Lecture Notes on Auctions and Nonlinear Pricing

by

Isabelle Perrigne and Quang Vuong
MOTIVATION

- Auctions: A market mechanism to sell/allocate goods widely used in practice

  → Basic Idea: Taking advantage of competition among heterogenous buyers/sellers

  → Examples: Gas leases, timber, real estate, cars, treasury bills, flowers, wine, eBay, collectibles, art, . . .
     but also procurements by public institutions (construction, maintenance, services,. . .)

  → Auction Theory: Models derived from game theory (Nash and Harsanyi, 1994 Nobel Laureates) to analyze
     bidders’ strategic behavior and optimal mechanism (efficient allocation and revenue comparison)

  → Data Availability

  → Beyond Auctions: Tools to analyze strategic behavior, allocation, markets and matching
MOTIVATION

• Nonlinear Pricing and Bundling: Pricing strategies widely used by firms

  → Basic Idea: Exploiting consumers heterogeneity to set prices and design products
      leading to price discount on large quantities or bundling products

  → Examples: Consumption goods (from paper towel to coffee), airline tickets, electricity, telecommunication,
      show tickets, insurance, . . .

  → Theory of Incentives: Models of incomplete information or *adverse selection* to determine firm’s optimal
      pricing through second and third degree price discrimination

  → Beyond Nonlinear Pricing: Tools to analyze differentiated products
ROADMAP

I. Structural Estimation of Auctions: The Basics

II. Extensions (reserve price, random reserve price, asymmetric bidders, bid preferences, affiliated values)

III. Unobserved Heterogeneity, Risk Aversion, Ascending Auctions

IV. Nonlinear Pricing, Bundling and Differentiated Products
STRUCTURAL APPROACH

• Rational behavior of agents

• A theoretical model leading to an econometric model

• Observables: Equilibrium outcomes

• Estimation of model primitives → Counterfactual Studies

• Three challenging problems/questions:

  (i) IDENTIFICATION: Can we recover uniquely the model primitives from the observables?

  (ii) MODEL TESTING: Does the model impose restrictions on observables? What are these restrictions?
      Can we distinguish models from observables?

  (iii) ESTIMATION: Can we find an estimation procedure that is computationally friendly?
I. STRUCTURAL ESTIMATION OF AUCTIONS: THE BASICS

Focus on first-price sealed-bid auctions

Model

• Assumptions:
  (i) \( I \geq 2 \) potential symmetric and risk neutral bidders
  (ii) indivisible good
  (iii) private value/willingness to pay \( v_i \sim F(\cdot) \) on \([v, \bar{v}]\), \( i = 1, \ldots, I \) the \( v_i \)s are iid
    → IPV model
  (iv) no reserve price

• Bidder’s optimization problem: \( b_i = s(v_i) \) equilibrium strategy
  \[
  \max_{b_i} (v_i - b_i) \Pr(\text{winning}) = (v_i - b_i) \Pr(b_i \geq b_j, j \neq i) = (v_i - b_i) F^{I-1}(s^{-1}(b_i))
  \]
  → Differential Equation (FOC)
  \[
  1 = (v_i - s(v_i))(I - 1) \frac{f(v_i)}{F(v_i)} \frac{1}{s'(v_i)}
  \]
  with boundary condition \( s(\bar{v}) = \bar{v} \).

Remarks: \( b_i = s(v_i, F, I) \), Bayesian Nash equilibrium, \( s(\cdot) \) strictly increasing, closed form solution for \( s(\cdot) \)
Estimation Method: Two Choices

- Direct approach: \( b_i = s(v_i, F, I) \), considers \( F(v_i; \gamma) \)

  Drawbacks: (i) requires numerical methods and computationally heavy
  
  (ii) explicit solutions do not always exist
  
  (iii) limited possible extensions

- Indirect approach: Exploit the one-to-one mapping between \( b_i \) and \( v_i \) through \( b_i = s(v_i, F, I) \) and rewrite the FOC to obtain an expression in terms of observables

Guerre, Perrigne and Vuong (2000, *Econometrica*)
Identification

Bid distribution $b_i \sim G(\cdot|I)$ on $[b, \bar{b}] = [v, s(\bar{v})]$, $i = 1, \ldots, I$

$$G(b|I) = \text{Pr}[B \leq b|I] = \text{Pr}[s^{-1}(B) \leq s^{-1}(b)|I] = \text{Pr}[V \leq s^{-1}(b)|I] = F(s^{-1}(b)) = F(v)$$

Bid density

$$g(b|I) = \frac{f(v)}{s'(v)}$$

Recall the FOC

$$1 = (v_i - s(v_i))(I - 1) \frac{f(v_i)}{F(v_i)} \frac{1}{s'(v_i)}$$

$$= \frac{g(b_i|I)}{G(b_i|I)}$$

$$\Rightarrow v_i = b_i + \left(\frac{1}{I - 1} \frac{G(b_i|I)}{g(b_i|I)}\right) \equiv \xi(b_i, G, I), \quad i = 1, \ldots, I$$

(1)

Mark-up
The RHS of (1) contains observables \((b_i, I, G(\cdot, I), g(\cdot|I), I)\) only

\[ \rightarrow F(\cdot) \text{ is nonparametrically identified} \]

i.e. we cannot find another value distribution \(\tilde{F}(\cdot)\) leading to the same bid distribution \(G(\cdot|I)\)

or \(F(\cdot)\) is recovered uniquely from observables
Model Restrictions/Rationalization

There exists a distribution $F(\cdot)$ that rationalizes the joint distribution of bids $G(\cdot, \ldots, \cdot)$ in a first-price sealed-bid auction with independent private values if and only if

(i) $G(b_1, b_2, \ldots, b_I) = \prod_{i=1}^{I} G(b_i | I)$ Independence and idness of bids

(ii) $\xi(\cdot, G, I)$ strictly increasing in $b$ Monotonicity of the equilibrium strategy

Remarks:

– If (i) and/or (ii) is not satisfied by the observables, then the model is inappropriate to explain the data

– Bid independence (i) is more restrictive

– A large class of bid distributions satisfies (ii)

– See recent papers for testing independence and monotonicity
Estimation

- A two-step procedure based on (1)

STEP 1: Estimate nonparametrically $G(\cdot|I)$ and $g(\cdot|I)$

$$\rightarrow \hat{G}(\cdot|I) \text{ and } \hat{g}(\cdot|I)$$

$$\rightarrow \text{estimated values}$$

$$\hat{v}_i = b_i + \frac{1}{I-1} \frac{\hat{G}(b_i|I)}{\hat{g}(b_i|I)}$$

STEP 2: Use the $\hat{v}_i$ to estimate nonparametrically the private value density $f(\cdot)$

$$\rightarrow \hat{f}(\cdot)$$
Review on Nonparametric Kernel Estimation

Kernel Density Estimation

• Problem: Estimate density $f(\cdot)$ of $X \in \mathbb{R}^p$ from a random sample $X_1, X_2, \ldots, X_n$

• Kernel density estimator of density $f(x)$ at arbitrary $x \in \mathbb{R}^p$

$$\hat{f}(x) = \frac{1}{nh^p} \sum_{i=1}^{n} K \left( \frac{x - X_i}{h} \right)$$

where $h > 0$ is a bandwidth and $K(\cdot)$ is a kernel, i.e. a function satisfying $\int_{\mathbb{R}^p} K(u) du = 1$

→ Motivation for (2): $p = 1$

$$f(x) = \lim_{h \downarrow 0} \frac{F(x + 0.5h) - F(x - 0.5h)}{h}$$

A natural estimator is

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} I(x - 0.5h < X_i \leq x + 0.5h) = \frac{1}{nh} \sum_{i=1}^{n} I(-0.5 \leq \frac{x - X_i}{h} < 0.5)$$

i.e. $K(u) = I(-0.5 \leq u < 0.5)$ the uniform kernel
Asymptotic Properties

Assumptions:
(i) $f(\cdot)$ is $R$-continuously differentiable
(ii) $h = h_n$, where $h_n \to 0$ and $nh_n^p / \log n \to +\infty$ as $n \to +\infty$.

For any compact subset $C$ inside the support of $f(\cdot)$

$$\sup_{x \in C} |\hat{f}(x) - f(x)| = O_{a.s.} \left( h^R + \sqrt{\frac{\log n}{nh^p}} \right)$$

The best rate of uniform convergence of $\hat{f}(\cdot)$ is $r^* = (n / \log n)^{R/(2R+p)}$

This rate is achieved when $h = h^* \propto (\log n/n)^{1/(2R+p)}$.

Remarks:
- Bias: $\mathbb{E}[\hat{f}(x)] - f(x) = O(h^R)$
- Variance: $\text{Var}[\hat{f}(x)] = O \left( \frac{1}{nh^p} \right)$
- minor assumptions on $K(\cdot)$
- boundary effects: $C$ must be inside the support of $f(\cdot)$

→ What is the best rate of uniform convergence among all possible estimators of $f(\cdot)$?

Stone (1982, *Annals of Statistics*) optimal rate $r^* = (n / \log n)^{R/(2R+p)}$
Kernel Estimation of Conditional Moments

- Problem: Estimate some conditional moments of \(X = (Y, Z) \in R^q \times R^d\) from a random sample \(X_1, X_2, \ldots, X_n\), where \(X_i = (Y_i, Z_i)\), \(Z\) continuous such as
  
  (i) \(E[Y|Z = \cdot]\) regression of \(Y\) on \(Z\)
  
  (ii) \(F_{Y|Z}(y|\cdot) \equiv \text{Pr}[Y \leq y|Z = \cdot]\) cdf of \(Y\) given \(Z\) at some fixed \(y\)

- Framework

Estimate the function \(m(\cdot) = E[\psi(Y)|Z = \cdot]\) where \(\psi(\cdot)\) is a given function such as

(i) \(\psi(Y) = Y\) in case (i)

(ii) \(\psi(Y) = \mathbb{I}(Y \leq y)\) in case (ii)

Remark: \(E[\psi(Y)|Z = z] = \int_{\mathbb{R}^d} \psi(y) f_{Y|Z}(y|z) dy\), where \(f_{Y|Z}(y|z) = f_{Y,Z}(y,z)/f(z)\)

A natural estimator is

\[
\hat{m}(z) = \int_{\mathbb{R}^d} \psi(y) \hat{f}_{Y|Z}(y|z) dy, \quad \text{where} \quad \hat{f}_{Y|Z}(y|z) = \frac{\hat{f}_{Y,Z}(y,z)}{\hat{f}(z)}
\]
• Kernel estimator of \( m(z) \) at arbitrary \( z \in \mathbb{R}^d \)

\[
\hat{m}(z) = \frac{\sum_{i=1}^{n} \psi(Y_i) K \left( \frac{z-Z_i}{h} \right)}{\sum_{i=1}^{n} K \left( \frac{z-Z_i}{h} \right)}
\]  

(3)

Remark: \( \hat{m}(z) = \hat{\phi}(z)/\hat{f}_Z(z) \) where

\[
\hat{\phi}(z) = \frac{1}{nh^d} \sum_{i=1}^{n} \psi(Y_i) K \left( \frac{z-Z_i}{h} \right) \quad \text{and} \quad \hat{f}_Z(z) = \frac{1}{nh^d} \sum_{i=1}^{n} K \left( \frac{z-Z_i}{h} \right)
\]

\[ \rightarrow \text{Motivation for (3): } q = d = 1 \]

\[
m(z) \equiv \mathbb{E}[\psi(Y)|Z = z] \approx \lim_{h \downarrow 0} \mathbb{E}[\psi(Y)|z - 0.5h < Z \leq z + 0.5h]
\]

A natural estimator is

\[
\hat{m}(z) = \frac{\frac{1}{n} \sum_{i=1}^{n} \psi(Y_i) \mathbb{I}(z - 0.5h < Z_i \leq z + 0.5h)}{\frac{1}{n} \sum_{i=1}^{n} \mathbb{I}(z - 0.5h < Z_i \leq z + 0.5h)}
\]

i.e. \( K(u) = \mathbb{I}(-0.5 \leq u < 0.5) \) uniform kernel
Asymptotic Properties

Assumptions:

(i) $m(\cdot)$ and $f_Z(\cdot)$ $R$-continuously differentiable
(ii) $h = h_n$, $h_n \to 0$ and $nh_d^n/\log n \to +\infty$ as $n \to +\infty$.

For any compact subset $C$ inside the support of $f_Z(\cdot)$

$$\sup_{z \in C} |\hat{m}(z) - m(z)| = O_{a.s.} \left( h^R + \left( \frac{\log n}{nh^d} \right)^{1/2} \right)$$

Thus, the best rate of uniform convergence of $\hat{m}(\cdot)$ is $r^* = (n/\log n)^{R/(2R+d)}$

This rate is achieved when $h = h^* \propto (\log n/n)^{1/(2R+d)}$.

Remarks:

- rate depends only on the dimension of $Z$
- boundary effects: $C$ must be inside the support of $f_Z(\cdot)$
- if $Z$ contains discrete variables, the rate depends only on the number of continuous variables in $Z$

→ What is the best rate of uniform convergence among all possible estimators of $m(\cdot)$?

A Two-Step Nonparametric Estimator

• Observations: \((B_{i\ell}, I_\ell, X_\ell), \ i = 1, \ldots, I_\ell, \ell = 1, \ldots, L\)

\[\rightarrow X_\ell: \text{Capturing observed heterogeneity of auctioned goods}\]

Examples: Engineering estimate, appraisal value, size, location, \ldots

• Assumptions:

  (i) \(X_\ell \in I^d\)

  (ii) \((X_\ell, I_\ell) \text{ iid } \sim F_m(\cdot, \cdot)\)

  (iii) Given \(\ell, V_{i\ell} \text{ iid conditionally upon } (X_\ell, I_\ell)\) as \(F(\cdot | X_\ell, I_\ell)\) on \([v(X_\ell, I_\ell), v(X_\ell, I_\ell)]\)

\[\rightarrow \text{more general than } F(\cdot | X_\ell): I_\ell \text{ can capture unobserved heterogeneity}\]
**Step 1:** Estimate $G(b|x, I)$ and $g(b|x, I)$

* Because of ratio, estimate $G(b, x, I)$ and $g(b, x, I)$

\[
\hat{G}(b, x, I) = \frac{1}{Lh_G^d} \sum_{\ell : I_\ell = I} \frac{1}{I_\ell} \sum_{i=1}^{I_\ell} \mathbb{I}(B_{i\ell} \leq b) K_G \left( \frac{x - X_\ell}{h_G} \right)
\]

\[
\hat{g}(b, x, I) = \frac{1}{Lh_g^{d+1}} \sum_{\ell : I_\ell = I} \frac{1}{I_\ell} \sum_{i=1}^{I_\ell} K_g \left( \frac{b - B_{i\ell}}{h_g} \right) K_g \left( \frac{x - X_\ell}{h_G} \right)
\]

**Remarks:**

- If $X$ is of dimension $d$, product of kernels
- If $X$ is discrete, treat it as $I$, i.e. perform estimation on the subset where $X$ equals a particular value

How to choose $K_G(\cdot)$, $K_g(\cdot)$, $h_G$ and $h_g$?

→ In view of previous review, relative freedom for choosing kernel, compact support is better

→ Choice of bandwidths requires attention, especially their vanishing rates
• Smoothness assumptions on $F(\cdot|\cdot, I)$:

(i) $F(\cdot|\cdot, I)$ is $R + 1$ continuously differentiable, $R \geq 1$

(ii) $f(\cdot|\cdot, I) > 0$

→ Because $F(\cdot|X, I)$ and $G(\cdot|X, I)$ are related through the BN equilibrium $b = s(v, F, I)$, it implies some smoothness properties on $G(\cdot|X, I)$ on $[b(X, I), \bar{b}(X, I)]$, namely

(iii) $g(\cdot|\cdot, I) > 0$

(iv) $G(\cdot|\cdot, I)$ is $R + 1$ continuously differentiable

(v) $g(\cdot|\cdot, I)$ is also $R + 1$ continuously differentiable (inside its support)!

From (1)

$$g(b|X, I) = \frac{G(b|X, I)}{(I - 1)\xi(b, G, I) - b}$$

$$R + 1$$

since $s(\cdot)$ is $R + 1$ (from the differential equation)
• Bandwidths:

\[
    h_G = \lambda_G L^{-\frac{1}{2R+d+2}}, \quad h_g = \lambda_g L^{-\frac{1}{2R+d+3}}, \quad \lambda = constant \times \hat{\sigma}
\]

where the constant is related to kernel choice and \(\hat{\sigma}\) is the standard deviation of the observations

\[
    \implies \hat{V}_{i\ell} = B_{i\ell} + \frac{1}{I_{\ell} - 1} \hat{G}(B_{i\ell}, X_{\ell}, I_{\ell}), \quad i = 1, \ldots, I_{\ell}, \quad \ell = 1, \ldots, L
\]

**STEP 2: Estimate \(f(v|x, I)\)**

\[
    \hat{f}(v|x, I) = \frac{\hat{f}(v, x, I)}{\hat{f}_m(x, I)} = \frac{\frac{1}{Lh_f^{d+1}} \sum_{\ell: I_{\ell}=I} \frac{1}{I_{\ell}} \sum_{i=1}^{I_{\ell}} K_f \left( \frac{v - \hat{V}_{i\ell}}{h_f} \right) K_f \left( \frac{x - X_{\ell}}{h_f} \right)}{\frac{1}{Lh_m^{d+1}} \sum_{\ell: I_{\ell}=I} \frac{1}{I_{\ell}} \sum_{i=1}^{I_{\ell}} K_m \left( \frac{x - X_{\ell}}{h_f} \right) }
\]

Two issues to solve: (i) Boundary effects in both steps

(ii) Choice of bandwidths \(h_f\) and \(h_m\)
• How to correct for boundary effects?

