Recurring Themes in Auction Theory and Mechanism Design

Part IV: Some Empirics of Auctions

Dan Quint
U of Wisconsin
Visitor, UTokyo and UTMD

February 21, 2023
University of Tokyo

Overview

- Last three lectures: theory
- Today: empirics
- (But from a theory point of view)

Identification

The usual question in empirical auctions

- Lots of questions we might like to answer:
  - What's the optimal reserve price, and how much does it matter?
  - What if we switched to a different auction format?
  - How much is each incremental bidder worth?
  - What if we increased the information available to bidders?
  - What if we charged an entry fee?
  - What if we gave bid preferences to small firms/minority-owned businesses?

- To answer, we need to know the details of the environment
  - For private values models, this is distribution of bidder valuations
  - **Can we learn it from observed bid data?**

Identification

- When is a model identified?

Unobserved primitives	Model →	Observables
Parts of model we don't know ex ante • Preferences • Parameter values • Here: ***distribution F***	Maintained Assumptions • IPV • Symmetry • ***Equilibrium play***	Joint distribution of observable outcomes • Purchase decisions • Prices, demand • Here: ***bids***

- A model is a mapping from primitives to probability distributions over observable outcomes
- **A model is identified if this mapping is invertible**
  - With enough data to learn exact distribution of outcomes, we can uniquely pin down unspecified parts of the model

Example:
first price auctions

Symmetric IPV model is identified from bid data in first price auctions

- *n* bidders, symmetric independent private values $v_i \sim F$
- Bidder *i* solves $\max_b (v_i - b)\Pr(win|b)$
- In symmetric equilibrium, this is

$$\max_b (v_i - b)\big(G(b)\big)^{n-1}$$

  where $G(b)$ is CDF of an opponent's bids $\beta(v_j)$
- First-order condition is

$$-\big(G(b)\big)^{n-1} + (v_i - b)(n-1)\big(G(b)\big)^{n-2} g(b) = 0$$

- In symmetric equilibrium, this must hold at $b = \beta(v_i)$
- Plugging in $v_i = \beta^{-1}(b)$ and simplifying,

$$\beta^{-1}(b) = b + \frac{1}{n-1}\frac{G(b)}{g(b)}$$

Symmetric IPV model is identified from bid data in first price auctions

- So with *n* bidders and symmetric IPV, equilibrium implies

$$\beta^{-1}(b) = b + \frac{1}{n-1}\frac{G(b)}{g(b)}$$

- Good news: right-hand side is “data”!
  - Observe *n* and the distribution of bids
  - Plug in on RHS and impute $v_i = \beta^{-1}(b)$ for each observed bid
  - Distribution of imputed valuations is *F*
- Same idea generalizes…
  - Observable covariates
  - Risk-averse bidders
  - Correlated values

E Guerre, I Perrigne and Q Vuong (2000), Optimal Nonparametric Estimation of First-Price Auctions, *Econometrica* 68 (3)

What about ascending auctions?

What about ascending auctions?

- Again, n bidders, symmetric independent private values, want to learn F from bid data

- Two commonly-used abstractions for ascending auction:
  - Second-price sealed bid auction
  - "Button" auction
  - In both: dominant strategy to bid (or drop out at) your valuation
  - → allocation is efficient, transaction price = second-highest value

- Theorem: in second-price sealed bid or button auction, with fixed (known) number of bidders, F is identified from transaction prices

S Athey and P Haile (2002), Identification in Standard Auction Models, *Econometrica* 70 (6)

Symmetric IPV model is identified from *n* and transaction price in SPA or button auction

- Let F_T be distribution of transaction prices
- Transaction price = second-highest valuation
- Probability second-highest valuation is below *v* is

Prob all *n* valuations are below *v*

$$(F(v))^n + n\big(F(v)\big)^{n-1}(1 - F(v))$$

Prob exactly one valuation is above *v*

$$= n(F(v))^{n-1} - (n-1)(F(v))^n$$

- Define $\varphi(x) = nx^{n-1} - (n-1)x^n$, then

$$F_T(v) = \varphi(F(v))$$

- And φ is strictly increasing, so invertible, so

What we want to know! $F(v) = \varphi^{-1}(F_T(v))$ **Data!**

- (Additional bids reveal additional order statistics of valuations, so model is over-identified from bid data)

S Athey and P Haile (2002), Identification in Standard Auction Models, *Econometrica* 70 (6)

So that's the good news

- In either sealed-bid second-price or button auction, transaction price identifies symmetric IPV model
  - Again, extend to deal with observable covariates…
  - …or asymmetric bidders
- *But…*
  - Real-world ascending auctions aren't actually second-price sealed-bid or button auctions
  - And, the mapping between F and distribution of order statistics only holds for *independent* values

How to model open-outcry auctions?

