φ_Y:

$$\Phi_Y(t) = \int_{\mathbb{R}} \exp(ity) g_Y(y) dy$$
$$g_Y(y) = \frac{1}{2\phi} \int_{\mathbb{R}} \exp(itx) \Phi_Y(t) dt$$

- define the characteristic functions $\Phi_1(t) = E[\exp(itA_1)]$ for $A_1 \sim G_1$ and $\Phi_2(t) = E[\exp(itA_1)]$ for $A_2 \sim G_2$

Non-Parametric Identification

- we observe the joint distribution of B_1 and B_2
- alternatively, we observe the characteristic function of the joint distribution of B_1 and B_2

$$\begin{aligned} \Psi(t_1, t_2) &= E[\exp(it_1 B_1 + it_2 B_2)] \\ &= \Phi_1(t_1)\Phi_2(t_2)\Phi_Y(t_1 + t_2) \qquad (10) \end{aligned}$$

- we can solve for Φ_Y:

$$\Phi_Y(t) = \exp\left(\int_0^t \left.\frac{\partial \ln \Phi(t_1, x)}{\partial t_1}\right|_{t_1=0} dx - iE[A_1]\right)$$

- plugging Φ_Y into (10) we get Φ_1 and Φ_2

Non-Parametric Estimation

- we have n observations $(B_{1j}, B_{2j})_{j=1}^{n}$
- we get the estimate $\hat{\Psi}$ using

$$\Psi(t_1, t_2) = E[\exp(it_1 B_1 + it_2 B_2)]$$
$$\Rightarrow \hat{\Psi}(t_1, t_2) = \frac{1}{n} \sum_{j=1}^{n} \exp(it_1 B_{1j} + it_2 B_{2j}) \tag{11}$$

- we get the estimate $\partial\hat{\Psi}/\partial t_1$ using

$$\frac{\partial \Psi(t_1, t_2)}{\partial t_1} = E[iB_1 \exp(it_1 B_1 + it_2 B_2)]$$
$$\Rightarrow \frac{\partial \hat{\Psi}(t_1, t_2)}{\partial t_1} = \frac{1}{n} \sum_{j=1}^{n} iB_{1j} \exp(it_1 B_{1j} + it_2 B_{2j}) \tag{12}$$

- recall that the log derivative is $\partial \ln \Psi/\partial t_1 = (\partial \Psi/\partial t_1)/\Psi$

Non-Parametric Estimation

- using $\partial \ln \hat{\Psi} / \partial t_1$ and the equation from the identification of Φ_Y, Φ_1, Φ_2, we get estimates $\hat{\Phi}_Y$, $\hat{\Phi}_1$, $\hat{\Phi}_2$
- we get $\hat{g}_1$ using the one-to-one mapping between f_1 and Φ_1
- we get $\hat{g}_2$ analogously
- we get $\hat{G}_1$ and $\hat{G}_2$ by integrating $\hat{g}_1$ and $\hat{g}_2$
- using $\hat{g}_1$, $\hat{g}_2$, $\hat{G}_1$, $\hat{G}_2$ we can use the method from Guerre, Perrigne, Vuong to estimate the inverse bidding functions β_1^{-1} and β_2^{-1}
- β_1^{-1} and β_2^{-1} give us the pseudo-valuations for every possible idiosyncratic bid
- $\Rightarrow$ the distributions of the idiosyncratic bids $A_1 \sim G_1$ and $A_2 \sim G_2$ give us the distributions of the idiosyncratic valuations $V_1 \sim F_1$ and $V_2 \sim F_2$