→ Trimming values too close to boundaries

\[
\tilde{\hat{V}}_{i\ell} = \begin{cases} 
\hat{V}_{i\ell} & \text{if } B_{\min} + h_g \leq B_{i\ell} \leq B_{\max} - h_g \\
+\infty & \text{otherwise}
\end{cases}
\]

Remarks:

– \(h_g\) can be too large and exclude too many values, especially low values if log normal shape

Solution: logarithm transformation \(\rightarrow\) log \(B_{i\ell}\)

– The \(\infty\) values ‘disappear’ in \(\hat{f}(\cdot|x,I)\) because of the compact support of kernel

– Requires to estimate \(B_{\min}\) and \(B_{\max}\)

If no \(X\), \(B_{\min} = \min_{i,\ell} B_{i\ell}\), \(B_{\max} = \max_{i,\ell} B_{i\ell}\)

If \(X\), make a partition of \(X\) and apply a similar idea
• How to choose $h_f$ and $h_m$?

Assuming $f_m(\cdot|x, I) R + 1$ continuously differentiable,  

$h_m = \lambda_m L^{-\frac{1}{2R+4+2}}$

* The vanishing rate of $h_f$ requires the derivation of the consistency rate of $\hat{f}(\cdot|x, I)$

Intuition: $f(v)$ and $g(b|I)$ are related through the BN equilibrium

$$f(v) = \frac{g(\xi^{-1}(v)|I)}{\xi'(\xi^{-1}(v))}$$

where $\xi'(b) = 1 + \frac{1}{I-1} \times \left[1 - \frac{G(b|I)g'(b|I)}{g^2(b|I)}\right]$

The derivative $g'(\cdot|I)$ is $R$ continuously differentiable. Thus $\hat{g}'(\cdot|I)$ has a bias of order $h^R$ and a variance of order $1/Lh^3$.

$$\to \text{MISE} = h^{2R} + (1/Lh^3), \quad \text{min}_h \text{MISE} \implies h \propto L^{-\frac{1}{2R+3}}$$

$g'(\cdot|I)$ has a consistency rate of $L^{\frac{R}{2R+3}}$. 

22
• Optimal convergence rate (Theorem 2, Guerre, Perrigne and Vuong (2000))

The best (optimal) rate for estimating \( f(\cdot|X,I) \) from observed bids is \( L^{\frac{R}{2R+d+3}} \)

Remarks:

- Rate slower than the standard optimal rate \( L^{\frac{R}{2R+d+1}} \) if the private values were observed
- By choosing \( h_f \propto L^{-\frac{1}{2R+d+3}} \), \( \hat{f}(\cdot|X,I) \) attains the optimal consistency rate
- Replace \( \hat{V}_{i\ell} \) by \( \tilde{V}_{i\ell} \) in estimator

• General comment on nonparametric estimation

\( L^{\frac{R}{2R+d+3}} \) is quite slow, in addition with \( d \) curse of dimensionality

Example: \( R = 2, d = 1, L = 200, L^{\frac{R}{2R+d+3}} = 200^{2/8} = 3.76 \)

If \( d = 2, L^{\frac{R}{2R+d+3}} = 200^{2/9} = 3.25 \)

Standard nonparametric rate \( 200^{2/6} = 5.84 \)

Parametric rate \( 200^{1/2} = 14.14 \)
• Assessing finite sample properties with Monte Carlo experiments

$L = 200, I = 5$

$v \sim \text{lognormal (+ truncation because of } [v, \bar{v}])$

→ Construct 1,000 bids using $b = s(v, F, I)$

→ 1,000 replications of the two-step estimator

Comments:

(i) pointwise 90% confidence interval

(ii) very good fit

(iii) slight boundary effect at the upper boundary

(iv) $\hat{\xi}(\cdot)$ increasing
Comments:

(i) wider confidence interval due to two-step estimation (accumulation of estimation errors)

(ii) substantial boundary effects on the LHS and RHS ⇒ trimming necessary

(iii) very good fit on the trimmed subset

• How to avoid trimming?

See Hickman and Hubbard (WP, 2012) who apply a boundary-corrected kernel density estimator method from the statistical literature, their Monte Carlo simulations show significant improvements
The case of Procurements

Allocating the project to the lowest bid

\[ c_i \sim F(\cdot) \text{ on } [c, \overline{c}] \text{ iid} \]

\[ \rightarrow \text{Firm's optimization problem: } b_i = s(c_i, F, I) \]

\[ \max_{b_i}(b_i - c_i)Pr(\text{winning}) = (b_i - c_i)Pr(b_i \leq b_j, j \neq i) = (b_i - c_i)(1 - F(s^{-1}(b_i)))^{I-1} \]

\[ \rightarrow \text{Differential equation (FOC)} \]

\[ 1 = (s(c_i) - b_i)(I - 1)\frac{f(c_i)}{1 - F(c_i)} \frac{1}{s'(c_i)} \]

with boundary condition \( s(\overline{v}) = \overline{v} \)

\[ c_i = b_i - \frac{1}{I - 1} \frac{1 - G(b_i|I)}{g(b_i|I)} \equiv \xi(b_i, G, I) \]

Mark-up

All the previous results apply!
II. EXTENSIONS

Reserve Price $p_0$

- Motivation: Widely used in practice, tool for optimal auction, can reduce collusion

- Assumptions
  
  (i) $p_0 > \underline{v}$ or $p_0 < \overline{c}$

  (ii) $p_0$ known by all bidders before bidding

Remarks: Binding reserve price

\[ \implies \text{Not all bidders participate, those with } v \in [\underline{v}, p_0) \text{ do not bid} \]

\[ \implies I^* \leq I \text{ observed number of participating or active bidders} \]

\[ \implies G^*(\cdot|I) \text{ truncated bid distribution as only } b_i \geq p_0, i = 1, \ldots, I^* \text{ are observed} \]

Question: How can we assess whether $p_0$ is binding?

Estimate the bid distribution in the neighborhood of $p_0$

\[ \implies \hat{G}^*(p_0 \leq b \leq \delta p_0) \text{ with } \delta = 1.05, 1.10, 1.15, \ldots \]

If this value is large, then $p_0$ is likely to be binding
• Model: Similar differential equation but different boundary condition \( s(p_0) = p_0 \)
  
  \[ b_i = s(v_i, F, p_0, I), \ i = 1, \ldots, I^* \]

• Identification

Truncated bid distribution: \( b_i \sim G^*(\cdot | I) \) on \([p_0, \bar{b}]\)

\[ G^*(b|I) = \Pr(B \leq b | I, B \geq p_0) = \Pr(s^{-1}(B) \leq s^{-1}(b) | I, B \geq p_0) = \Pr(V \leq s^{-1}(b) | I, V \geq p_0) = \frac{F(v) - F(p_0)}{1 - F(p_0)} \]

Truncated bid density

\[ g^*(b|I) = \frac{f(v)}{s'(v)(1 - F(p_0))} \text{ with } 1 - F(p_0) = \Pr(\text{being 'active'}) \]

Recall the FOC

\[ 1 = (v_i - s(v_i))(I - 1) \frac{f(v_i)}{F(v_i)} \frac{1}{s'(v_i)} = g^*(b_i|I)(1 - F(p_0)) \]

\[ = \frac{G^*(b_i|I)(1 - F(p_0))}{G^*(b_i|I)(1 - F(p_0)) + F(p_0)} \]

\[ \Rightarrow v_i = b_i + \frac{1}{I - 1} \left( \frac{G^*(b_i|I)}{g^*(b_i|I)} + \frac{F(p_0)}{1 - F(p_0)} \frac{1}{g^*(b_i|I)} \right) \equiv \xi(b_i, G, I, F(p_0)), \ i = 1, \ldots, I^* \] (1)
In the RHS of (1), we do not know $I$ and $F(p_0)$ but we observe $I^*$!

Solution: $I^* \sim \mathcal{B}(I, 1 - F(p_0))$

$\implies I$ and $F(p_0)$ are identified

Because $E(I^*) = I(1 - F(p_0))$ and $\text{Var}(I^*) = I(1 - F(p_0))F(p_0)$

$F(\cdot)$ is nonparametrically identified on $[p_0, \bar{v}]$

Remark: No data variation below $p_0$ to allow for identification
• Model Restrictions/Rationalization

Independence of bids (up to some technical details), \( \xi(\cdot, G, I, F(p_0)) \) increasing, 
\( I^* \sim \Pi(\cdot) \) with \( \Pi(\cdot) \) Binomial with parameters \((I, 1 - F(p_0))\)

• Estimation

Observations: \((B_{i\ell}, I^*_{\ell}, p_{0\ell}, X_{\ell}), i = 1, \ldots, I^*, \ell = 1, \ldots, L\)

Similar assumptions as before

A Two-Step Estimation Procedure

**STEP 1:** Estimate nonparametrically \( g^*(b, x, I) \) and \( G^*(b, x, I) \) as well as \( I \) and \( F(p_0|x) \)

\[ \rightarrow \hat{g}^*(b, x, I), \hat{G}^*(b, x, I), \hat{I}, \hat{F}(p_0|x) \]
Estimator for $I$ and $F(p_0|x) \equiv \phi(c)$ (with $p_0$ function of $x$ to avoid conditioning on $p_0$)

(i) Moments of the Binomial and nonparametric regression estimators for $\hat{E}[I^*_\ell|X_\ell]$ and $\hat{\text{Var}}[I^*_\ell|X_\ell]$ but $\hat{I}$ needs to be an integer!

(ii) $\hat{I} = \max_\ell I^*_\ell$ and $\hat{E}[I^*|x] = \hat{I}[1 - \hat{\phi}(x)]$ where

$$
\hat{\phi}(x) = 1 - \frac{1}{\hat{I}} \frac{1}{Lh_x} \sum_{\ell=1}^{L} I^*_\ell \frac{K_x \left( \frac{x-X_\ell}{h_x} \right)}{Lh_x} \sum_{\ell=1}^{L} K_x \left( \frac{x-X_\ell}{h_x} \right)
$$

with $h_x \propto L^{-\frac{1}{2p+1+2}}$

Remark: Maximum estimator converges faster but requires than $I$ constant across auctions

$$
\Rightarrow \hat{V}_{i\ell} = B_{i\ell} + \frac{1}{\hat{I} - 1} \left( \frac{\hat{G}^*(B_{i\ell}, X_\ell, \hat{I})}{\hat{g}^*(B_{i\ell}, X_\ell, \hat{I})} + \frac{\hat{\phi}(X_\ell)}{1 - \hat{\phi}(X_\ell)} \frac{1}{\hat{g}^*(B_{i\ell}|X_\ell, \hat{I})} \right), i = 1, \ldots, I^*_\ell, \ell = 1, \ldots, L
$$
**STEP 2:** Trim the estimated values $\hat{V}_{i\ell}$ and estimate nonparametrically $f^*(\cdot|x, \hat{I})$

$$\rightarrow \hat{f}(v|x, \hat{I}) = \hat{f}^*(v|x, \hat{I})[1 - \phi(\hat{x})]$$

Remarks:

(i) $\hat{f}(\cdot|\cdot, \cdot)$ still attains the optimal rate $L \frac{R}{2R+3+d+3}$

because $\hat{I}$ converges at rate $L$ and $\phi(\cdot)$ converges at rate $L \frac{R+1}{2R+d+2}$ up to some smoothness assumptions on $p_0$ as function of $x$

(ii) Estimator of $I$ assumes a fixed number of potential bidders. Estimation can be refined by defining subsets through $X$.


Luo, Perrigne and Vuong (2011, WP)

Subsets based on characteristics, location, . . .

(iii) See GPV (2000) for technical details because $g^*(p_0|x, I) = +\infty$
Random Reserve Price

Assumption: $p_0$ is unknown by bidders before bidding, secret reserve price

Examples: Sotheby’s, some timber auctions, some procurements, treasury bills, etc

Auction Rule: The auctioneer rejects ex post some bids that are below his reserve price


Assumptions:

(i) Seller’s value $v_0 \sim H(\cdot)$ on $[\underline{v}, \overline{v}]$

(ii) Risk neutral seller

(iii) $v_i$ and $v_0$ are independent

(iv) $[H(\cdot), F(\cdot)]$ common knowledge
• Model

Optimal strategy for the seller: \( p_0 = p(v_0) = v_0 \)

Intuition: \( p_0 < v_0 \) is irrational and \( p_0 > v_0 \) would lead to lost transactions

Bidder’s optimization problem: \( b_i = s(v_i, F, H, I) \) equilibrium strategy

\[
\max_{b_i} (v_i - b_i) \Pr(\text{winning}) = (v_i - b_i) \Pr(b_i \geq b_j, j \neq i, b_i \geq p_0) = (v_i - b_i) F^I(s^{-1}(b_i)) H(b_i)
\]

Differential Equation (FOC)

\[
1 = (v_i - s(v_i)) \left[ (I - 1) \frac{f(v_i)}{F(v_i)} \frac{1}{s'(v_i)} + \frac{h(s(v_i))}{H(s(v_i))} \right]
\]

with boundary condition \( s(\underline{v}) = \underline{v} \).

Remark: No closed form solution exists. The indirect approach becomes very useful!
• Identification

\[ F(\cdot), \ H(\cdot), \ I \quad G(\cdot|I), \ P(\cdot), \ I \]

Model Primitives \hspace{1cm} Observations

Can we identify?

Remarks:

(i) \( P(\cdot) \) distribution of reserve prices \( p_0 \) (if \( p_0 \) is not observed, one can assume \( H(\cdot) = F(\cdot) \))

(ii) \( I \) is observed since all bidders participate!

(iii) Because \( p_0 = v_0 \), \( P(\cdot) = H(\cdot) \)
Using \( G(b_i) = F(v_i) \), \( g(b_i) = f(v_i)/s'(v_i) \) and the observed \( H(\cdot) \), the FOC becomes

\[
v_i = b_i + \frac{1}{(I - 1)} \frac{g(b_i|I)}{G(b_i|I)} + \frac{h(b_i)}{H(b_i)} \equiv \xi(b_i, G, H, I), i = 1, \ldots, I
\]  

(2)

The RHS of (2) contains observables only. \( F(\cdot) \) is nonparametrically identified

- **Model Restrictions/Rationalization**

Independence of bids, \( \xi(\cdot, G, H, I) \) increasing and \( p_0 \) independent of \((b_1, \ldots, b_I)\)

- **Estimation**

Observables \((B_{i\ell}, p_{0\ell}, X_{\ell}, I_{\ell})\), \( i = 1, \ldots, I_{\ell}, \ell = 1, \ldots, L \)

**A Two-Step Procedure**

**STEP 1:** Estimate nonparametrically \( G(b, x, I) \) and \( g(b, x, I) \) as well as \( H(p_0, x) \) and \( h(p_0, x) \)

\[
\rightarrow \hat{V}_{i\ell} = \hat{B}_{i\ell} + \frac{1}{(I_{\ell} - 1)} \frac{\hat{g}(B_{i\ell}, X_{\ell}, I_{\ell})}{\hat{G}(B_{i\ell}, X_{\ell}, I_{\ell})} + \frac{\hat{h}(B_{i\ell}, X_{\ell})}{\hat{H}(B_{i\ell}, X_{\ell})}, \ i = 1, \ldots, I_{\ell}, \ell = 1, \ldots, L
\]
**Step 2:** Similar as before $\rightarrow \hat{f}(v|x, I)$

**Remarks:**
- Bandwidths for estimating $H(\cdot, x, I)$ and $h(\cdot, x, I)$ are the same as those for estimating $G(\cdot, x, I)$ and $g(\cdot, x, I)$ up to some smoothness assumptions on $h(\cdot, x, I)$, i.e. $R + 1$ differentiability
- Optimal rate for $\hat{f}(v|x, I)$ applies

**Application to French Timber Auctions**

- Data: French public timber, 74 auctions for a total of 212 bids

<table>
<thead>
<tr>
<th>Variable</th>
<th>Number of Observations</th>
<th>Mean</th>
<th>S.D.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bids (FF/m³)</td>
<td>212</td>
<td>124.73</td>
<td>83.57</td>
<td>7.63</td>
<td>660.00</td>
</tr>
<tr>
<td>Winning bids (FF/m³)</td>
<td>46</td>
<td>139.87</td>
<td>94.06</td>
<td>21.80</td>
<td>660.00</td>
</tr>
<tr>
<td>Reserve price (FF/m³)</td>
<td>74</td>
<td>141.73</td>
<td>111.69</td>
<td>21.80</td>
<td>900.00</td>
</tr>
<tr>
<td>Area (ha)</td>
<td>74</td>
<td>11.23</td>
<td>7.53</td>
<td>1.10</td>
<td>33.00</td>
</tr>
<tr>
<td>Volume (m³)</td>
<td>74</td>
<td>799.36</td>
<td>484.39</td>
<td>102.00</td>
<td>2416.00</td>
</tr>
<tr>
<td>Density (m³/ha)</td>
<td>74</td>
<td>94.81</td>
<td>68.95</td>
<td>24.35</td>
<td>311.18</td>
</tr>
<tr>
<td>Appraisal value (FF/m³)</td>
<td>74</td>
<td>165.82</td>
<td>243.48</td>
<td>26.92</td>
<td>1051.72</td>
</tr>
<tr>
<td>Percentage of saw timber</td>
<td>74</td>
<td>70.34</td>
<td>22.57</td>
<td>2.24</td>
<td>100.00</td>
</tr>
<tr>
<td>Number of bidders</td>
<td>74</td>
<td>2.87</td>
<td>1.15</td>
<td>2</td>
<td>7</td>
</tr>
</tbody>
</table>
Comments:

- 38% of auctions are rejected (47% for auctions with $I = 2$, 23% for auctions with $I = 4$, 0% after)
- important heterogeneity across lots due to different species
  → Appraisal value captures observed heterogeneity

• Empirical Results

(i) $\hat{v}_i \in [33.5, 639.6]$ after 10% trimming
(ii) comparison of $\hat{E}(v|x)$ and $\hat{E}(p_0|x)$ shows that bidders value timber more than the seller
(iii) $\hat{\xi}(\cdot)$ strictly increasing
(iv) $\hat{h}(\cdot|x, I)$ and $\hat{f}(\cdot|x, I)$ quite different (see next slide)
(v) IR (informational rent) $\sim 28\%$ because of low competition
(vi) IR increasing in appraisal value and decreasing in $I$
Counterfactuals: Secret reserve prices, a puzzle for theorists as the seller’s commitment in $p_0$ increases profit

Optimal announced reserve price: $p_0^*$ solving $p_0^* = v_0 + \frac{1-F(p_0^*)}{f(p_0^*)}$

Based on empirical results $p_0^* \in [74.8, 358.6]$, 46% larger than observed $p_0$!