Bidding in open-outcry ascending auctions

- Suppose you attended an art auction and the bidding looked like this:

Dan	15		21		26					
Fuhito		20		25		29		35		50
Kenzo							30		38	

- Not exactly clear what any of our valuations are!
- What would you infer?
  - If you believe private values and rational behavior…
  - Probably $v_{Kenzo} \geq 38$, $v_{Dan} \geq 26$, and $v_{Fuhito} \geq 50$
  - And perhaps $v_{Kenzo}, v_{Dan} \leq 51$
- Is that enough to work with?

Simple "behavioral" assumptions lead to upper and lower bounds on *F*

- Let
  - $v^{(k)}$ be *k^th* highest valuation out of *n* bidders, so $v^{(1)} \geq v^{(2)} \geq \cdots$
  - F_k be distribution of $v^{(k)}$
  - $b^{(k)}$ be highest bid from *k^th* highest bidder
  - G_k be distribution of $b^{(k)}$, and G_1^δ the distribution of $b^{(1)} + \delta$
  - φ_k be mapping from *F* to distribution of *k^th* highest of *n* independent draws from *F*
- If we assume $b^{(k)} \leq v^{(k)}$, this implies

$$G_k(v) \geq F_k(v) = \varphi_k(F(v)) \longrightarrow F(v) \leq \varphi_k^{-1}(G_k(v))$$

- And for *k* > 1, if we assume $v^{(k)} \leq b^{(1)} + \delta$, then

$$F_k(v) \geq G_1^\delta(v) \longrightarrow F(v) \geq \varphi_k^{-1}\left(G_1^\delta(v)\right)$$

- So we get upper and lower bounds on *F* from data!

P Haile and E Tamer (2003), Inference with an Incomplete Model of English Auctions, *Journal of Political Economy* 111 (1)

Simple "behavioral" assumptions lead to upper and lower bounds on *F*

- So for auctions of a given size *n*, we get…
  - *n* separate pointwise upper bounds for *F*(*v*)
  - one pointwise lower bound for *F*(*v*)
- If we have auctions with (exogenously) different numbers of bidders, we get additional bounds on *F*
- And bounds on *F* lead to bounds on optimal reserve price
- Bidding assumptions are pretty easy to swallow
- ***But,*** this still requires bidder valuations be independent (after controlling for observables)

P Haile and E Tamer (2003), Inference with an Incomplete Model of English Auctions, *Journal of Political Economy* 111 (1)

Ascending auctions with correlated values

What to do if bidder values are not independent?

- Without independence, no unique mapping between marginal and order statistic distributions

- Assume private values, just potentially correlated
  - Bidders themselves might perceive valuations as correlated…
  - …or as independent, conditional on observables they see but seller doesn't
  - For first-price auctions, these are different models…
  - …but for ascending auctions, observationally equivalent

- Unobserved primitive is no longer a marginal distribution, but entire *joint* distribution of bidders' valuations…

- …although only some parts matter for some purposes

Preliminaries

- Change notation: let $v^{k:n}$ be k^{th} *lowest* valuation, so $v^{n:n}$ is highest, $v^{n-1:n}$ second-highest, etc.
- Let $F_{k:n}$ be CDF of $v^{k:n}$
- For simplicity, let's assume transaction price is exactly second-highest valuation
  - Could work with behavioral assumptions of Haile and Tamer
- Revenue is $v^{n-1:n}$, or r if $v^{n:n} > r > v^{n-1:n}$, so

$$\pi(r,n) = (r - v_0)(F_{n-1:n}(r) - F_{n:n}(r)) + \int_r^\infty (v - v_0)dF_{n-1:n}(v)$$

- Depends only on two *marginal* distributions $F_{n-1:n}$ and $F_{n:n}$
- $F_{n-1:n}$ is "data" – so if we can put bounds on $F_{n:n}$, that suffices for reserve price counterfactuals

What can bid data tell us about $F_{n:n}$?

- With *independent* values, $F_{n-1:n} \rightarrow F \rightarrow F_{n:n}$
- Specifically,

$$F_{n:n}(v) = (F(v))^n = \varphi_1\left(\varphi_2^{-1}(F_{n-1:n}(v))\right)$$

where $\varphi_1(x) = x^n$ and $\varphi_2(x) = nx^{n-1} - (n-1)x^n$

- With standard formulations of symmetric, *correlated* values, this gives the lower bound ("best case scenario") for $F_{n:n}(v)$
  - For intuition, suppose $v_i \sim$ *i.i.d.* $F(\cdot\,|\theta)$
  - Then $F_{n:n}(v) = E_\theta(F(v|\theta))^n = E_\theta\varphi_1(\varphi_2^{-1}(F_{n-1:n}(v|\theta)))$
  - The function $\varphi_1 \circ \varphi_2^{-1}$ is convex
  - So by Jensen's Inequality,

$$F_{n:n}(v) = E_\theta\varphi_1 \circ \varphi_2^{-1}(F_{n-1:n}(v|\theta)) \geq \varphi_1 \circ \varphi_2^{-1}(E_\theta F_{n-1:n}(v|\theta))$$

$$= \left(\varphi_2^{-1}(F_{n-1:n}(v))\right)^n$$

D Quint (2008), Unobserved Correlation in Private-Value Ascending Auctions, *Economics Letters* 100 (3)

What can bid data tell us about $F_{n:n}$?

- With *independent* values, $F_{n-1:n} \to F \to F_{n:n}$
- Specifically,

$$F_{n:n}(v) = (F(v))^n = \varphi_1\left(\varphi_2^{-1}(F_{n-1:n}(v))\right)$$

where $\varphi_1(x) = x^n$ and $\varphi_2(x) = nx^{n-1} - (n-1)x^n$

- With standard formulations of symmetric, *correlated* values, this gives the lower bound (“best case scenario”) for $F_{n:n}(v)$
- Upper bound (“worst case”) is $F_{n:n}(v) = F_{n-1:n}(v)$ (perfect corr)
- This gives upper and lower bounds for $\pi(r,n)$ and optimal reserve price – but may be too wide to be useful
  - Optimal reserve ranges from $r^* = v_0$ to $r^* = r^*_{IPV}$
  - Losing bids can tighten upper bound on $\pi(r,n)$, but not the lower
  - Can’t falsify “all bidders had same valuation in each auction”

D Quint (2008), Unobserved Correlation in Private-Value Ascending Auctions, *Economics Letters* 100 (3)

What can we do to get point identification or tighter bounds?

- Three approaches for correlated values/unobserved heterogeneity in ascending auctions

1. Assume losing bids reveal more than one valuation
   - Suppose we're willing to assume the two highest losing bidders both bid all the way up to their valuations… or three… or more…
   - Very reasonable in button auction (where we'd observe all but highest order statistic), second-price auction
   - Several recent working papers give positive identification results
   - (Some for $v_i = \theta + \varepsilon_i$, some for general correlation)

E Mbakop, Identification of Auctions with Incomplete Bid Data in the Presence of UH
Y Luo and R Xiao, Identification of Auction Models Using Order Statistics
Y Luo, P Sang and R Xiao, Order Statistics Approaches to Unobserved Heterogeneity in Auctions
JH Cho, Y Luo and R Xiao, Deconvolution from Two Order Statistics

What can we do to get point identification or tighter bounds?

- Three approaches for correlated values/unobserved heterogeneity in ascending auctions

1. Assume losing bids reveal more than one valuation

2. Use variation in reserve price
   - Suppose $v_i = \theta + \varepsilon_i \ldots$
   - …and suppose seller knows θ, and reserve price is increasing in θ

J Roberts (2013), Unobserved Heterogeneity and Reserve Prices in Auctions, *RAND Journal of Economics* 44 (4)
J Freyberger and B Larsen (2022), Identification in Ascending Auctions, with an Application to Digital Rights Management, *Quantitative Economics* 13 (2)

What can we do to get point identification or tighter bounds?