→ only 24 auctions will receive bids

→ simulated optimal revenue = 5,364,019 and profit = 1,964,519

actual auction revenue = 5,281,378 and profit = 876,668

→ IR become 15%

Alternative justifications for random reserve prices: bidders’ risk aversion, entry, common value
Asymmetric Bidders

- Motivation: Bidders vary in size, location, capacity/backlog, ... 
  i.e. exogenous factors affect their value distribution

Examples:

(i) large vs small firms (efficiency, scale economies, ...) 

(ii) distance matters in construction procurements

(iii) collusion introduces asymmetry!


Campo, Perrigne and Vuong (2003, *Journal of Applied Econometrics*): Solo vs cartel in gas leases

and many others who apply a similar method such as


• Model

Assumptions:

(i) two types of bidders: Weak or type ‘0’, strong or type ‘1’

(ii) Bidders’ type known to all bidders

(iii) $v_{0i} \sim F_0(\cdot) \text{ iid, } v_{1i} \sim F_1(\cdot) \text{ iid, both distributions on } [\underline{v}, \overline{v}]$

(iv) $I_0$ bidders of type ‘0’, $I_1$ bidders of type ‘1’, $I_0 + I_1 = I \geq 2$

(v) $(v_{01}, \ldots, v_{0I_0}, v_{11}, \ldots, v_{1I_1})$ mutually independent (see CPV (2003) for asymmetry with affiliation)

$F_1(\cdot)$ stochastically dominates $F_0(\cdot)$

for $\alpha \in (0, 1)$, $v_0(\alpha) < v_1(\alpha)$

‘0’ bidders tend to draw lower values than ‘1’ bidders
Bidder ‘1’ optimization problem, $i = 1, \ldots, I_1$

$$\max_{b_{1i}}(v_{1i} - b_{1i}) \Pr(\text{winning}) = (v_{1i} - b_{1i}) \Pr(b_{1i} \geq b_{1j} \text{ and } b_{1i} \geq b_{0j}, j \neq i) = (v_{1i} - b_{1i}) F_1^{-1}(s_1^{-1}(b_{1i})) F_0^{-1}(s_0^{-1}(b_{1i}))$$

Differential equation (FOC)

$$(v_{1i} - b_{1i}) \left[ (I_1 - 1) \frac{1}{s_1'(s_1^{-1}(b_{1i}))} \frac{f_1(s_1^{-1}(b_{1i}))}{F_1(s_1^{-1}(b_{1i}))} + I_0 \frac{1}{s_0'(s_0^{-1}(b_{1i}))} \frac{f_0(s_0^{-1}(b_{1i}))}{F_0(s_0^{-1}(b_{1i}))} \right] = 1$$

Similarly, for a ‘0’ bidder, $i = 1, \ldots, I$

$$(v_{0i} - b_{0i}) \left[ I_1 \frac{1}{s_1'(s_1^{-1}(b_{0i}))} \frac{f_1(s_1^{-1}(b_{0i}))}{F_1(s_1^{-1}(b_{0i}))} + (I_0 - 1) \frac{1}{s_0'(s_0^{-1}(b_{0i}))} \frac{f_0(s_0^{-1}(b_{0i}))}{F_0(s_0^{-1}(b_{0i}))} \right] = 1$$

$\rightarrow$ Equilibrium strategies $s_0(\cdot, F_0, F_1, I_0, I_1)$ and $s_1(\cdot, F_0, F_1, I_0, I_1)$ solutions of this system of differential equations with boundary conditions $s_0(\nu) = s_1(\nu) = \nu$ and $s_0(\nu) = s_1(\nu)$
Equilibrium Properties:

(i) No explicit solution to the intractable system but equilibrium exists and is unique

(ii) Because of their ‘disadvantages’, ‘0’ bidders bid more aggressively to compensate for their low values

(iii) In view of (ii), the winner might not be the bidder with the largest value

⇒ Inefficient allocation

(iv) Affects competition, ‘1’ bidders are more sensitive to competition in their own group
Identification

Key assumption: Bidders’ type known to the analyst

(if anonymous bidders, Lamy (2012, Journal of Econometrics))

\[ F_0(\cdot), F_1(\cdot), I_0, I_1 \quad G_0(\cdot|I_1, I_0), G_1(\cdot|I_0, I_1), I_0, I_1 \]

Model Primitives Observations

Can we identify?

Using \( G_1(b_1|i|I_1, I_0) = F_1(v_{1i}) \), \( g_1(b_1|i|I_0, I_1) = f_1(v_{1i})/s'_1(v_{1i}) \), \( G_0(b_0|i|I_1, I_0) = F_0(v_{0i}) \), \( g_0(b_0|i|I_0, I_1) = f_0(v_{0i})/s'_0(v_{0i}) \)

\[ v_{1i} = b_{1i} + \frac{1}{(I_1 - 1) \frac{g_1(b_{1i}|I_0, I_1)}{G_1(b_{1i}|I_0, I_1)} + I_0 \frac{g_0(b_{1i}|I_0, I_1)}{G_0(b_{1i}|I_0, I_1)}} = \xi_1(b_{1i}, G_0, G_1, I_0, I_1), i = 1, \ldots, I_1 \]

\[ v_{0i} = b_{0i} + \frac{1}{I_1 \frac{g_1(b_{0i}|I_0, I_1)}{G_1(b_{0i}|I_0, I_1)} + (I_0 - 1) \frac{g_0(b_{0i}|I_0, I_1)}{G_0(b_{0i}|I_0, I_1)}} = \xi_0(b_{0i}, G_0, G_1, I_0, I_1), i = 1, \ldots, I_0 \]

(3)
The RHS of (3) contain observables only. 

$F_0(\cdot)$ and $F_1(\cdot)$ are nonparametrically identified

- Model Restrictions/Rationalization: Bid independence, 
  
  $\xi_1(\cdot, G_0, G_1, I_0, I_1)$ and $\xi_0(\cdot, G_0, G_1, I_0, I_1)$ are both increasing

- Estimation:

Observables $(B_{1i\ell}, B_{0j\ell}, I_{1\ell}, I_{0\ell}, X_{\ell}), i = 1, \ldots, I_{1\ell}, j = 1, \ldots, I_{0\ell}, \ell = 1, \ldots, L$

Follow the two-step procedure, conduct estimation for types 0 and 1

- How to detect asymmetry?
  
  - Define asymmetry in terms of exogenous variables
  
  - Test equality of the bid distributions given $X$ (Kolmogorov-Smirnov test)

  - Strong bidders tend to win auctions more often than weak ones

  - Bid regressions show different patterns of sensibility to competition, etc
APPLICATION TO SNOW REMOVAL PROCUREMENTS

- Data: 61 procurements organized by the City of Montreal, 457 bids

- Empirical Evidence:
  (i) high level of competition $I^* \in [5, 11]$
  (ii) binding reserve price: $\Pr(\delta p_0 \leq b \leq p_0) = 0.19$ for $\delta = 0.95$ ($0.33$ for $\delta = 0.90$)
  (iii) heavy equipments (trucks, snowploughs, sweepers) requiring close-by storage in winter
  (iv) west Montreal urbanized, east Montreal industrial $\rightarrow$ cost advantage/disadvantage relative to the firm’s position to the job location
  (v) 32 east tracts, 29 west tracts, 8.2 bidders on east tracts, 6.8 on west tracts
  (vi) east firms tend to bid more on east tracts (4.7 vs 1.8),
       west firms tend to bid more on west tracts (2.8 vs 1.4)
  (vii) east tracts: 57% of bids by east firms but 72% won by east firms
        west tracts: 41% of bids by west firms but 62% won by west firms
(viii) different bid sensitivity on west and east tracts

(ix) bid distributions: West tracts display stochastic dominance

- Empirical Results

(i) \( \hat{\xi}_0 W(\cdot) \) and \( \hat{\xi}_1 W(\cdot) \) increasing

(ii) \( \hat{\xi}_1 W(\cdot) > \hat{\xi}_0 W(\cdot) \) at \( c \) given

(iii) \( \hat{\xi}_0 E(\cdot) \approx \hat{\xi}_1 E(\cdot) \)

(iv) cost difference of 69 cents per meter on west tracts

(v) \( IR \approx 8\% \) for strong firms and 4\% for weak firms

(vi) cost densities differ on west tracts
• Counterfactuals:

  - 24% of auctions are inefficient, the winner has a cost 6.6% larger than the most efficient firm

  - discriminatory prices for '0' and '1' → cost reduction by 1.1%, weak firms are more likely to win

  - subsidy of 65 cents per meter to weak firms (to reestablish symmetry)
    → cost reduction of 2.4% (1% with cost of public funds at 0.3) at actual participation
Bid Preference

• Motivation: Public institutions grant a bid preference/discount/credit to some firms such as minority, small or domestic firms, . . .
but can negatively affect participation of nonfavored firms and lead to inefficient allocation!

Examples:

Penalty of 50% to foreign firms on defense contracts
Credit of 25% to small firms in the FCC spectrum auctions
Discount of 5% to small firms in Caltrans procurements

Krasnokutskaya and Seim (2011, American Economic Review)

• Model

Assumptions

(i) $I_f$ favored (small) firms, $c_f \sim F_f(\cdot)$
    $I_n$ nonfavored (large) firms, $c_n \sim F_n(\cdot)$

(ii) a ‘f’ bidder wins the auction if $b_f < (1 + \delta)b_n$, $b_n$ lowest ‘n’ bid and is paid $b_f$
Favored firm optimization problem: \[ \max_{b_{fi}} (b_{fi} - c_{fi})(1 - F_f (s_f^{-1}(b_{fi}))^{I_f-1})(1 - F_n (s_n^{-1}(b_{fi}/(1 + \delta))))^{I_n} \]

Nonfavored firm optimization problem: \[ \max_{b_{ni}} (b_{ni} - c_{ni})(1 - F_f (s_f^{-1}((1 + \delta)b_{ni}))^{I_f}(1 - F_n (s_n^{-1}(b_{ni})))^{I_n-1} \]

\[ \implies \] complex system of differential equations with appropriate boundary conditions as in asymmetric auctions

Using bid distributions \( G_f(\cdot|I_f, I_n) \) and \( G_n(\cdot|I_f, I_n) \)

\[ c_{fi} = b_{fi} - \frac{1}{(I_f - 1) \frac{g_f(b_{fi}, I_f, I_n)}{1 - G_f(b_{fi}, I_f, I_n)} + \frac{I_n}{1 + \delta} \frac{g_n(b_{fi}/(1+\delta), I_f, I_n)}{1 - G_n(b_{fi}/(1+\delta), I_f, I_n)}} = \xi_f(b_{fi}, G_f, G_n, I_f, I_n), i = 1, \ldots, I_f \]

\[ c_{ni} = b_{ni} - \frac{1}{I_f \frac{(1 + \delta)g_f(b_{ni}, I_f, I_n)}{1 - G_f((1 + \delta)b_{ni}, I_f, I_n)} + (I_n - 1) \frac{g_n(b_{ni}, I_f, I_n)}{1 - G_n(b_{ni}, I_f, I_n)}} = \xi_f(b_{ni}, G_f, G_n, I_f, I_n), i = 1, \ldots, I_n \]

• Identification: With \( \delta \) known, RHS of (4) contain observables only
  \[ \implies F_f(\cdot) \text{ and } F_n(\cdot) \text{ are nonparametrically identified} \]

• Model Restrictions/Rationalization: Same as for asymmetric auctions

• Estimation: Observations \( (B_{f,i}, B_{n,j}, I_{f,i}, I_{n,j}, X_{\ell}), i = 1, \ldots, I_{f,\ell}, j = 1, \ldots, I_{n,\ell}, \ell = 1, \ldots, L \)
  Follow the two-step procedure for asymmetric auctions
APPLICATION TO CALTRANS PREFERENCE PROCUREMENTS

• Data:
  (i) 4,136 highway procurements by Caltrans (2,091 state-funded, 2,045 federal-funded),
  (ii) $\delta = 0.05$ to small firms in state-funded procurements
  (iii) small firm: independently owned and operated, California based, < 100 employees, < $10,000K$ revenue
  (iv) to establish comparison, focus on 2612 procurements with estimate < $1$ million

• Empirical Evidence:
  – small firms tend to participate more to state contracts (2.42 vs 2.33)
    but less participation of large firms (2.89 vs 3.66)
  – small firms tend to win more state contracts (45% vs 33%)
  – 7% of state contracts winners have changed because of preference
  – small firms tend to bid more aggressively
  – $X$: Engineering estimate and number of items
  – preference affects the number of small and large bidders
Empirical Results ($2 \leq I_n \leq 5, 1 \leq I_f \leq 4$)

(i) cost difference in federal contracts
(ii) little cost difference in state contracts except for low costs
(iii) 42 out of 895 state project allocations are inefficient
(iv) 6 out of 417 federal project allocations are inefficient
(v) preference add inefficiency but remains small $\sim 0.09\%$

Counterfactuals (using numerical methods and Weibull specification):

(i) increase large firms participation in state contracts $\rightarrow$ lower procurement cost thanks to large firm bids
(ii) is $\delta = 0.05$ optimal? Increasing $\delta$ leads to more inefficiency
Application to Caltrans Preference Procurements with Endogenous Participation

- Motivation: Why consider endogenous participation?
  
  (i) procurement cost varies much more with endogenous participation and conclusions differ
  20% to small firms vs 50% to large firms!

  (ii) Caltrans allocative goal of 25% ($) to small firms with endogenous participation, goal reached with
  \[ \delta = 0.2 \] and procurement cost +3.2%

- Data: 697 state contracts, 39% allocated to small firms, 5% of contracts winners have changed because of preference

- Empirical Evidence:
  
  - Probit for potential bidder to submit a bid: More competition decreases entry, especially for small firms
    Is it due to larger entry costs?
  
  - Bid regression, small firms tend to bid less aggressively
    Is it due to larger costs or preference?
• Model

Assumptions:
- Entry costs: $K_{fi} \sim H_f(\cdot)$, $K_{ni} \sim H_n(\cdot)$
- Potential bidders: $I_f, I_n$
- Cost asymmetry: $c_{fi} \sim F_f(\cdot)$, $c_{ni} \sim F_n(\cdot)$
- Actual bidders (after entry): $I^*_f, I^*_n$

Two-stage game: 1) Bidder occurs an entry cost and decides to enter based on expected profit from entry
   2) Upon entering, bidder knows his production cost and bids

Entry Game: Computation of ex ante expected profit with respect to cost and actual number of bidders
   $\rightarrow$ For ‘f’ bidder, $\Pr(I^*_f - 1, I^*_n|I_f, I_n)$
   $\rightarrow$ For ‘n’ bidder, $\Pr(I^*_f, I^*_n - 1|I_f, I_n)$

Using Binomial with $p_f, p_n$ probabilities of entering
   $\rightarrow \tilde{\Pi}_f(p_f, p_n) - K_f = 0 \rightarrow K_f(p_f, p_n)$
   $\rightarrow \tilde{\Pi}_n(p_f, p_n) - K_n = 0 \rightarrow K_n(p_f, p_n)$
Equilibrium: $p_f = H_f(K_f(p_f, p_n))$ and $p_n = H_n(K_n(p_f, p_n))$

* Equilibrium exists but is not unique!

Bidding Game: As in asymmetric auction

- Estimation:
  - Observed number of plan holders $I_f = 3.9, I_n = 6.6$ ($I_f^* = 1.7, I_n^* = 2.6$)
  - Because of a large number of covariates (work type, location), parameterization of $G_f(\cdot), G_n(\cdot), H_f(\cdot), H_n(\cdot)$
  - GMM estimator
  - Costs estimated using GPV

- Empirical Results:
  
  (i) Costs larger for small firms except for rural road work
  
  (ii) Entry costs account for 2.2% to 3.9% of engineering estimate
  
  (iii) Entry costs tend to increase with project size for small firms but decrease in size for large bidders
  
  (iv) Entry cost distributions differ between small and large firms
Counterfactuals (using numerical algorithms for asymmetric auctions)

(i) preference to small firms favors their entry but reduces that of large ones but overall + entry
(ii) with endogenous entry, small firms more likely to win
(iii) procurement costs increases despite more competition

What is the optimal discount to minimize procurement cost?

→ Favor large firms instead of small firms
→ Favor the entry of large firms because of their lower costs

How to attain the 25% target?

→ \( \delta = 0.45 \) and cost increases by 6.6% with no entry model
→ \( \delta = 0.15 \) and cost increases by 1.4% with endogenous entry model
Affiliated Values

• Motivation: independence of private values is a strong assumption

Affiliation: A large value for a bidder makes the other bidders more likely to draw large than small values, more general than positive dependence

Li, Perrigne and Vuong (2002, *Rand Journal of Economics*)

• Model

Assumptions:

(i) \((v_1, \ldots, v_I) \sim F(\cdot, \ldots, \cdot)\) on \([\underline{v}, \overline{v}]^I\)

(ii) \(F(\cdot, \ldots, \cdot)\) is exchangeable (symmetric) and affiliated

→ APV model

Bidder optimization problem: \(b_i = s(v_i, F, I)\) equilibrium strategy

\[
\max_{b_i}(v_i - b_i) \Pr(\text{winning}) = (v_i - b_i) \Pr(B_i \geq b_i | v_i) = (v_i - b_i) F_{y_i|v_i}(v_i, v_i)
\]

where \(B_i = s(y_i)\), \(y_i = \max_{j \neq i} v_j\), boundary condition \(s(\underline{v}, F, I) = \underline{v}\)

31
→ Differential Equation (FOC): Considering bidder ‘1’ since they are all alike (symmetry)

\[ 1 = (v_i - s(v_i)) \frac{f_{y|v_i}(v_i|v_i)}{F_{y|v_1}(v_i|v_i)} \frac{1}{s'(v_i)} \]

with boundary condition \( s(v) = v \)

Remarks:

- A closed form solution exists
- \( F_{y_1,v_1}(y,v) = F(v,y,...,y) \)

- Identification

\[
F(\cdot,\ldots,\cdot), I \quad G(\cdot,\ldots,\cdot), I
\]

Model Primitives \quad Observations

Can we identify?
Using the bid distributions

\[ G_{B_1|b_1}(X_1|x_1) = \Pr(B_1 \leq X_1|b_1 = x_1) = \Pr(y_1 \leq s^{-1}(X_1)|v_1 = s^{-1}(x_1)) = F_{y_1|v_1}(s^{-1}(X_1)|s^{-1}(x_1)) \]

\[ g_{B_1|b_1}(X_1|x_1) = \frac{f_{y_1|v_1}(s^{-1}(X_1)|s^{-1}(x_1))}{s'(s^{-1}(X_1))} \]

\[ \implies v_i = b_i + \frac{G_{B_1|b_1(b_i|b_i)}}{g_{B_1|b_1(b_i|b_i)}} \equiv \xi(b, G), i = 1, \ldots, I \quad (5) \]

The RHS of (5) contains observables only

\[ F(\cdot, \ldots, \cdot) \] is nonparametrically identified

Remark: APV is the most general framework identified from observed bids as any affiliated model and hence a CV model is observationally equivalent to some APV model

In AV model, bidder’s utility \( U(v_1, \ldots, v_I, c) \) and \( (v_1, \ldots, v_I, c) \sim F(v_1, \ldots, v_I, c) \)

In CV model, bidder’s utility \( U(v_1, \ldots, v_I, c) = c \) and the \( v_i \)s are independent conditionally upon \( c \)

Laffont and Vuong (1996, American Economic Review)
• Model Restrictions/Rationalization

A joint bid distribution $G(\cdot, \ldots, \cdot)$ can be rationalized by a symmetric APV model if and only if

(i) $G(\cdot, \ldots, \cdot)$ is exchangeable and affiliated

(ii) $\xi(b, G)$ is increasing in $b$

• Estimation

Observations $(b_{i\ell}, i = 1, \ldots, I_\ell, \ell = 1, \ldots, L)$

Two-Step Procedure: Given $I, L_I$ number of auctions with $I$ bidders

STEP 1: Estimate $G_{B_1, b_1}(\cdot, \cdot)$ and $g_{B_1, b_1}(\cdot, \cdot)$ using $B_{i\ell} = \max_{j \neq i} b_{j\ell}$

\[
\hat{G}_{B_1, b_1}(B, b) = \frac{1}{L_I h_G} \sum_{\ell=1}^{L_I} \frac{1}{I} \sum_{i=1}^{I} \mathbb{I}(B_{i\ell} \leq B) K_G \left( \frac{b - b_{i\ell}}{h_G} \right)
\]

\[
\hat{g}_{B_1, b_1}(B, b) = \frac{1}{L_I h_g^2} \sum_{\ell=1}^{L_I} \frac{1}{I} \sum_{i=1}^{I} K_g \left( \frac{B - B_{i\ell}}{h_G} \right) K_g \left( \frac{b - b_{i\ell}}{h_G} \right)
\]

\[
\Rightarrow \hat{V}_{i\ell} = b_{i\ell} + \frac{1}{I - 1} \frac{\hat{G}_{B_1, b_1}(b_{i\ell}, b_{i\ell})}{\hat{g}_{B_1, b_1}(b_{i\ell}, b_{i\ell})}, i = 1, \ldots, I, \ell = 1, \ldots, L_I
\]
Step 2: \( \hat{f}(v_1, \ldots, v_I) = \frac{1}{L h_f} \sum_{\ell=1}^{L} K_f \left( \frac{v_1 - \hat{V}_{1\ell}}{h_f} \right) \ldots K_f \left( \frac{v_I - \hat{V}_{I\ell}}{h_f} \right) \)

Remark: Impose exchangeability by averaging on \( I! \) permutations, i.e. for \( I = 2 \)
\( \tilde{f}(v_1, v_2) = [\hat{f}(v_1, v_2) + \hat{f}(v_2, v_1)]/2 \)

Consistency Rate: If \( R > I - 2 \), \( \hat{f}(\cdot, \ldots, \cdot) \) converges at rate \( L^{\frac{R(R+I-2)}{(R+I+1)(2R+2I-2)}} \), quite slow but does not require to specify a priori the dependence among values!

\( I = 2, R = 2, L = 200 \rightarrow \text{rate } L^{4/30} = 2.03 \)

If the private values were observed, the rate would be \( L^{2/6} = 5.85! \)

How to choose the bandwidths?
\( \rightarrow h_G \propto L^{-\frac{1}{2R+2I-3}}, h_g \propto L^{-\frac{1}{2R+2I-2}}, h_f \propto L^{-\frac{R+I-2}{(2R+2I-2)(R+I+1)}} \)

- Monte-Carlo Experiments

\( v_i = \gamma + u_i, \gamma \sim U_{[a,b]}, u_i \sim U_{[-\epsilon, \epsilon]}, a = 0.25, b = 0.75, \epsilon = 0.25 \)

\( \rightarrow [v, \bar{v}] = [0, 1] \) and \( \text{corr}(v_i, v_j) = 0.5 \)

\( L = 100, I = 2, I = 3, \) draw private values, compute equilibrium bids and apply the estimation procedure

1,000 replications
(i) marginal density within confidence interval for $I = 2$

(ii) narrower confidence interval for $I = 3$

(iii) problem at .5 because of smoothness issue

(iv) tends to underestimate the average IR

(v) IR smaller for $I = 3$ than for $I = 2$

(vi) When estimating the model under independence

→ overestimate $v$ when large

underestimate $v$ when small

→ true density does not lie in the confidence interval

(vii) trimming issue, Hickman and Hubbard (2013, WP)
III. UNOBSERVED HETEROGENEITY, RISK AVersion, ASCENDING AUCTIONS

Unobserved Heterogeneity

• Motivation: A typical problem in many empirical studies
  → The analyst does not observe all the characteristics or observes them but imperfectly (measurement error)
  → Construction projects involve specifications that are too many and/or impossible to quantify
  → Goods have also a large number of characteristics hard to quantify

• Basic Idea: Multiplicative decomposition of private values into 2 components
  (1) Common component: Heterogeneity observed by all bidders included unobserved to the analyst
  (2) Individual component: Private information to bidder

Remarks:
  – if asymmetry, the distribution of (2) varies across types
  – APV also introduces dependence of bids, can we distinguish models?
  – bids ∼ indicators in a measurement error model
Krasnokutskaya (2011, Review of Economic Studies)
See also Li, Perrigne and Vuong (2000, Journal of Econometrics), similar idea but in the context of CV

- Model

Assumptions:
(i) I bidders common knowledge
(ii) \( v_i = y \epsilon_i, \ i = 1, \ldots, I \)
\( \rightarrow y \) common and known to all bidders
\( \rightarrow \epsilon_i \) private information to bidder \( i \), iid
(iii) \( (y, \epsilon) \sim F(\cdot, \cdot) \) on \([y, \bar{y}] \times [\underline{\epsilon}, \bar{\epsilon}]\)
(iv) \( y \) and \( \epsilon \) are independent, i.e. \( F(y, \epsilon) = F_y(y) \times F_\epsilon(\epsilon) \) with \( f_y(\cdot) > 0, f_\epsilon(\cdot) > 0 \)
(v) \( E(\epsilon) = 1 \) normalization
(vi) IPV framework, \([F_y(\cdot), F_\epsilon(\cdot)]\) common knowledge

Remarks:
- in the APV model bidders could not distinguish \( y \) from \( \epsilon \)
- if asymmetry \( \epsilon_0 \sim F_{\epsilon_0}(\cdot), \epsilon_1 \sim F_{\epsilon_1}(\cdot) \)
- (iv) crucial for identification


Bidder’s optimization problem: $b_i = s(v_i)$, $i = 1 \ldots, I$

$$\max_{b_i} (b_i - y \epsilon_i) \Pr(b_i \geq b_j, j \neq i | \epsilon_i, y)$$

Remark: Equilibrium exists and is unique, $s(\cdot)$ strictly increasing

Solution: $\alpha(\cdot)$ equilibrium with $y = 1$, i.e. $b_i = s(v_i) = s(y \epsilon_i) = y \alpha(\epsilon_i) = \alpha(\epsilon_i)$, $i = 1, \ldots, I$

With $y = 1$, $\Pr(b_i \geq b_j, j \neq i) = F_{\epsilon}(\alpha^{-1}(b_i))^{I-1}$

$$\implies 1 = (\epsilon_i - b_i)(I - 1) \frac{f_{\epsilon}(\alpha^{-1}(b_i))}{F_{\epsilon}(\alpha^{-1}(b_i))} \frac{1}{\alpha'(\alpha^{-1}(b_i))}$$ with $\alpha^{-1}(b_i) = \epsilon_i$

with appropriate boundary conditions.

Equilibrium: Any value of $y$ will lead to the same differential equation, i.e. $b_i = s(v_i) = y \alpha(\epsilon_i) = y \alpha_i$, $i = 1, \ldots, I$, with $\alpha(\cdot)$ strictly increasing
• Identification

\[ F_y(\cdot), F_\epsilon(\cdot), I \quad G(\cdot, \ldots, \cdot), I \]

Model Primitives \quad Observations

Can we identify?

Remark: (i) Use \( G(\cdot, \ldots, \cdot) \) instead of \( G(\cdot) \) because \( b_i = y\alpha_i \)

(ii) The bids are ‘indicators’ in a measurement error model

Proceed in two steps: (1) Identification of \( f_y(\cdot) \) and \( f_\alpha(\cdot) \)

(2) Identification of \( F_\epsilon(\cdot) \) through \( \alpha_i = \alpha(\epsilon_i) \)
Step 1 Using Kotlarski (1966), $f_y(\cdot)$ and $f_\alpha(\cdot)$ are nonparametrically identified

$$I = 2, \quad \begin{cases} 
\log B_1 = \log Y + \log A_1 \\
\log B_2 = \log Y + \log A_2 
\end{cases}$$

Idea: The characteristic function of the sum of two independent variables is equal to the product of the characteristic functions of these two variables

* Independence of $Y$ and $\epsilon$ crucial assumption!

$\Psi(\cdot, \cdot)$ characteristic function of $(\log B_1, \log B_2)$

$\phi_{\log y}$ characteristic function of $\log Y$

$\phi_{\log \alpha}$ characteristic function of $\log A$
\[ \phi_{\log y}(t) = \exp \left( \int_0^t \frac{\partial \Psi(0, u_2)}{\partial u_1} du_2 - itE[\log A] \right) \]

\[ \phi_{\log \alpha}(t) = \frac{\Psi(t, 0)}{\phi_{\log y}(t)} = \frac{\Psi(0, t)}{\phi_{\log y}(t)} \]

\[ \Rightarrow f_y(\cdot) \text{ and } f_\alpha(\cdot) \text{ are nonparametrically identified} \]

Remarks:
- Assume \( E[\log A] = 0 \) and then make adjustments on \( Y, A_1, A_2 \) to achieve \( E(\epsilon) = 1 \)
- We cannot identify the realizations of \( Y, A_1 \) and \( A_2 \) because of 2 equations for 3 unknown!

STEP 2: \( \alpha_i = \alpha(\epsilon_i) \), since \( \alpha(\cdot) \) is increasing and using the distribution of \( \alpha \) the FOC gives

\[ \epsilon_i = \alpha_i + \frac{1}{I-1} F_\alpha(\alpha_i) \equiv \xi(\alpha_i), \ i = 1, \ldots, I \]

\[ \Rightarrow F_\epsilon(\cdot) \text{ is nonparametrically identified} \]
Model Restrictions

Not all sufficient and necessary conditions are known, some testable implications

Assuming $I = 3$, then $\log(B_1/B_3) = \log A_1 - \log A_3$ and $\log(B_2/B_3) = \log A_2 - \log A_3$, i.e. $A_3$ similar as $Y$

$\rightarrow$ characteristic functions $\lambda_{\log A_3}(t)$, $\lambda_{\log A_1}(t)$

(1) $\lambda_{\log A_3}(t) = \lambda_{\log A_1}(t)$

(2) $\phi_{\log A} = \lambda_{\log A_i}(t)$, $i = 1, 3$

(3) $B_1/B_2$ and $B_3/B_4$ are independent with $I = 4$

(4) $\phi_{\log y}(t)$ should be the same whether we use $(1, 2)$ or $(1, 3)$ with $I = 3$

(5) $\xi(\cdot)$ strictly increasing in $\alpha_i$

How to distinguish with the APV model?

Intuition: The set of affiliated distributions is larger than the set of conditionally independent distributions

$\rightarrow$ Restrictions

Assume $I = 4$, then $B_1/B_2 \perp B_3/B_4$, i.e. independence of bid ratios, not true in APV
• Estimation

Observables \((B_{i\ell}, i = 1, \ldots, I, \ell = 1, \ldots, L)\) (no \(X_{\ell}, \) could be introduced in \(F_{\ell}(\cdot)\)) \(\rightarrow \log B_{i\ell}\)

A Two-Step Procedure:

**Step 1:** Fix \(I, L_I\)

\[
\hat{\Psi}(t_1, t_2) = \frac{1}{I(I - 1)} \sum_{1 \leq i \neq j \leq I} \frac{1}{L_I} \sum_{\ell = 1}^{L_I} \exp(it_1 \log B_{i\ell} + it_2 \log B_{j\ell}) \quad \longrightarrow \quad \frac{\partial \hat{\Psi}(t_1, t_2)}{\partial t_1}
\]

\[
\hat{\phi}_{\log y}(t) = \exp \left( \int_0^t \frac{\partial \hat{\Psi}(0, u_2)/\partial u_1}{\hat{\Psi}(0, u_2)} du_2 \right)
\]

\[
\hat{\phi}_{\log \alpha}(t) = \frac{\hat{\Psi}(t, 0)}{\hat{\phi}_{\log y}(t)} = \frac{\hat{\Psi}(0, t)}{\hat{\phi}_{\log y}(t)}
\]
\[ \tilde{f}_{\log \alpha}(u_1) = \frac{1}{2\pi} \int_T^T d(t) \exp(-it u_1) \hat{\phi}_{\log \alpha}(t) dt \quad \rightarrow \quad \hat{f}(\alpha) = \frac{\tilde{f}_{\log \alpha}(\log \alpha)}{\alpha} \]

\[ \tilde{f}_{\log y}(u_2) = \frac{1}{2\pi} \int_T^T d(t) \exp(-it u_2) \hat{\phi}_{\log y}(t) dt \quad \rightarrow \quad \hat{f}(y) = \frac{\tilde{f}_{\log y}(\log y)}{y} \]

Remarks:

- the log $B$ can be adjusted to a smaller interval to avoid oscillations

- damping factor $d(t) = 1 - |t|/T$ if $|t| \leq T$ to smooth the tails of estimated densities, typical problem with truncated inverse Fourier transformations

- smoothing parameter $T$, could be chosen from comparisons of mean and variance values of log $Y$ and log $A$ from bids with those obtained from estimated densities
STEP 2: Using \( \hat{F}_\alpha(\alpha) = \int_0^\alpha \hat{f}_\alpha(u)du \),

\[
\hat{\xi}(\alpha) = \alpha + \frac{1}{I-1} \frac{\hat{F}_\alpha(\alpha)}{\hat{f}_\alpha(\alpha)} \rightarrow \alpha = \hat{\xi}^{-1}(\epsilon) \rightarrow \hat{F}_\epsilon(\epsilon) = \hat{F}_\alpha(\hat{\xi}^{-1}(\epsilon))
\]

Alternatively, draw \( \alpha \) as from \( \hat{G}_\alpha(\cdot) \) and estimate \( f_\epsilon(\cdot) \) using \( \epsilon = \hat{\xi}(\alpha) \)

\[
\rightarrow \hat{f}_\epsilon(v) = \int_\frac{v}{y} \frac{1}{y} \hat{f}_\epsilon \left( \frac{v}{y} \right) \hat{f}_y(y)dy
\]

Remark: Some adjustments need to be made to attain normalization \( E(\epsilon) = 1 \)

APPLICATION TO MICHIGAN HIGHWAY PROCUREMENTS

• Data: Resurfacing of roads, 3947 contracts
  
  \( \rightarrow \) Unobserved heterogeneity: road condition (elevation, curvature), traffic, conditions of existing surface

• Empirical Evidence:
  
  – two types of bidders: regular/large, fringe
  
  – Independence of bid pairs rejected (IPV vs unobserved heterogeneity)
  
  – Independence of bid ratios not rejected (APV vs unobserved heterogeneity)
- bid regression on engineering estimate, duration, number of items, $R^2 = 0.17$
  → room for unobserved heterogeneity!

• Empirical Results: $300K \leq X \leq 580K$, completion 2–6, 2 fringe, 2 regular, $L = 226$

Comments:
- asymmetry in $f_\epsilon(\cdot)$, fringe bidders have higher mean but lower variance than regular ones
- $\epsilon$ counts for 31% of cost variation
- IR (measured as $(\alpha - \epsilon)/\epsilon$): 8.4% for regular, 6.1% for fringe
- 3.6% to 6.2% of inefficient auctions, leading to a 2% increase in procurement costs
– APV and IPV lead to lower cost estimates and larger IR.
– IPV and APV lead to larger variance in addition to different modes and mean.