- Three approaches for correlated values/unobserved heterogeneity in ascending auctions

1. Assume losing bids reveal more than one valuation

2. Use variation in reserve price

3. Use variation in number of bidders

A Aradillas-López, A Gandhi and D Quint (2013), Identification and Inference in Ascending Auctions with Correlated Private Values, *Econometrica* 81 (2)
D Coey, B Larsen, K Sweeney and C Waisman (2017), Ascending Auctions with Bidder Asymmetries, *Quantitative Economics* 8 (1)

Using variation in *n*

Goal: use knowledge of $F_{n-1:n}$ for various n to bound $F_{n:n}$

- Why should this work?
- As we add bidders, distribution of transaction prices shifts to the right
- If valuations are highly correlated, adding another bidder doesn't change transaction price much; if valuations are close to independent, it does
- "How fast" $F_{n-1:n}$ shifts with n tells how correlated values are, so how close $F_{n:n}$ is to $F_{n-1:n}$

Goal: use knowledge of $F_{n-1:n}$ for various n to bound $F_{n:n}$

- Thought experiment:
  - Start with auction with 6 bidders, possibly correlated values
  - Pick 5 of them at random, look at highest value among those 5
  - With probability 1/6, you dropped the one with the highest value, so highest remaining is second-highest of the original 6
  - With probability 5/6, you didn't drop the highest one, so highest remaining is highest of original 6
- Turns out that

$$F_{5:5}(v) = \frac{1}{6} F_{5:6}(v) + \frac{5}{6} F_{6:6}(v)$$

- Or more generally,

$$F_{n:n}(v) = \frac{1}{n+1} F_{n:n+1}(v) + \frac{n}{n+1} F_{n+1:n+1}(v)$$

So, for example…

$$
\begin{aligned}
F_{3:3}(v) &= \frac{1}{4}F_{3:4}(v) + \frac{3}{4}F_{4:4}(v) \\
&= \frac{1}{4}F_{3:4}(v) + \frac{3}{20}F_{4:5}(v) + \frac{3}{5}F_{5:5}(v) \\
&= \frac{1}{4}F_{3:4}(v) + \frac{3}{20}F_{4:5}(v) + \frac{1}{10}F_{5:6}(v) + \frac{3}{6}F_{6:6}(v) \\
&= \frac{1}{4}F_{3:4}(v) + \frac{3}{20}F_{4:5}(v) + \frac{1}{10}F_{5:6}(v) + \frac{1}{14}F_{6:7}(v) + \frac{3}{7}F_{7:7}(v) \\
&= \frac{1}{4}F_{3:4}(v) + \frac{3}{20}F_{4:5}(v) + \frac{1}{10}F_{5:6}(v) + \frac{1}{14}F_{6:7}(v) + \frac{3}{56}F_{7:8}(v) + \frac{3}{8}F_{8:8}(v)
\end{aligned}
$$

What we want

Something we can bound

"data"

Vanishing!

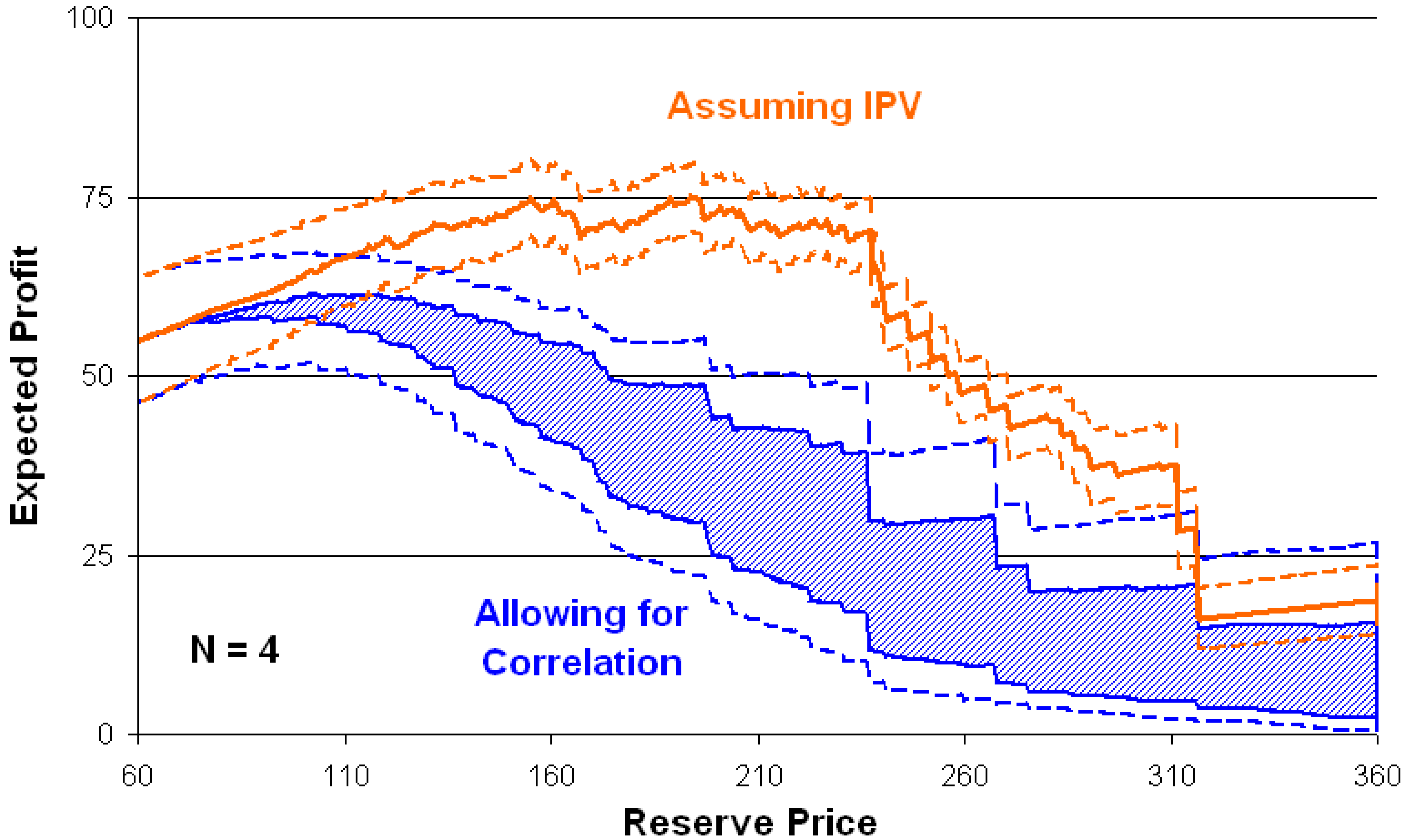
How well does it work?

- Data from US Forest Service timber auctions
  - Auctions for logging rights
  - “Scaled sales” (bids are per unit harvested)
  - Region 6 (Oregon), where bidders don’t conduct their own “cruises”
  - Short-term contracts, so little worry about resale

- 1,113 observations
  - Control for appraisal value and other key covariates
  - Number of bidders ranges from 2 to 11 (average 5.3)
  - Top two bids typically very close together

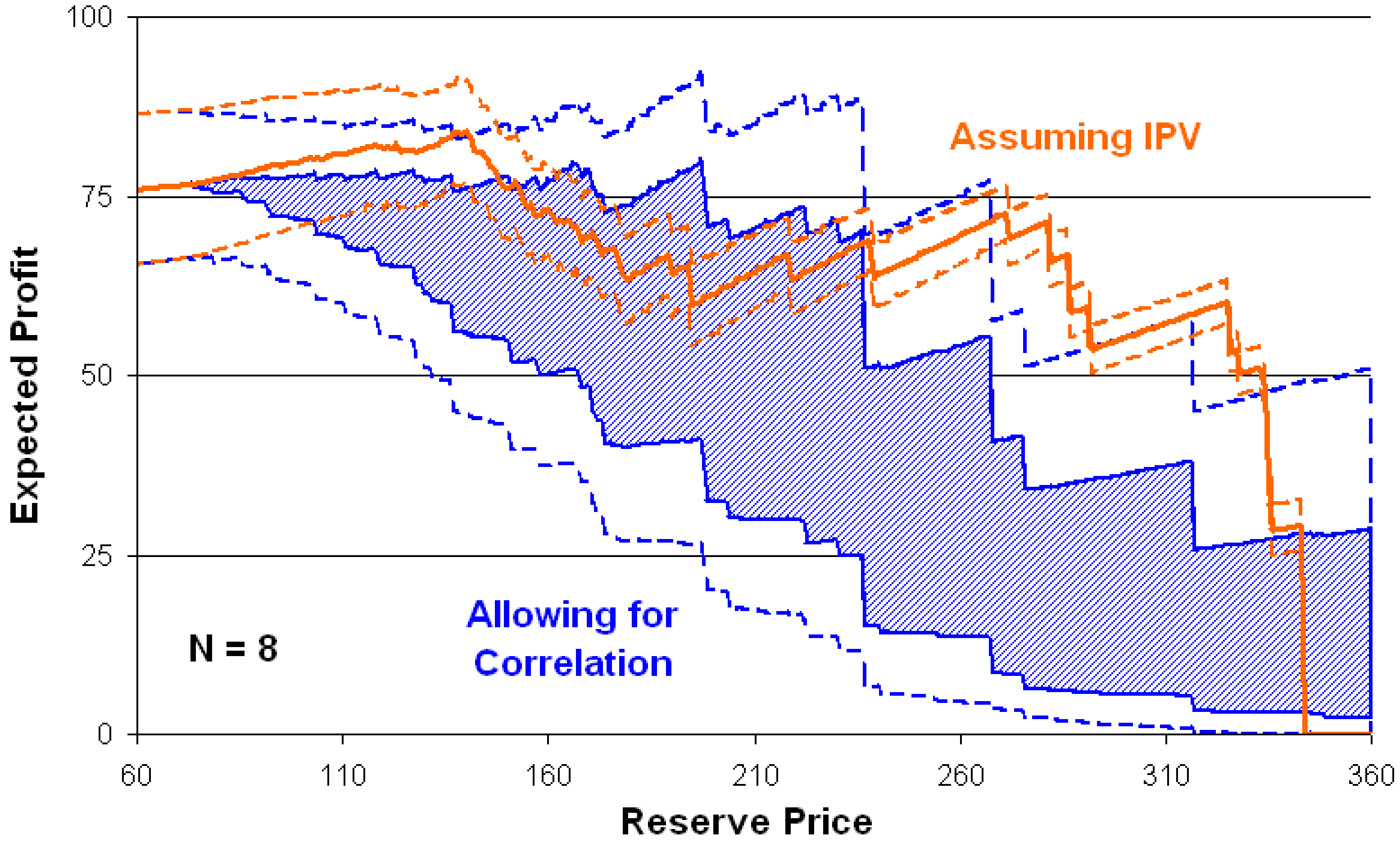
A Aradillas-López, A Gandhi and D Quint (2013), Identification and Inference in Ascending Auctions with Correlated Private Values, *Econometrica* 81 (2)