This method has been successfully applied in a number of subsequent papers for various models (entry, collusion, etc).
Risk Aversion

- Motivation: Important component of individual behavior
  - In experiments, large empirical evidence of risk aversion
  - In auction data, evidence of risk diversification in timber auctions, success of buy-it-now option in eBay

Guerre, Perrigne and Vuong (2009, *Econometrica*)
Campos, Guerre, Perrigne and Vuong (2011, *Review of Economic Studies*)

- Model

Assumptions:

  (i) IPV, \( v_i \sim F(\cdot|I) \) on \([v(I), \bar{v}(I)]\)

  (ii) bidder’s utility \( U(\cdot) \), \( U(0) = 0 \), \( U'(\cdot) > 0 \), \( U''(\cdot) \leq 0 \) (risk aversion), \( U(1) = 1 \) (normalization)

Remarks:

- \( F(\cdot|I) \) conditioning upon \( I \) important, source of the identifying restriction
- risk neutrality \( U(\cdot) = \cdot \)
Bidder’s optimization problem: Equilibrium strategy \( b_i = s(v_i, F, U, I) \)

\[
\max_{b_i} U(v_i - b_i) \Pr(b_i \geq b_j, j \neq i) = U(v_i - b_i) F^{I-1}(s^{-1}(b_i))
\]

\[\rightarrow\] Differential Equation (FOC)

\[1 = (I - 1) \frac{f(v_i|I)}{F(v_i|I)} \frac{1}{s'(v_i)} \lambda(v_i - b_i)\]

where \( \lambda(\cdot) = U(\cdot)/U'(\cdot) \) with boundary condition \( s(\underline{v}(I)) = \underline{v}(I) \) since \( U(0) = 0 \).

Equilibrium Properties:

- \( s(\cdot, F, U, I) \) exists and is strictly increasing
- bids tend to be larger than under risk neutrality

Remark: \( \lambda'(\cdot) = 1 - \frac{U(\cdot) U''(\cdot)}{U'^2(\cdot)} \geq 1 \) since \( U''(\cdot) \leq 0 \)
• Identification

Using bid distribution and density, the FOC becomes

\[ 1 = (I - 1) \frac{g(b_i|I)}{G(b_i|I)} \lambda(v_i - b_i) \]

\[ \implies v_i = b_i + \lambda^{-1} \left( \frac{1}{I - 1} \frac{G(b_i|I)}{g(b_i|I)} \right) = \xi(b_i, U, G, I) \text{ for } i = 1, \ldots, I \quad (1) \]

\[ U(\cdot), F(\cdot|I), I \quad G(\cdot|I), I \]

Model Primitives \hspace{1cm} Observations

Can we identify?

A priori no without additional restriction, data, \ldots
\([U, F]\) is not identified

Counterexample: Construct \([\tilde{U}, \tilde{F}]\) with \(\tilde{U}(\cdot) = \left[\frac{U(\cdot/\delta)}{U(1/\delta)}\right]^\delta, \delta \in (0, 1)\). Thus \(\tilde{\lambda}(\cdot) = \frac{\lambda(\cdot/\delta)}{\delta}\)

By choosing \(\tilde{F}(\cdot)\) distribution of \(\tilde{v} = (1 - \delta)b + \delta \xi(b) = \tilde{\xi}(b)\), \([U, F]\) and \([\tilde{U}, \tilde{F}]\) are observationally equivalent, i.e. they lead to the same bid distribution \(G(\cdot)\)

Intuition: The nonidentification arises from the ‘soft’ model restrictions

• Model Restrictions/Rationalization

Smoothness assumptions: – \(F(\cdot|I)\) is \(R + 1\), \(f(\cdot|I) > 0\)

– \(U(\cdot)\) is \(R + 2\), \(U(0) = 0, U(1) = 1, U'(\cdot) > 0, U''(\cdot) \leq 0\)

+ smoothness of \(\lambda(\cdot)\) at 0

\(\implies G(\cdot|I)\) is \(R + 1\), \(g(\cdot|I) > 0, g(\cdot|I)\) is \(R + 1\) up to smoothness of \(G(b|I)/g(b|I)\) at \(b(I)\)

There exists an IPV model with risk aversion that rationalizes \(G(\cdot, \ldots, \cdot)\) if and only if

(i) \(G(b_1, \ldots, b_I|I) = \prod_{i=1}^{I} G(b_i|I)\) Bid independence

(ii) \(\exists \lambda(\cdot) R + 1\) differentiable with \(\lambda(0) = 0, \lambda'(\cdot) \geq 1\) such that \(\xi'(\cdot) > 0\).
Additional result: Any bid distribution $G(\cdot|I)$ satisfying the above smoothness assumptions can be rationalized by a structure $[U,F]$ with risk aversion

Comments:
- restriction (ii) is redundant
- risk aversion can rationalize a large set of bid distributions

Question: How to achieve identification?

Three possibilities: (1) Exclusion restriction $F(v|I) = F(v)$

(2) Parsimonious parameterization of $[U,F]$

(3) Additional data from ascending auctions

**Identification under exogenous participation:** $F(v|I) = F(v)$

Intuition: Exploit variations of the bid quantiles in $I$ while the value quantiles remain invariant

Assumptions:
- $I_2 > I_1 \geq 2$
- $s(\cdot)$ increasing in participation (not crucial but easier for presentation)
- $s_j(\cdot,U,F), G_j(\cdot), j = 1,2$
Denoting $v(\alpha)$ $\alpha$-quantile of $F(\cdot)$, $b_j(\alpha)$ $\alpha$-quantile of $G_j(\cdot)$

Compatibility Conditions

$$v(\alpha) = b_2(\alpha) + \lambda^{-1}\left( \frac{1}{I_2 - 1} \frac{\alpha}{g_2(b_2(\alpha))} \right) = b_1(\alpha) + \lambda^{-1}\left( \frac{1}{I_1 - 1} \frac{\alpha}{g_1(b_1(\alpha))} \right) \forall \alpha \in [0, 1].$$

Remarks:

- $R_j(\alpha) = \alpha/[(I_j - 1)g_j(b_j(\alpha))] = \lambda[v(\alpha) - s_j(v(\alpha))]$ with range $[0, \tau_j]$

- Because $s_1(\cdot) < s_2(\cdot)$, $R_1(\cdot) > R_2(\cdot) \Rightarrow \tau_2 < \tau_1$

- $[0, \max_v v - s_1(v)]$ domain of $\lambda(\cdot)$

- Identifying $[\lambda, F]$ is equivalent to identify $[U, F]$ since $U(x) = \int_t^x [1/\lambda(t)]dt$
Consider $\alpha_0$
\[ v(\alpha_0) - b_1(\alpha_0) = \lambda^{-1}[R_1(\alpha_0)] \]
But $b_1(\alpha_0) + \lambda^{-1}[R_1(\alpha_0)] = b_2(\alpha) + \lambda^{-1}[R_2(\alpha_0)]$
\[ \Rightarrow \lambda^{-1}[R_1(\alpha_0)] = b_2(\alpha_0) - b_1(\alpha_0) + \lambda^{-1}[R_2(\alpha_0)] \]
\[ = \Delta b(\alpha_0) + \lambda^{-1}[R_2(\alpha_0)] \]
By continuity of $R_j(\cdot)$, $\exists \alpha_1 : R_1(\alpha_1) = R_2(\alpha_0)$
But $\lambda^{-1}[R_1(\alpha_1)] = \Delta b(\alpha_1) + \lambda^{-1}[R_2(\alpha_1)]$
\[ \Rightarrow \lambda^{-1}[R_1(\alpha_0)] = \Delta b(\alpha_0) + \Delta b(\alpha_1) + \lambda^{-1}[R_2(\alpha_1)] \]
\[ \vdotswith{2} \]
\[ \lambda^{-1}[R_1(\alpha_0)] = \sum_{t=1}^{T} \Delta b(\alpha_t) + \lambda^{-1}[R_2(\alpha_T)] \]
\[ \vdotswith{2} \]
\[ \lambda^{-1}[R_1(\alpha_0)] = \sum_{t=1}^{\infty} \Delta b(\alpha_t) \]
$\lambda^{-1}(\cdot)$ is identified!

The sequence $\{\alpha_t\}$ exists but not necessarily unique, but $\lambda^{-1}(u)$ is unique.
\[ v(\alpha_0) = b_1(\alpha_0) + \sum_{t=1}^{\infty} \Delta b(\alpha_t) \Rightarrow F(\cdot) \text{ is identified} \]
• Model Restrictions/Rationalization

(i) bid independence

(ii) $\xi(\cdot)$ strictly increasing

(iii) compatibility conditions satisfied for every $\alpha \in [0, 1]$

• Extension to Endogenous Participation: $I = I(W, Z, \epsilon)$

Assumptions:

- $v \perp Z|W, \epsilon \Rightarrow F(v|W, Z, \epsilon) = F(v|W, \epsilon)$
- $\epsilon = I - E[I|W, Z]$ because $\epsilon$ is unobserved
- $Z \sim$ instruments, in practice reserve price, seller’s reputation

• Estimation: Not studied yet, can rely on the quantiles sequence or compatibility conditions (with sieve estimator)

N.B. Results extend to reserve price, APV, asymmetry in preferences and/or values
SEMIParametric Identification

Natural to parametrize $U(\cdot)$:
- CRRA with $U(x) = x^{1-c}$, $0 \leq c \leq 1$
- CARA with $U(x) = (1 - \exp(-ax))/(1 - \exp(-a))$, $a > 0$

Any distribution $G(\cdot|I)$ satisfying some smoothness can be rationalized by some CRRA or CARA structure

Not so good result: We cannot distinguish a CARA from a CRRA structure!

Any CARA or CRRA structure is not identified.

Intuition: Construct $[U, F]$ and $[\tilde{U}, \tilde{F}]$ such that
- $U(x) = x^{1-c}$ and $\tilde{U}(x) = x^{1-\tilde{c}}$, $\tilde{c} > c$
- $\tilde{F}(\cdot)$ is the distribution of $\tilde{v} = b + \frac{1-\tilde{c}}{1-c} \frac{G(b)}{g(b)} = \frac{\tilde{c} - c}{1-c} b + \frac{1-\tilde{c}}{1-c} \xi(b)$
- $\tilde{\xi}'(\cdot) > 0$, thus $[\tilde{U}, \tilde{F}]$ and $[U, F]$ are observationally equivalent
- $\tilde{v}(\alpha) < v(\alpha) \rightarrow$ all the quantiles of $\tilde{F}(\cdot)$ ‘shrink’ to compensate for the increase in risk aversion

Solution: Pin down a quantile by a conditional quantile restriction

Remark: Parameterizing $F(\cdot)$ only does not help
• Semiparametric Identification

Assumptions:
(i) $U(\cdot) = U(\cdot; \theta), \theta \in \mathbb{R}^p$
(ii) $v(\alpha; X, I) = v(\alpha; X, I, \gamma), \gamma \in \mathbb{R}^q$
(iii) unique solution of the system of equations in $(\theta, \gamma)$

Intuition: Variation in $(X, I)$ will achieve identification of $\theta$ through conditional $\alpha$-quantile restriction

At some $\alpha$, for any $(x, I)$

$$g(b(\alpha; x, I)|x, I) = \frac{1}{I - 1} \frac{\alpha}{\lambda(v(\alpha; x, I, \gamma) - b(\alpha; x, I; \theta))}.$$ 

Under assumptions (i)–(iii), the semiparametric model is identified.

Remark: A quantile restriction is not as strong as parameterizing the full distribution. Moreover, the $\alpha$-quantile can a polynomial of high degree

• Semiparametric Estimation

To simplify, $\alpha = 1$ and $v(1, x, I, \gamma) = \overline{v}$, i.e. constant upper boundary
Observations: \((B_{i\ell}, X_\ell, i = 1, \ldots, I_\ell, \ell = 1, \ldots, L)\)

Multistep Procedure:

**STEP 1:** (i) Estimate \(\bar{b}(x, I)\) by partitioning \(x\) and find the minimal cylinder containing the observations

(\(ii\)) Estimate \(g(\bar{b}(x, I)|x, I)\) with a one-sided kernel because of boundary effect

\[\to \hat{b}(X_\ell, I_\ell) \text{ and } \hat{g}(\hat{b}(X_\ell, I_\ell)|X_\ell, I_\ell)\]

**STEP 2:** Define the model

\[Y_{i\ell} = m(X_\ell, I_\ell; v, \theta) + e_{i\ell} + \epsilon_{i\ell}, i = 1, \ldots, I_\ell, L = 1, \ldots, L, \quad e_{i\ell}: \text{bias}, \quad \epsilon_{i\ell}: \text{error term}\]

where \(Y_{i\ell} = g(\bar{b}(X_\ell, I_\ell)|X_\ell, I_\ell)\), \(m(X_\ell, I_\ell; v, \theta) = \frac{1}{I_\ell - 1} \frac{1}{\lambda(v - \bar{b}(X_\ell, I_\ell); \theta)}\)

\[\to \text{Weighted NLLS}: \min_{v, \theta} \sum_{\ell=1}^{L} \sum_{i=1}^{I_\ell} \omega(X_\ell, I_\ell) [\hat{Y}_{i\ell} - \hat{m}(X_\ell, I_\ell; v, \theta)]^2\]

\[\to (\hat{\theta}, \hat{v})\]

**STEP 3:** Estimate nonparametrically the conditional bid distribution and density to obtain

\[\hat{V}_{i\ell} = B_{i\ell} + \lambda^{-1} \left( \frac{1}{I_\ell - 1} \hat{G}(B_{i\ell}|X_\ell, I_\ell), \hat{\theta} \right)\]

Estimate nonparametrically the density \(\to \hat{f}(\cdot|x, I)\)
\((\hat{\theta}, \hat{v})\) converges at the best (optimal) rate \(L^{\frac{R+1}{2(R+3)}} < \sqrt{L}\)

Why not \(\sqrt{L}\) as for most semiparametric estimators?

\[\rightarrow \text{Var}(\epsilon_{i\ell}) \text{ diverging} \]

\[\rightarrow \text{rate independent of } d \text{ though, no curse of dimensionality} \]

Asymptotic distribution for \(\hat{\theta}\): normality and variance allowing to test for risk aversion as \(\hat{\theta} = 1 - \hat{c}\)

N.B. Results extend to reserve price, APV, asymmetry in values and in preferences (no need of quantile restriction because of compatibility conditions)

**APPLICATION TO US TIMBER AUCTIONS**

- Data: 378 auctions, 1400 bids, focus on the 300 auctions with appraisal <300K.

Remarks:

- \(\alpha = 1\) and \(\alpha = .5\) provide similar results.
- the constant specification for the upper boundary provides a poor fit
- retained specification (see table): Quadratic with \(\hat{\theta} = 0.72\) (for \(\alpha = 1\), \(\hat{\theta} = 0.71\), \(\hat{c} = 0.28\) significant risk aversion
- private values would be overestimated with a risk neutral model
<table>
<thead>
<tr>
<th></th>
<th>Constant</th>
<th>Linear</th>
<th>Quadratic</th>
<th>Cubic</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRRA</td>
<td>$\hat{\theta} = 0.5866$</td>
<td>$\hat{\theta} = 0.6006$</td>
<td>$\hat{\theta} = 0.7223$</td>
<td>$\hat{\theta} = 0.5560$</td>
</tr>
<tr>
<td></td>
<td>(0.1150)</td>
<td>(0.1502)</td>
<td>(0.3586)</td>
<td>(0.2788)</td>
</tr>
<tr>
<td></td>
<td>$\hat{\gamma}_0 = 114.9999$</td>
<td>$\hat{\gamma}_0 = 25.0037$</td>
<td>$\hat{\gamma}_0 = 28.6067$</td>
<td>$\hat{\gamma}_0 = 22.1137$</td>
</tr>
<tr>
<td></td>
<td>(1.0379)</td>
<td>(6.1497)</td>
<td>(12.7786)</td>
<td>(9.7024)</td>
</tr>
<tr>
<td></td>
<td>$\hat{\gamma}_1 = 1.3339$</td>
<td>$\hat{\gamma}_1 = 1.4245$</td>
<td>$\hat{\gamma}_1 = 1.4978$</td>
<td>$\hat{\gamma}_1 = 1.4245$</td>
</tr>
<tr>
<td></td>
<td>(0.0658)</td>
<td>(0.0899)</td>
<td>(0.0016)</td>
<td>(0.0016)</td>
</tr>
<tr>
<td></td>
<td>$\hat{\gamma}_2 = -0.0016$</td>
<td>$\hat{\gamma}_2 = -0.0016$</td>
<td>$\hat{\gamma}_2 = -0.00346$</td>
<td>$\hat{\gamma}_2 = -0.00007$</td>
</tr>
<tr>
<td></td>
<td>(0.0004)</td>
<td>(0.0004)</td>
<td>(0.000006)</td>
<td>(0.000006)</td>
</tr>
<tr>
<td></td>
<td>$SSE/TSS = 0.8585$</td>
<td>$SSE/TSS = 0.3261$</td>
<td>$SSE/TSS = 0.3208$</td>
<td>$SSE/TSS = 0.3206$</td>
</tr>
<tr>
<td>CARA</td>
<td>$\hat{\theta} = 0.000005$</td>
<td>$\hat{\theta} = 0.000002$</td>
<td>$\hat{\theta} = 0.000003$</td>
<td>$\hat{\theta} = 0.000002$</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.007)</td>
<td>(0.018)</td>
<td>(0.0183)</td>
</tr>
<tr>
<td></td>
<td>$\hat{\gamma}_0 = 249.0001$</td>
<td>$\hat{\gamma}_0 = 40.2831$</td>
<td>$\hat{\gamma}_0 = 38.3498$</td>
<td>$\hat{\gamma}_0 = 38.0018$</td>
</tr>
<tr>
<td></td>
<td>(4.5300)</td>
<td>(3.7615)</td>
<td>(11.0000)</td>
<td>(1.8562)</td>
</tr>
<tr>
<td></td>
<td>$\hat{\gamma}_1 = 1.2898$</td>
<td>$\hat{\gamma}_1 = 1.4662$</td>
<td>$\hat{\gamma}_1 = 1.5103$</td>
<td>$\hat{\gamma}_1 = 1.5103$</td>
</tr>
<tr>
<td></td>
<td>(0.1230)</td>
<td>(0.2020)</td>
<td>(0.3131)</td>
<td>(0.0037)</td>
</tr>
<tr>
<td></td>
<td>$\hat{\gamma}_2 = -0.0015$</td>
<td>$\hat{\gamma}_2 = -0.0023$</td>
<td>$\hat{\gamma}_2 = -0.00346$</td>
<td>$\hat{\gamma}_2 = -0.00007$</td>
</tr>
<tr>
<td></td>
<td>(0.0009)</td>
<td>(0.0009)</td>
<td>(0.000006)</td>
<td>(0.000006)</td>
</tr>
<tr>
<td></td>
<td>$SSE/TSS = 0.7834$</td>
<td>$SSE/TSS = 0.3233$</td>
<td>$SSE/TSS = 0.3211$</td>
<td>$SSE/TSS = 0.3211$</td>
</tr>
</tbody>
</table>
Identification with Additional Data

Basic Idea: Combining ascending and first-price sealed-bid auction data

In ascending auctions, winning bid \( \sim \) second-highest value

\[ \rightarrow F(\cdot|x, I) \] is identified (see next slide on ascending auctions)

\[ \rightarrow \hat{F}(\cdot|x, I) \] from observables \((b_{\ell}^{w}, X_{\ell}, I_{\ell}, \ell = 1, \ldots, L_{A})\)

From sealed-bid auction

\[ v(\alpha, x, I) - b(\alpha, x, I) = \lambda^{-1}\left(\frac{1}{I - 1} g(b(\alpha, x, I)|x, I)\right) \]

Since \( v(\alpha, x, i) = F^{-1}(\alpha|x, I), \lambda^{-1}(\cdot) \) is identified

Using nonparametric estimators, \( \hat{\lambda}(\cdot) \) converges at rate \( L_{F}^{2R+4+3} \)

Application to Timber Auction Data

Potential selection of bidders and tract characteristics into auction mechanisms

A CRRA approximation from \( \hat{\lambda}(\cdot) \) gives \( \hat{\theta} = 0.64 \) (CRRA) and \( \hat{a} = 0.00004 \) (CARA)

\( \hat{U}(\cdot) \) has irregularities partly captured by CRRA, parametric specification would tend to overestimate \( U(\cdot) \).
Ascending Auctions

Motivation: Widely used in practice for collectible, art, eBay, cars, . . .

Practice of ascending auctions quite different from theory

Button Model: When the value raises, bidders release the button, the last bidder pays the value at which the bidder before him releases the button, IPV

→ Winning bid ∼ Second-highest value among I bidders
→ Playing his value is a dominant strategy (still valid with risk aversion, asymmetry)

Practice: Bidders can bid several times, minimum increment, . . ., eBay has even more complex rules

Question: How can we rationalize the observations (winning and losing bids)?

First Approach: Winning bid ∼ second-order statistics among I bidders, I observed

Order Statistics: 

\[
F_{n:I}(v) = \frac{I!}{(n-1)!} \int_0^{F(v)} t^{n-1} (1-t)^{I-n} dt
\]

\[n = I - 1 \rightarrow F_{I-1:I}(v) = IF^{I-1}(t) - (I-1)F^I(t)\]

→ \(F(\cdot)\) is nonparametrically identified (this argument does not work with affiliation)
Estimation: Estimate nonparametrically the second-order statistics from winning bids and solve for $F(\cdot)$

\[ \rightarrow \hat{F}(v|x, I) \]

SECOND APPROACH: Consider an ‘incomplete’ model


- Basic Idea: Use some assumptions of rational behavior rationalizing observations

  \[ \rightarrow \text{Bounds on } F(\cdot) \]

- Assumptions:
  - IPV model, $I$ potential bidders, $F(v)$
  - $\Delta$ bid increment

Two assumptions of rational behavior

(i) Bidders do not bid more than they are willing to pay

(ii) Bidders do not allow an opponent to win at a price they can beat

Ordered bids $(B_{1:I}, \ldots, B_{I:I}), b_{i:I} \sim G^{i:I}(\cdot)$
Upper Bound: (i) implies \( b_i \leq v_i \) or \( G(v) \geq F(v) \) for all \( v \)

In addition, \( b_{i:I} \leq v_{i:I} \) for all \( I \Rightarrow F_{i:I}(v) \leq G_{i:I}(v) \)

Since \( F(v) = \phi(F_{i:I}(v), i, I) \) (see above) with \( \phi(\cdot) \) monotonic
\[
F(v) \leq F_U(v) = \min_{i,I} \phi(G_{i:I}(v), i, I)
\]

Lower Bound: (ii) implies that \( v_{I-1:I} \) is less than \( b_{I:I} + \Delta \) distributed as \( G_{\Delta}^{I,I}(\cdot) \) or \( F_{I-1:I}(v) \geq G_{\Delta}^{I,I}(v) \)
\[
F(v) \geq F_L(v) = \max_{I} \phi(G_{\Delta}^{I,I}(v), I - 1, I)
\]

Remarks:

– In the button auction model \( \Delta = 0 \) and \( b_{I-1:I} = v_{I-1:I} = b_{I:I} \), thus lower and upper bounds collapse

– \( F(\cdot) \) is set identified

• Estimation: Use nonparametric estimators for bid distributions, to avoid \( \hat{F}_U(\cdot) \) and \( \hat{F}_L(\cdot) \) crossing in finite samples, a correction can be made

N.B. No extensions so far, seems difficult to apply to affiliated values, asymmetry, . . .
• Monte Carlo Experiments
  → Overall very good performance except for low values (problem with order statistics distribution)
  → Larger $\Delta$ tend to lead to wider bounds
  → Comparison with using losing bids only, poor performance
  → Comparison with using winning bids only good results if $\Delta$ not too large

APPLICATION TO US TIMBER AUCTIONS

$\Delta = 5$ cents per thousand board feet (MBF), bid show jump bidding
$X$: year, species concentration, harvesting costs, appraisal value, location, six monts inventory

Bootstrapped confidence intervals
$\hat{f}(\cdot) \sim \text{lognormal shape}$
Rejection of classic button model
Bonds derived for optimal reserve price
IV. NONLINEAR PRICING, BUNDLING, DIFFERENTIATED PRODUCTS

Nonlinear Pricing: The Basics

Perrigne and Vuong (2013, working paper)

- Assumptions:
  
  (i) Consumer taste/willingness to pay (adverse selection) $\theta \sim F(\cdot)$ on $\Theta \equiv [\underline{\theta}, \overline{\theta}]$, $f(\cdot) > 0$, iid
  
  (ii) Consumer utility $U(q, \theta)$ on $\mathbb{R}^+ \times \Theta$ continuously differentiable
  
  (iii) $U(0, \theta) = 0$ (normalization), $U_q(q, \theta) > 0$, $U_{qq}(q, \theta) \leq 0$, $U_\theta(q, \theta) > 0$, $U_{\theta\theta}(q, \theta) \leq 0$, $U_{q\theta}(q, \theta) > 0$
  
  (iv) $U_q(q, \theta) - \frac{1 - F(\theta)}{f(\theta)} U_{q\theta}(q, \theta)$ nondecreasing in $\theta$
  
  (v) Firm cost $C(q)$, $C_q(\cdot) > 0$
  
  (vi) $\theta$: private information, the firm knows $[U, F]$
  
  (vii) Tariff $T(q)$, $T(0) = 0$, $T_q(\cdot) > 0$
  
  (viii) Single good, no exclusion of consumers
Model

Consumer optimization and participation

\[ \max_q U(q, \theta) - T(q) \]

\[ \quad \rightarrow U_q(q, \theta) = T_q(q) \quad \text{(IC)} \]

\[ \quad \rightarrow U(q, \theta) - T(q) \geq U(0, \theta) - T(0) = 0 \quad \text{(IR)} \]

Remark:

- (IC) is equivalent to set \( U(q(\theta), \theta) - T(q(\theta)) \geq U(q(\tilde{\theta}), \theta) - T(q(\tilde{\theta})), \forall \theta, \tilde{\theta} \in \Theta \)
- no countervailing incentives

Firm optimization problem: Choose \( T(\cdot) \) and \( q(\cdot) \)

\[ \max_{T(\cdot),q(\cdot)} \Pi = \int_\Theta T(q(\theta)) f(\theta) d\theta - \int_\Theta C(q(\theta)) f(\theta) d\theta \quad \text{s.t.} \quad \text{(IC) \& (IR)} \]

Remarks:

- The firm faces one consumer. One could also consider the firm facing a population of consumers of size one with the cost equal to \( C[\int_\theta q(\theta) f(\theta) d\theta] \equiv C(Q) \)
- (IR) \( \rightarrow U(q(\theta), \theta) = T(q(\theta)) \), i.e. no rent for the \( \theta \) consumer. assuming \( q_\theta(\cdot) > 0 \)
Assuming $q_\theta(\cdot) > 0$, define $\tilde{U}(\theta) = U(q(\theta), \theta) - T(q(\theta))$

\[ \rightarrow \tilde{U}_\theta(\theta) = [U_q(q(\theta), \theta) - T_q(q(\theta))]q_\theta(\theta) + U_\theta(q(\theta), \theta) = U_\theta(q(\theta), \theta) \text{ and } \tilde{U}(\theta) = 0 \]

$q(\cdot)$: Control function, $\tilde{U}(\cdot)$: State Variable, $\mu(\cdot)$: Multiplier associated to (IC) constraint

Hamiltonian: \[ H = \{ U(q(\theta), \theta) - \tilde{U}(\theta) - C(q(\theta)) \} f(\theta) + \mu(\theta) U_\theta(q(\theta), \theta) \]

\[ \rightarrow H_q = \{ U_q(q(\theta), \theta) - C_q(q(\theta)) \} f(\theta) + \mu(\theta) U_{\theta q}(q(\theta), \theta) = 0 \]

\[ \rightarrow -H_{\tilde{U}} = f(\theta) = \mu_\theta(\theta) \]

Using tranversality condition $\mu(\theta) = 0$, $\mu(\theta) = -[1 - F(\theta)]$, the FOC are

\[ U_q(q(\theta), \theta) = C_q(q(\theta)) + \frac{1 - F(\theta)}{f(\theta)} U_{\theta q}(q(\theta), \theta) \quad (1) \]

\[ T_q(q(\theta)) = U_q(q(\theta), \theta) \quad (2) \]

with boundary condition $T(q(\theta)) = U(q(\theta), \theta)$
Remarks:

- $\theta$ consumer gets no rent
- (1) characterizes the optimal quantity schedule $q(\cdot)$ on $[q(\theta), q(\bar{\theta})] \equiv [\underline{q}, \bar{q}]$
- $\bar{\theta}$ consumer gets the first-best
- $\frac{1-F(\theta)}{f(\theta)} U_{\theta q}(q, \theta)$ distortion due to incomplete information
- once $q(\cdot)$ is known, (2) characterizes the optimal price schedule $T(\cdot)$ on $[T(q(\theta)), T(q(\bar{\theta}))] \equiv [\underline{t}, \bar{t}]$
- under additional assumptions, $q_{\theta}(\cdot) > 0$ and $T_{qq}(\cdot) < 0$ (concave tariff)
  $\rightarrow$ one-to-one mapping between $q$ and $\theta$
- second-degree price discrimination, i.e. tariff does not depend on consumers characteristics $X$

• Identification

Assumptions:

- single market data
- $T(\cdot)$ is observed/known
The model is **not** identified without identifying assumptions.

Discussion: (2) appears in hedonic models where $U(\cdot, \cdot)$ and $F(\cdot)$ are not identified. Identification is achieved under separability of $U(\cdot, \cdot)$ in $\theta$ and some variations in $X$ independently of $\theta$.

$\star$ Impossible to find a variable independent of $\theta$ but correlated with $q$ since $q = q(\theta)$

How to achieve identification?

$\rightarrow$ Exploit (1) and (2), i.e. FOC of both consumer and firm.

$\rightarrow$ Identifying assumption: $U(q, \theta) = \theta U_0(q)$, $U_0(\cdot) > 0$, $U_{0q}(\cdot) > 0$, $U_{0qq}(\cdot) < 0$

$\rightarrow$ Exploit the one-to-one equilibrium mapping $q = q(\theta)$
Structure \([U_0(\cdot), F(\cdot), C(\cdot)]\)

FOC become with \(q = q(\theta)\) and for all \(q \in [q, \bar{q}]\)

\[\theta U_{0q}(q) = C_q(q) + \frac{1 - F(\theta)}{f(\theta)} U_{0q}(q)\]  

\[T_q(q) = \theta U_{0q}(q)\]  

Normalization: \([U, F, C]\) and \([\tilde{U}, \tilde{F}, \tilde{C}]\) with \(\tilde{U}_{0q}(\cdot) = U_{0q}(\cdot)/\alpha\), \(\tilde{\theta} = \alpha \theta\) and \(\tilde{F}(\cdot) = F(\cdot/\alpha), \alpha > 0\). These two structures are observationally equivalent since they lead to the same FOC

\[\bar{\theta} = 1\]

Remarks:

- (3) and (4) at \(\bar{\theta}\) and \(q(\bar{\theta})\) give \(T_q(\bar{q}) = C_q(\bar{q}) \equiv \gamma\)

- Since (3) involves \(C_q(q)\), one needs to assume constant marginal cost, i.e. \(C_q(q) = \gamma\)

- If considering \(C(\int_{\Theta} q(\theta)f(\theta)d\theta)\), i.e. the cost for the total amount produced, no need for further assumption
Purchase/quantity distribution and density: \( G^g(q) = F(\theta) \) and \( g^q(q) = \theta_q(q) f(\theta) \)

(3) and (4) \( \rightarrow T_q(q) = \gamma + \frac{1-G^g(q)}{g^q(q)} \theta_q(q) U_0q(q) \)

Differentiating (4) w.r.t. \( q \) \( \rightarrow \theta(q) U_0q(q) = T_{qq}(q) - \theta(q) U_{0qq}(q) \)

Combining both equations gives \( \theta(q) U_{0qq}(q) = T_{qq}(q) - \frac{g^q(q)}{1-G^g(q)} [T_q(q) - \gamma] \)

Dividing the RHS by \( T_q(q) \) and the LHS by \( \theta(q) U_0q(q) \) (= \( T_q(q) \) by (4)) gives

\[
\frac{U_{0qq}(q)}{U_0q(q)} = \frac{T_{qq}(q)}{T_q(q)} - \frac{g^q(q)}{1-G^g(q)} \left( 1 - \frac{\gamma}{T_q(q)} \right)
\]

Integrating from \( q \) to \( \bar{q} \), taking the exponential and after some computations

\[
U_{0q}(q) = \frac{T_q(q)}{\xi(q)}, \quad \theta(q) = \xi(q) \quad \text{(5)}
\]

\[
\xi(q) = [1 - G^q(q)]^{\frac{\gamma}{1-\gamma}} \exp \left\{ -\gamma \int_q^{\bar{q}} \frac{T_{qq}(x)}{T_q^2(x)} \log[1 - G^q(x)] dx \right\} \quad \text{(6)}
\]

The RHS contain observables \( G^q(\cdot), T(\cdot) \) and \( \gamma \)

\( \rightarrow U_{0q}(\cdot) \) and \( F(\cdot) \) are nonparametrically identified on \( [q, \bar{q}] \) and \( [\theta, \bar{\theta}] = [\theta, 1] \)

\( \rightarrow C(\cdot) \) is identified by \( \gamma \) under constant marginal cost

7
Remarks:

- $\xi(q) = 1, \xi(q) = \theta$

- $U_0(q)$ identified by $U_0(q) = U_0(q) + \int_q^q U_0(x)dx$, where $U_0(q) = T(q)/\theta$ by boundary condition

• Model Restrictions/Rationalization

Not studied yet

Remark:

- preliminary step toward a test of asymmetric information as restrictions should be also derived for models (i) under complete information (first-best) and (ii) with cost-scale economies

• Estimation

Observations: $q_i, i = 1, \ldots, N$, $T(\cdot)$
A Multi-Step Procedure:

STEP 1: Estimate nonparametrically $\gamma, G^q(\cdot)$

$\rightarrow \hat{\gamma}, \hat{G}^q(\cdot)$

$\rightarrow \hat{\xi}(\cdot), \hat{U}_{0q}(\cdot), \hat{\theta}(\cdot)$

STEP 2: Estimate nonparametrically $f(\cdot)$

$\rightarrow \hat{f}(\cdot)$

STEP 1: $\hat{\gamma} = T_q(q_{\text{max}})$ with $q_{\text{max}} = \max_i q_i$, $\hat{G}^q(q) = \frac{1}{N} \sum_{i=1}^{N} I(q_i \leq q)$

$\rightarrow \hat{\xi}(q) = [1 - \hat{G}^q(q)]^{\frac{\hat{\gamma}}{T_q(q)}} \exp \left\{ \hat{\gamma} \left( \frac{1}{T_q(q^{j+1})} - \frac{1}{T_q(q^j)} \right) \log[1 - \hat{G}^q(q^j)] + \hat{\gamma} \sum_{t=j+1}^{J} \left( \frac{1}{T_q(q^{t+1})} - \frac{1}{T_q(q^t)} \right) \log[1 - \hat{G}^q(q^t)] \right\}$

for $q \in [q^j, q^{j+1}), j = 0, \ldots, J - 1$ with $q^0 = q$ and $q^J = q_{\text{max}}$

Remarks:

- $\hat{G}^q(\cdot)$ is a step function with steps at $q^1 < \ldots < q^J$

- $\hat{\xi}(q)$ for $q \in [q, q_{\text{max}})$
\[
\hat{U}_0(q) = \frac{T_q(q)}{\xi(q)}, \quad \hat{\theta}(q) = \hat{\xi}(q)
\]

**Step 2:** Estimate \(\hat{\theta}_i = \hat{\theta}(q_i), i, \ldots, N\)

\[
\hat{f}(\theta) = \frac{1}{Nh_f} \sum_{i=1}^{N} K \left( \frac{\theta - \hat{\theta}_i}{h_f} \right)
\]