How well does it work?



A Aradillas-López, A Gandhi and D Quint (2013), Identification and Inference in Ascending Auctions with Correlated Private Values, *Econometrica* 81 (2)

How well does it work?



A Aradillas-López, A Gandhi and D Quint (2013), Identification and Inference in Ascending Auctions with Correlated Private Values, *Econometrica* 81 (2)

What if you worry *n* isn't exogenous?

- If auctions vary and bidders endogenously choose which to enter, valuations will not be independent of *n*
- Plausible case: more bidders when object is more valuable
  - Choose *k* bidders at random out of an *n*-bidder auction
  - If (probability at least one of the *k* has valuation $\geq v$) is increasing in *n*, we say "valuations stochastically increasing in *n*"
  - In that case, *upper* bound on $\pi(r,n)$ is still valid

A Aradillas-López, A Gandhi and D Quint (2013), Identification and Inference in Ascending Auctions with Correlated Private Values, *Econometrica* 81 (2)

What if you want point estimates rather than bounds?

- Suppose you're willing to assume…
  - $v_i = \theta + \varepsilon_i$, with θ, $\{\varepsilon_i\}$ independent of each other and *n*
  - transaction price = second-highest valuation

- We show…
  - If you observe $F_{T|n}$ for two values of *n*, the model is identified
  - If you don't observe *n* but you have an instrument *x*,
    know distribution of *n*|*x*,
    and observe $F_{T|x}$ for two values of *x*, the model is identified
  - If you observe "filtered *n*,"
    have the correct model of how real *n* maps to observed *n*,
    and have an instrument, then the model is identified

C Hernández, D Quint and C Turansick (2020), Estimation in English Auctions with Unobserved Heterogeneity, *RAND Journal of Economics* 51 (3)

What if you want point estimates rather than bounds?

- We use data from 15,000 eBay Motors sales
- Use “prime time” ending times as participation shifter
- Propose a model for how number of “potential bidders” leads to number of observed bids
- Semi-nonparametrically estimate distributions of θ and ε_i
- We decompose variation in log transaction prices into…
  - 83% variation in observables
  - 11% unobserved heterogeneity
  - 6% variation in idiosyncratic valuations
- We find consumer surplus estimate would be **260%** too high if we assumed IPV (conditional on observables)

C Hernández, D Quint and C Turansick (2020), Estimation in English Auctions with Unobserved Heterogeneity, *RAND Journal of Economics* 51 (3)

“Why not just control for observables better?”