- Asymptotic Properties (using empirical processes and normalization \(\bar{\theta} = 1\))

\(\hat{U}_0(q)\) converges at rate \(\sqrt{N}\) with estimated variance equal to \(\hat{U}_0^2(q)\hat{\omega}(q)\)

with \(\hat{\omega}(q) = \frac{1}{N} \sum_{i=1}^{N} 1_{\{q_i \leq q\}} \left( \frac{\hat{\gamma}}{T_q(q_i)} - 1 \right)^2 \frac{1}{(1 - G_q(q_i))^2}\)

Remarks:

- variance can be used for pointwise confidence intervals
- similar result for \(\hat{\theta}(\cdot)\)
- variance increasing in \(q\)
- \(\hat{f}(\cdot)\) converges at rate \(N^{\frac{R}{2R+2}}\) with \(h_f \propto N^{-\frac{1}{2R+2}}\)
- this rate is faster than that of GPV for auctions but still slower than \(N^{\frac{R}{2R+1}}\) for observed types
Exclusion of Consumers

- Motivation: Not all consumers purchase the good because of low taste or willingness to pay for it

Outside option

Examples: cellular phone services, advertising (standard listing)

- Assumption:
  - threshold type \( \theta^* \), if \( \theta \geq \theta^* \), consumers buy \( q(\cdot) > 0 \), if \( \theta < \theta^* \), consumers buy nothing or \( q = 0 \)

- Model: It is optimal for the firm to exclude consumers with low \( \theta \)

Consumer: \( U_q(q, \theta) = T_q(q) \) for \( \theta \in [\theta^*, \overline{\theta}] \) \( \rightarrow \) (IC)

Firm: \( \max_{q(\cdot), T(\cdot), \theta^*} \Pi = \int_{\theta^*}^{\overline{\theta}} T(q(\theta)) f(\theta) d\theta - \int_{\theta^*}^{\overline{\theta}} C(q(\theta)) f(\theta) d\theta \) s.t. (IC) & (IR)

Remarks:
  - boundary condition \( U(q(\theta^*), \theta^*) = T(q(\theta^*)) \), no rent for the \( \theta^* \) consumer
  - \( \theta^* \) endogenous as well as minimal quantity \( \underline{q} = q(\theta^*) \)
FOC (1) and (2) with boundary condition as well as exclusion condition

\[ \theta^* \text{ satisfying } U(q(\theta^*), \theta^*) - C(q(\theta^*)) - \frac{1 - F(\theta^*)}{f(\theta^*)} U_\theta(q(\theta^*), \theta^*) = 0 \]

- Identification

Truncated distribution \( G^{*q}(q) = \frac{F(\theta) - F(\theta^*)}{1 - F(\theta^*)} \equiv F^*(\theta) \)

Assuming \( U(q, \theta) = \theta U_0(q) \) and a constant marginal cost, (3) and (4) with \([1 - F(\theta)]/f(\theta) = [1 - F^*(\theta)]/f^*(\theta)\) identify \( U_{0q}(q) \) and \( f^*(\theta) \) on \([q, q]\) and \([\theta^*, \overline{\theta}]\)

Remarks:
- \( \xi(q) = \theta^* \)
- Exclusion condition can identify an additional cost parameter

Assuming \( C(q) = \kappa + \gamma q \) if \( q > 0 \) and \( C(q) = 0 \) if \( q = 0 \), \( \gamma = T_q(\overline{q}) \), using (3), (4) and the boundary condition \( T(q) = \theta^* U_0(q) \)

\[ \kappa = U_0(q) \left( \theta^* - \frac{1 - F(\theta^*)}{f(\theta^*)} \right) - \gamma q = \gamma \left( \frac{t}{T_q(q)} - q \right) \]

Remark:
- result extends to other cost functions, i.e. \( C(q) = \kappa(1 + q)^\gamma, \quad * \text{MC in FOC no longer a constant!} \)
• Estimation

Similar method applies, to estimate $\kappa$, $q_{\text{min}} = \min_i q_i$

APPLICATION TO YELLOW PAGE ADVERTISING

• Data: Phone book by Verizon, 6,823 advertisements/listings, tariff obtained from Yellow Page Association, standard listing free to all businesses

Comments:

– Concave tariff, reduction up to 66% of price per unit
– 4,671 businesses buy advertising $\rightarrow$ ‘exclusion’ of 31.5%
– color options $\rightarrow$ adjusted-quality quantities
– similar model but $q$ exogenous leading to boundary conditions $\lim_{\theta \downarrow \theta^*} q(\theta) = q$ and $\lim_{q \downarrow q} T(q) = 0$
– normalization $\theta^* = 1$

• Empirical Results

$\hat{\gamma} = 7.9$, i.e. an additional line costs at the margin $\$35$ and is charged $\$101$, an additional full-page page costs at the margin $\$11,919$ and is charged $\$18,513$
Comment: $\theta - \frac{1 - \hat{F}(\theta)}{f(\theta)}$ strictly increasing

- Counterfactuals

Assumptions: Constant marginal cost + functional forms of $U_0(\cdot)$ and $f(\cdot)$ on $[0, q]$ and $[\theta, \theta^*]$

Actual profit = $2,806,800$, actual businesses payoff+profit = $6,375,700$

Informational rent 37%
Linear Pricing $T(q) = pq$

$\rightarrow p = 15.8$

$\rightarrow$ less ‘exclusion’ at 18%

$\rightarrow$ an additional line costs $70 and $\bar{q} = 1,500$ (instead of $101 and 6,367$), penalizes large $\theta$ and favors low $\theta$

$\rightarrow$ profit − 21%, businesses payoffs + 10.4%, sum − 3.44%

Third-Degree Price Discrimination $T(q, X)$ with $U_0(q, X)$ and $F(\theta|X)$

$\rightarrow X$: B2C, B2B, Nonprofit organizations

$\rightarrow$ B2C firms have the largest types

$\rightarrow$ more ‘exclusion’, 58%, 55% and 41%

$\rightarrow$ tariff functions lower for the three groups, quantity schedules higher for the three groups

$\rightarrow$ profit + 4.8%, businesses payoff + 13.9%, sum + 9.9%
Observed Heterogeneity

How to introduce observed consumers characteristics $X$ in the model?

☆ In second-degree price discrimination $T(\cdot)$ independent of $X$

$\rightarrow$ 1st solution: Consider $U(q, \theta, X)$, $F(\theta|X)$ but impose $T(\cdot)$ independent of $X$

$\rightarrow$ 2nd solution: $\theta = r(X, \epsilon)$ with $\epsilon \sim F_{\epsilon|X}(\cdot|\cdot)$

• Assumptions and Identification

  (i) $\epsilon \perp X_1|X_2$ (to avoid independence of $X$ and $\epsilon$)

  (ii) $\exists x_1^0$ such that $r(x_1^0, x_2, \epsilon) = \epsilon$

  (iii) $r(x_1, x_2, \cdot)$ and $F_{\epsilon|X_1,X_2}(\cdot|x_1, x_2)$ are both strictly increasing

Under (i)-(iii), $r(x_1, x_2, \cdot)$ and $F_{\epsilon|X_1,X_2}(\cdot|x_1, x_2)$ are identified as

$$r(x_1, x_2, \epsilon) = F^{-1}_{\theta|X_1,X_2}(F_{\epsilon|X_1,X_2}(\epsilon|x_1, x_2)|x_1, x_2)$$

$$F_{\epsilon|X_1,X_2}(\epsilon|x_1, x_2) = F_{\theta|X_1,X_2}(\epsilon|x_1^0, x_2)$$

Remarks:

- $\theta = r_0(X) + \epsilon$, $E(\epsilon|X) = 0$, $r_0(\cdot)$ is identified

- adjustments for truncation necessary if consumers exclusion
Unobserved Heterogeneity

• Motivation: $q$ is observed imperfectly
  
  the analyst cannot observe all consumption associated with price paid
  
  the analyst observes the prices paid but not the tariff $T(\cdot)$
  
  model likely rejected because $q_i \neq q(t_i)$

$\rightarrow \epsilon$: measurement error or unobserved product heterogeneity

Luo (2013, working paper)

• Assumptions and Model
  
  (i) $U(q, \theta, \epsilon) = \theta U_0(q\epsilon)$
  
  (ii) $C(q, \epsilon) = C(q\epsilon)$
  
  (iii) $\theta$ and $\epsilon$ are independent
  
  (iv) $E(\log \epsilon) = 0$

Remark:

- $\epsilon$ acts as a quantity multiplier, $Q \equiv q\epsilon$: ‘effective’ quantity

FOC (3) and (4) still valid with $Q = Q(\theta)$
Identification and Estimation

Observations: \( t_i, q_i, i = 1, \ldots, N \)

\[ q = \frac{Q(\theta)}{\epsilon} \quad \text{or} \quad \log q = \log Q(\theta) - \log \epsilon \]

Basic Idea: Exploit the mappings \( T \leftrightarrow \theta \leftrightarrow Q \)

Using the mapping between \( \theta \) and \( t \), (iii) and (iv), we have \( \log Q(t) = \mathbb{E}[\log q | T = t] \), i.e. \( Q(\cdot) \) is identified and can be estimated using nonparametric regression

How to recover \( T(\cdot) \)?

\[ t = T(Q) \Rightarrow Q = T^{-1}(t) \text{ with } \log Q(t) = \mathbb{E}[\log q | T = t] \]

\( T(\cdot) \) is identified

\( T(\cdot) \) is increasing and concave or \( T^{-1}(\cdot) \) is increasing and convex

\( \longrightarrow \) Sieve estimator for \( T^{-1}(\cdot) \), i.e. \( Q(t) \approx \beta_0 + \beta_1 t + \sum_{k=1}^{K_N} \delta_k \psi_k(t) \), with \( K_N \) interior knots for \( t \in [\underline{t}, \bar{t}] \).

\([\underline{t}, \bar{t}]\) partitioned into \( K_N + 1 \) bins \([\tau_{k-1}, \tau_k), k = 1, \ldots, K_N + 1, \tau_0 = \underline{t}, \tau_{K_N+1} = \bar{t}\)
\[
\psi_k(t) = \begin{cases} 
0 & \text{if } t \in [-\infty, \tau_{k-1}] \\
(t - \tau_{k-1})^3/[6(\tau_k - \tau_{k-1})] & \text{if } t \in [\tau_{k-1}, \tau_k] \\
((t - \tau_{k+1})^3/[6(\tau_k - \tau_{k+1})]) + a_1 t + a_0 & \text{if } t \in [\tau_k, \tau_{k+1}] \\
a_1 t + a_0 & \text{if } t \in [\tau_{k+1}, +\infty] 
\end{cases}
\]

where \( a_1 = (\tau_{k+1} - \tau_{k-1})/2 \) and \( a_0 = ((\tau_k - \tau_{k-1})^2 - (\tau_k - \tau_{k+1})^2 + 3\tau_k(\tau_{k+1} - \tau_{k-1}))/6 \)

\[
\min_{\beta_0, \beta_1, \delta_1, \ldots, \delta_K, \beta_i \geq 0, \delta_k \geq 0} \frac{1}{N} \sum_{i=1}^{N} \left( \log q_i - \log \left( \beta_0 + \beta_1 t_i + \sum_{k=1}^{K_N} \delta_k \psi_k(t_i) \right) \right)^2,
\]

\[\rightarrow \hat{\tilde{T}}^{-1}(t) = \hat{\beta}_0 + \hat{\beta}_1 t + \sum_{k=1}^{K_N} \hat{\delta}_k \psi_k(t)\]

Remark:

- \( \beta_1 \geq 0, \delta_k \geq 0 \) to insure convexity
- \( \hat{\epsilon}_i = \hat{Q}(t_i)/q_i \rightarrow \hat{f}_\epsilon(\cdot) \) using kernel density estimator
APPLICATION TO CELLULAR PHONE

- Data: a random sample of 2,000 cellular phone consumers
  
  \[ t_i, q_{ji}, j = 1, 2, 3 \text{ and } i = 1, \ldots, 2,000 \]

  aggregation of \( q_1, q_2, q_3 \) into \( q \)

- Empirical results
Comments:
- $\hat{\kappa} = 1.78 \ (\sim 7\% \ of \ \hat{t}), \ \hat{\gamma} = 0.18$
- important variability of $\epsilon$
- IR $\sim 29\%$

- Conterfactuals (considering same set of consumers)
  - Two-part tariff: firm’s profit $-3.2\%$ and consumer net surplus $-9\%$, quite close to nonlinear pricing but consumers tend to buy more $+1.4\%$
  - Linear pricing: firm’s profit $-17.9\%$ and consumer net surplus $-32.2\%$, consumers buy much less $-47.6\%$
  - Phone plans: firm’s profit $-16.1\%$ and consumer net surplus $-29.3\%$, consumers buy less $-42.5\%$
Multiproduct Nonlinear Pricing

- Motivation: Firms produce and sell several products
  Voice and SMS by cellular phone providers

Luo, Perrigne and Vuong (2012, working paper)

- Assumptions:
  (i) to simplify the presentation, two products $q_1$ and $q_2$
  (ii) $(\theta_1, \theta_2) \sim F(\cdot, \cdot)$ on $[\underline{\theta}_1, \bar{\theta}_1] \times [\underline{\theta}_2, \bar{\theta}_2] = \Theta$, $f(\cdot, \cdot) > 0$
  (iii) $U(q_1, q_2, \theta_1, \theta_2)$ strictly increasing in all its arguments, continuous and convex in $(\theta_1, \theta_2)$, $U(0, 0, \theta_1, \theta_2) = 0$
  (iv) $U(\cdot, \cdot, \cdot, \cdot)$ homogenous of degree one in $(\theta_1, \theta_2)$
  (v) $C(q_1, q_2)$ increasing and continuously differentiable

Remarks:
- two products/types involve multidimensional screening, difficult to solve and no single crossing property
  $\rightarrow$ pooling at equilibrium
- (iv) necessary for multidimensional types
Hereafter, \( \theta = (\theta_1, \theta_2), q = (q_1, q_2) \)

How to solve for multidimensional screening?

1. Aggregation of types into a single one, i.e. \( h = h(\theta_1, \theta_2) \)

2. Rewrite the model in terms of cost \( c = C(q_1, q_2) \) leading to a cost-based tariff \( T(c) \)

\[ \rightarrow \text{cost-based indirect utility} \quad V(c, \theta) = \max_{q: C(q) \leq c} U(q, \theta) \]

- Model

More Assumptions:

(vi) \( V(c, \theta) = h(\theta)V_0(c), V_0(\cdot) \) increasing and concave, \( h(\cdot) \geq 0 \)

(vii) following (iii)-(iv), \( h(\cdot) \) is increasing and homogenous of degree one

(viii) \( f(\theta) = f^h(h(\theta))f^0(\theta), f^0(\cdot) \) homogenous of degree zero in \( \theta \)

(ix) \( h \sim \Phi(\cdot) \) on \([h, \bar{h}]\), \( \phi(h) = kf^h(h) > 0 \)

(x) \( 1 - \frac{1-\Phi(h)}{h\phi(h)} \) nondecreasing in \( h \) (stronger that \( 1 - \frac{1-\Phi(h)}{\phi(h)} \) nondecreasing)
The optimal tariff is cost-based. A $\theta$ consumer pays $t = T(C(q))$ and chooses $q$ such that

$$\max_{q: C(q) \leq c} U(q, \theta)$$

with $c$ solving

$$\max_{c \geq 0} \left[ h - \frac{1 - \Phi(h)}{\phi(h)} \right] V_0(c) - c$$

$\rightarrow$ one-to-one mapping between $c$ and $h$ as $c = C(h)$

$\rightarrow$ exclusion condition, there is a threshold level $h^*$ defined as

$$\left[ h^* - \frac{1 - \Phi(h^*)}{\phi(h^*)} \right] V_0(C(h^*)) - C(h^*) = 0$$

and $T(C(h^*)) = h^*V_0(C(h^*))$ (boundary condition)

FOC are

$$T_c(c) = hV_{0c}(c)$$

$$hV_{0c}(c) = 1 + \frac{1 - \Phi(h)}{\phi(h)}V_{0c}(c)$$
Remarks:

- (7) results from consumer optimization, i.e. $\max_c hV_0(c) - T(c)$ when offered a cost-based tariff
- (8) results from firm optimization
- one-to-one mapping between $t$ and $c$ as $t = T(c)$, $T(\cdot)$ increasing and concave

- Identification

\[ U(\cdot, \cdot, \cdot, \cdot), f(\cdot, \cdot), h(\cdot, \cdot), C(\cdot, \cdot) \]
\[ G^{q_1,q_2}(\cdot, \cdot), G'(\cdot) \]

Model Primitives Observables

Can we identify?

The model is not identified. Moreover, likely to be rejected with data, i.e. $(q_{1i}, q_{2i}) \neq (q_1(t_i), q_2(t_i))$

\[ \rightarrow \text{unobserved heterogeneity } (\epsilon_1, \epsilon_2) \sim f_\epsilon(\cdot, \cdot) \]
Identifying assumptions:

(i) \( U(q_1, q_2, \theta_1, \theta_2, \epsilon_1, \epsilon_2) = U(q_1\epsilon_1, q_2\epsilon_2, \theta_1, \theta_2) \equiv U_0(Q_1, Q_2, \theta_1, \theta_2) \)

(ii) \( C(q_1, q_2, \epsilon_1, \epsilon_2) = C(Q_1, Q_2) = \kappa(1 + Q_1)^\gamma(1 + Q_2)^{1-\gamma} \)

(iii) \( \text{E}[\log \epsilon_1] = \text{E}[\log \epsilon_2] = 0 \)

(iv) \( \bar{h} = 1 \) (normalization)

Remarks:

- FOC (7) and (8), exclusion and boundary conditions remain with \((Q_1, Q_2)\) instead of \((q_1, q_2)\)

- \( t = T[\kappa(1 + Q_1(h))^{\gamma}(1 + Q_2(h))^{1-\gamma}] \), \( q_1 = Q_1(h)/\epsilon_1 \), \( q_2 = Q_2(h)/\epsilon_2 \)

- in (viii), \( f^0(\theta, \epsilon) \Rightarrow \epsilon \) and \( h \) are independent

Identification in two steps:

1. Assume \( c \) known (and hence \( T(c) \)), identify \( Q_1(t), Q_2(t), f_\epsilon(\cdot, \cdot) \) \( V_{0c}(\cdot) \) and \( \phi(\cdot) \)

2. Identify \( \kappa, \gamma \) and \( T(\cdot) \)
Using independence of $\epsilon$ with $h$ and hence $t$, log $Q_1(t) = E[\log q_1|T = t]$ and log $Q_2(t) = E[\log q_2|T = t]$

$\rightarrow Q_1(\cdot)$ and $Q_2(\cdot)$ are identified on $[t, \bar{t}]$

$\rightarrow f_{\epsilon|t}(\cdot, \cdot|t)$ and hence $f_{\epsilon}(\cdot, \cdot)$ are identified

Since (7) and (8) are similar to (3) and (4), $V_0(c) = \frac{T_c(c)}{\xi(c)}$ and $h(c) = \xi(c)$, with

$$
\xi(c) = (1 - G^c(c))^{-1} \exp \left\{ - \int_c^\bar{c} \frac{T_{cc}(x)}{T_c(x)^2} \log(1 - G^c(x)) dx \right\}, \quad \xi(\bar{c}) = 1 \text{ and } \xi(c) = h^*
$$

$\rightarrow V_0(c)$ and $h(\cdot)$ are both identified on $[c, \bar{c}]$

$\rightarrow \phi^*(\cdot)$ is identified on $[h^*, \bar{h}]$ as well as $f^h(\cdot)$ up to a multiplicative constant since $\phi(h) = khf^h(h)$

From the one-to-one mapping between $h$ and $t$, $G^t(t) = \Phi(h(t))$ and $g_t(t) = \phi(h(t))h'(t)$

Using the above in the exclusion condition in addition to $c = T^{-1}(t)$, $V_0(\underline{c}) = \underline{t}/h^*$ and $T_t^{-1}(t) = 1/T_c(c)$ gives

$$
\underline{c} = T_t^{-1}(t)\underline{t}
$$
Using $T^{-1}(t) = \kappa(1 + Q_1(t))^\gamma(1 + Q_2(t))^{1-\gamma}$ and taking the logarithm gives

$$T_t^{-1}(t) = T^{-1}(t) \left( \gamma \frac{d \log(1 + Q_1(t))}{dt} + (1 - \gamma) \frac{d \log(1 + Q_2(t))}{dt} \right)$$

leading to

$$\gamma = \frac{1}{\kappa} - \frac{d \log(1 + Q_2(t))}{dt} \frac{d \log(1 + Q_1(t))}{dt}$$

Evaluating (7) and (8) at $\bar{t}$ and $\bar{c}$ gives $T_c(\bar{c}) = 1$. After some algebra

$$\frac{1}{\kappa} = (1 + Q_1(\bar{t}))^\gamma(1 + Q_2(\bar{t}))^{1-\gamma} \left( \gamma \frac{d \log(1 + Q_1(\bar{t}))}{dt} + (1 - \gamma) \frac{d \log(1 + Q_2(\bar{t}))}{dt} \right)$$

$\rightarrow$ $\kappa$ and $\gamma$ are identified

$\rightarrow T(\cdot)$ is identified from $t = T[\kappa(1 + Q_1(t))^\gamma(1 + Q_2(t))^{1-\gamma}]$

$\rightarrow$ assuming $U(Q_1, Q_2, \theta_1, \theta_2) = h(\theta_1, \theta_2)U_0(Q_1, Q_2)$, $U_0(\cdot, \cdot)$ is identified at values $Q_1(t)$ and $Q_2(t)$, $t \in [t, \bar{t}]$
• Estimation

Observations \((q_{1i}, q_{2i}, t_i), i = 1, \ldots, N\)

Multistep Procedure

STEP 1:

Estimate \(Q_1(t)\) and \(Q_2(t)\) using sieve estimators (easier to compute derivatives) \(\rightarrow \hat{Q}_1(\cdot), \hat{Q}_2(\cdot)\)

Estimate \(\hat{\epsilon}_{1i}, \hat{\epsilon}_{2i}, i = 1, \ldots, N \rightarrow \hat{f}_{\epsilon_1, \epsilon_2}(\cdot, \cdot)\)

Estimate \(t\) and \(\bar{t}\) using minimum and maximum estimators over the \(t_i\)s \(\rightarrow t_{\min}\) and \(t_{\max}\)

Using \(\hat{Q}_1(t_{\min}), \hat{Q}_1(t_{\max}), \hat{Q}_2(t_{\min}), \hat{Q}_2(t_{\max})\), estimate \(\kappa\) and \(\gamma \rightarrow \hat{\kappa}, \hat{\gamma}\)

Using \(T^{-1}(t) = \kappa(1 + Q_1(t))^{\gamma}(1 + Q_2(t))^{1-\gamma}\), \(\hat{\kappa}\) and \(\hat{\gamma}\), estimate \(T^{-1}(\cdot)\) by a sieve constrained (increasing and convex) estimator \(\rightarrow \hat{T}^{-1}(\cdot)\)

STEP 2: Provide estimators using \(t\) instead of \(\hat{c}\)

Noting that \(c = T^{-1}(t)\) and \(T_{tt}^{-1}(t) = -(T_{tt}[T^{-1}(t)]/T_{t}^{2}(T^{-1}(t))) \times T^{-1}(t)\), rewrite \(\xi(c)\) as \(\xi(T^{-1}(t))\) and by decomposing the integral into a sum

\[
\hat{\xi}(\hat{T}^{-1}(t)) = \left[1 - \hat{G}^{t}(t)\right]^{\hat{T}^{-1}(t)-1} \exp \left\{ \left[\hat{T}^{-1}_{t}(t^{j+1}) - \hat{T}^{-1}_{t}(t)\right] \log [1 - \hat{G}^{t}(t^{j})] + \sum_{s=j+1}^{J-1} \left[\hat{T}^{-1}_{t}(t^{s+1}) - \hat{T}^{-1}_{t}(t^{s})\right] \log [1 - \hat{G}^{t}(t^{s})] \right\}
\]
\[ V_{0t}(t) \equiv \hat{V}_{0t}(\hat{T}^{-1}(t)) = \frac{1}{\hat{T}^{-1}(t)\xi(\hat{T}^{-1}(t))}, \quad \hat{h}(t) \equiv \hat{h}(\hat{T}^{-1}(t)) = \hat{\xi}(\hat{T}^{-1}(t)) \]

\[ \hat{V}_0(t) = \int_{t_{min}}^{t} \frac{1}{\hat{T}^{-1}(y)\xi(\hat{T}^{-1}(y))} dy + \frac{t_{min}}{\hat{h}^*} = \hat{U}_0(\hat{Q}_1(t), \hat{Q}_2(t)) \text{ using } \hat{V}_0(t_{min}) = t_{min}/\hat{h}^* \]

\[ \hat{\phi}^* (h) \]

**APPLICATION TO VOICE AND SMS**

- Data: 4,601 observations, bill $t$, quantity of phone calls $q_1$ and number of SMS $q_2$
  - large number of additional services, low correlation between $q_1$ and $t$ and $q_2$ and $t$, concave tariff
Empirical results

\( \hat{\gamma} = 0.175, \epsilon_m \) includes multimedia and data usage

1 \( \frac{\Phi^*(h)}{h\phi^*(h)} \) increasing

IR \( \sim 0.32 \) with important variability

weak correlation between \( \epsilon_1 \) and \( \epsilon_2 \) at 0.03
Bundling and Nonlinear Pricing

• Motivation: Products offered in a bundle or separately, i.e. mixed bundling, such as internet and phone

Luo (2012, working paper)

• Assumptions:

(i) \(q \geq 0, j = \{0, 1\} \in J\) (can be generalized to \(J \geq 2\))
(ii) \((\theta, \beta) \sim F(\cdot, \cdot)\) on \([\theta, \bar{\theta}] \times [\underline{\beta}, \overline{\beta}]\) with \(\underline{\beta} < 0 < \overline{\beta} \leq 1\), \(f(\cdot, \cdot) > 0\)
(iii) \(H(\theta|j) \equiv \theta - \frac{1-F(\theta|D(\beta)=j)}{f(\theta|D(\beta)=j)}\) nondecreasing in \(\theta\) with \(D(\beta) \equiv \min\{j \in J : j \geq \beta\}\)
(iv) \(\frac{1-F(\theta|D(\beta)=j)}{f(\theta|D(\beta)=j)}\) increasing in \(j\)
(v) \(\tilde{U}(q, j, \theta, \beta) = U(q, j, \theta) \equiv U_0(q, \theta) + U_1(q, j)\) for \(j \geq \beta\)
(vi) \(U(0, 0, \theta) = 0, U(q, j, \theta)\) increasing and concave in both \(q\) and \(\theta\), \(U_{q\theta}(q, j, \theta) > 0\)
(vii) \(\frac{\partial}{\partial \theta} - \frac{U_{qq}(q, j, \theta)}{U_q(q, j, \theta)} \leq 0\)
(viii) \(C(q)\) increasing satisfying \(\frac{C_{qq}(q, j)}{C_q(q, j)} > \frac{U_{qq}(q, j, \theta)}{U_q(q, j, \theta)}\)
(ix) \(U_0(0, \theta) = 0, U_1(0, j) > C(0, j)\)
(x) \(U_q(q, 1, \theta) - U_q(q, 0, \theta) \leq C_q(q, 1) - C_q(q, 0), U_1(0, 1) - U_1(0, 0) \leq C(0, 1) - C(0, 0)\)
Remarks:

- (iv) creates some affiliation between $\theta$ and $j$
- (v) implies $U_{\theta j} (\cdot, \cdot, \cdot)$ independent of $\theta$
- (ix) implies utility independent of \textsl{theta} when $q = 0$ and insures that selling $j$ only is possible

\textbullet Model

Firm’s optimization: Find $T(\theta, \beta), q(\theta, \beta), j(\theta, \beta)$ that maximizes profit s.t. (IC), (IR) and $j(\theta, \beta) \geq \beta$

Proceed in two steps: (1) assume $\beta$ known, and (2) show mechanism optimality with $\beta$ unknown to the firm

Under (v) and (x), $j(\theta, \beta) = D(\beta)$, $q(\theta, \beta) = q(\theta, D(\beta))$, $t(\theta, \beta) = t(\theta, D(\beta))$, i.e. one can use $D(\beta)$ or $j$

Exclusion: $\theta^*_j$ solves

$$U_0(q(\theta, j), \theta) + U_1(q(\theta, j), j) - U_1(0, j) - [C(q(\theta, j), j) - C(0, j)] - U_{0\theta}(q(\theta, j), \theta) \frac{1 - F(\theta|D(\beta))}{f(\theta|D(\beta))} = 0$$

$\rightarrow$ if $\beta \leq 0$, then $D(\beta) = 0$

if $\theta < \theta^*_0$, then $q = 0$, i.e. consumers buy nothing

if $\theta > \theta^*_0$, then $q > 0$, consumers pay $T_0(q)$
if $\beta > 0$, then $D(\beta) = 1$

If $\theta < \theta_1^*$, then $q = 0$, i.e. consumers buy 1 only and pay $t_1$

If $\theta > \theta_1^*$, then $q > 0$, consumers buy bundle $(q, 1)$ and pay $T_1(q)$

For $\theta \in [\theta_1^*, \bar{\theta}]$ and given $D(\beta) = j$ FOC defining $q(\cdot, j)$ and $T(\cdot, j)$

\[
U_{0q}(q(\theta, j), \theta) + U_{1q}(q(\theta, j), j) = C_q(q(\theta, j), j) + U_{0q}(q(\theta, j), \theta) \frac{1 - F(\theta|D(\beta) = j)}{f(\theta|D(\beta) = j)} \tag{9}
\]

\[
T_q(q, j) = U_{0q}(q(\theta, j), \theta) + U_{1q}(q(\theta, j), j) \tag{10}
\]

Remarks:
- $q(\theta, j)$ decreasing in $j$ and $\theta_j^*$ increasing in $j$, $T(q, j) - U_1(q, j)$ increasing in $j$
- under incomplete information in $\beta$, $q(\cdot, \cdot)$ and $T(\cdot, \cdot)$ satisfy (9) and (10)

• Identification

$F(\cdot|D(\beta) = j) \equiv F_j(\cdot)$, $C(\cdot, j) \equiv C_j(\cdot)$, $T(\cdot, j) \equiv T_j(\cdot)$

Remark:
- To simplify, $T_j(\cdot)$ is known.
The model is **not** identified

Identifying assumptions:

(i) $U_0(q, \theta) = \theta U_0^0(q)$, $U^0(0) = 0$, $U^0_0(\cdot) \geq 0$, $U^0_{qq}(\cdot) \leq 0$

(ii) $U_1(q, 0) = 0$ (normalization)

(iii) $C_j(q) = \kappa_j + \gamma_j q$ if $q > 0$ and $0$ if $q = 0$

(9) and (10) become

\[
\theta U^0_q(q) + U_1q(q, j) = \gamma_j + \frac{1 - F_j(\theta)}{f_j(\theta)} U^0_q(q) \quad (11)
\]
\[
T_{jq}(q) = \theta U^0_q(q) + U_1q(q, j) \quad (12)
\]

one-to-one mapping between $q$ and $\theta$ given $j$

$q_j = q(\theta^*_j, j)$, $q_\ast = q(\bar{\theta}, j)$, $t^*_j = \theta^*_j U^0(q_j) + U_1(q_j, j)$, $j = 0, 1$

$t_1 = U_1(0, 1)$ (1 only) + exclusion conditions defining $\theta^*_0$ and $\theta^*_1$
\[\gamma_j = T_{jq}(q_j), \quad t_1 = U_1(0, 1), \quad \text{proportion of consumers buying 1 only equals } F(\theta_1^*|1)\]

\[\rightarrow \gamma_j, U_1(0, 1) \text{ and } F(\theta_1^*|1) \text{ are identified}\]

\(j = 0\), i.e. consumers buying \(q > 0\) only, \(U_1(q, 0) = 0\) by assumption

\[\text{FOC } \rightarrow \theta U_q^0(q) = \gamma_j + \frac{1-F_0(\theta)}{f_0(\theta)} \text{ and } T_0q(q) = \theta U_q^0(q), \quad q \in [q_0, \bar{q}_0] \text{ as in the basic model!}\]

\[\rightarrow U_q^0(q) = T_0q(q)/\xi(q) \text{ and } \theta_0(q) = \xi(q) \text{ (under normalization } \bar{\theta} = 1)\]

\[\xi(q) = [1 - G_0^*(q)]^{\gamma_0(q)} - 1 \exp \left\{ -\gamma_0 \int_{q_0}^q \log[1 - G_0^*(x)] \frac{T_{0qq}(x)}{T_{0q}(x)} \, dx \right\}\]

\[\rightarrow U_0^0(q) = \frac{\theta_0}{\bar{\theta}} + \int_{q_0}^q U_q^0(x) \, dx \text{ and } F_0^*(\theta) \text{ are identified on } [q_0, \bar{q}_0] \text{ and } [\theta_0, \bar{\theta}]\]

\(j = 1\), i.e. consumers buying \(q > 0\) and 1, exploit the mapping between \(q\) and \(\theta\) to rewrite FOC (11) and (12)

\[\rightarrow \theta_1(q) = 1 - \int_q^{q_1} \frac{T_{1q}(x)}{U_q^0(x)} \frac{q_1^*(x)}{1-G_1^*(x)} \, dx, \quad U_{1q}(q, 1) = T_{1q}(q) - \theta_1(q)U_q^0(q)\]

\[\rightarrow U_{1q}(\cdot, 1) \text{ is identified on } [\max\{q_1, q_0\}, \bar{q}_1]\]

\[F_1^*(\cdot) \text{ is identified on } [\theta_1(\max\{q_1, q_0\}), \bar{\theta}]\]

36
Using exclusion condition and FOC, $\kappa_0$ and $\kappa_1$ are identified by

$$
\kappa_0 = \gamma_0 \left( \frac{t_0^*}{T_0q(q_0)} - q_0 \right), \quad \kappa_1 = t_1^* - t_1 - \gamma_1 q_1 - U^0(q_1) \frac{1 - F_1(\theta_1^*)}{f_1(\theta_1^*)}
$$

- Estimation

Observables: $(q_{0i}, t_{0i})$ for those consuming $q$ only, $(q_{1i}, t_{1i})$ for those consuming $q$ and 1, $t_1$ and their number for those consuming 1 only

Follow the steps of identification using minimum, maximum, empirical cdf and kernel estimators

$$
\rightarrow \hat{U}^0(q), \hat{U}_1q(\cdot), \hat{f}_0^*(\cdot), \hat{f}_1^*(\cdot), \hat{\gamma}_0, \hat{\kappa}_0, \hat{\gamma}_1, \hat{\kappa}_1
$$

**Application to China Telecom Phone and Internet**

- Data: phone consumption from land line $q$ and 2 levels of internet, i.e. $j = 0, 1, 2$, fixed fee for internet only, concave tariff in $q$ for $j = 0, 1, 2$, rate per minute increases with $j$

Remark: unobserved heterogeneity $\epsilon$ to estimate $T_j(\cdot)$ with sieve estimator
Comments:
- $\hat{H}(\cdot|j)$ increasing as assumed in the model
- $\hat{U}_{1j\delta}(\cdot) <, j = 1, 2$ suggesting complementarity between phone and internet
- IR $\sim$ 29.8% decreasing in $j$
- component pricing: firm’s profit $- 10.1\%$, consumer net surplus $- 17.2\%$, more exclusion, larger $T_0(\cdot)$ and $T_1(\cdot)$ but lower $T_2(\cdot)$