- In both papers just cited, we controlled for observable variation in a fairly basic way
- Would apparent correlation vanish with better controls?
- eBay listings in 14 product categories
  - OLS analysis of “standard” dataset explained 0-15% of price dispersion
  - Machine learning model on full eBay listing (literally all the information buyers had) explains 48% of price dispersion
- But most people aren’t doing this
- Allowing for unobserved heterogeneity (or correlation) “lowers the stakes” of controlling for observables

A Bodoh-Creed, J Boehnke and B Hickman (2020), Using Machine Learning to Explain Violations of the “Law of One Price,” working paper

Early empirical puzzle: why are real-world reserve prices so low?

- Empirical takeaway from these papers: correlation or unobserved heterogeneity favor lower reserve prices
- So do…
  - Uncertainty in estimates of primitives
  - Endogenous entry
  - Competition between sellers
  - Common values
- Lots of deviations from "baseline" IPV model suggest lower optimal reserve prices

DJ Kim (2013), Optimal Choice of a Reserve Price under Uncertainty, *IJIO* 31 (5)
D Levin and J Smith (1994), Equilibrium in Auctions with Entry, *AER* 84 (3)
M Peters and S Severinov (1997), Competition among Sellers Who Offer Auctions Instead of Prices, *JET* 75
D Quint (2017), Common Values and Low Reserve Prices, *JINDEC* 65 (2)

Takeaways

Takeaways

- Today's question: when is an auction model identified?
  - what combinations of *modeling assumptions* and *observables* allow you to uniquely recover unobserved primitives of model?
  - (separate from: how to estimate on finite samples)
- Focus on ascending auctions
  - Under IPV assumption, button/second-price auctions identified from transaction prices and *n*
  - Under IPV and realistic bidding assumptions, ascending auction is set-identified, with useful bounds for many counterfactuals
  - Without IPV, things are harder
    - identification or useful bounds from multiple losing bids, endogenously-varying reserve price, or variation in number of bidders
  - Empirical work suggests correlation matters!
  - Correlation, among other things, favors lower reserve prices

Thank you!

References

E Guerre, I Perrigne and Q Vuong (2000), Optimal Nonparametric Estimation of First-Price Auctions, *Econometrica* 68 (3)

S Athey and P Haile (2002), Identification in Standard Auction Models, *Econometrica* 70 (6)

P Haile and E Tamer (2003), Inference with an Incomplete Model of English Auctions, *Journal of Political Economy* 111.1

D Quint (2008), Unobserved Correlation in Private-Value Ascending Auctions, *Econ Letters* 100.3

A Aradillas-López, A Gandhi and D Quint (2013), Identification and Inference in Ascending Auctions with Correlated Private Values, *Econometrica* 81 (2)

C Hernández, D Quint and C Turansick (2020), Estimation in English Auctions with Unobserved Heterogeneity, *RAND Journal of Economics* 51 (3)

A Bodoh-Creed, J Boehnke and B Hickman (2020), Using Machine Learning to Explain Violations of the "Law of One Price," working paper

DJ Kim (2013), Optimal Choice of a Reserve Price under Uncertainty, *International Journal of Industrial Organization* 31 (5)

D Levin and J Smith (1994), Equilibrium in Auctions with Entry, *Amer Economic Review* 84 (3)

M Peters and S Severinov (1997), Competition among Sellers Who Offer Auctions Instead of Prices, *Journal of Economic Theory* 75

D Quint (2017), Common Values and Low Reserve Prices, *J of Industrial Economics* 65 (2)

Thank you!

More References

E Mbakop, Identification of Auctions with Incomplete Bid Data in the Presence of UH

Y Luo and R Xiao, Identification of Auction Models Using Order Statistics

Y Luo, P Sang and R Xiao, Order Statistics Approaches to Unobserved Heterogeneity in Auctions

JH Cho, Y Luo and R Xiao, Deconvolution from Two Order Statistics

J Roberts (2013), Unobserved Heterogeneity and Reserve Prices in Auctions, *RAND J Econ* 44.4

J Freyberger and B Larsen (2022), Identification in Ascending Auctions, with an Application to Digital Rights Management, *Quantitative Economics* 13 (2)

D Coey, B Larsen, K Sweeney and C Waisman (2017), Ascending Auctions with Bidder Asymmetries, *Quantitative Economics* 8 (1)