1 Outline.

1. Motivation.

2. Two-Step Estimation of Auction Models.

3. Two-Step Estimation of Demand in Differentiated Product Markets.

4. Two-Step Estimation of Games.

- Structural models are becoming increasingly used in industrial organization.

- These methods are useful in markets because they allow for a logic link between game theory and empirical practice.

- Unfortunately, they are often quite difficult to estimate.

- Frequently estimation of structural models requires computation of the equilibrium within the estimation procedure.

- Nonlinear optimization and difficult simulation procedures are also required.

- Numerical reliability is frequently an issue.

- Making the estimator feasible often requires the economist to make strong parametric assumptions.

- Recently, two-step nonparametric approaches are being increasingly applied when estimating structural models.

- Started in auctions (see Guerre, Perrign and Vuong (2000, Econometrica)).

- In a first step, the economist estimates some part of the policy function or constraint for an agent.

- In auctions, you nonparametrically estimate bid distributions.

- In demand, you nonparametrically estimate a hedonic.

- In games, you nonparametrically estimate policy functions.

- In a second step, you apply the first order conditions for optimality to find structural parameters to rationalize the first step.

- In some cases, this may involve inequalities instead of first order conditions.

- Two step estimators are much easier to compute and require fewer parametric assumptions.

- Can be a loss of efficiency and bias due to first step errors.

2 Example #1: Auctions.

- The first example we consider is first-price sealed-bid auctions.

- The logic of this estimator works in far more general auctions.

- See Pesendorfer and Jofre-Bonet (Econometrica 2003), Hortacsu (2004), Bajari and Ye (2003) and many papers by Vuong with co-authors.

- Let v be a valuation and F be the distribution of valuations.

- First price sealed bid auction.

- $b = \mathbf{b}(v)$ denote the equilibrium bid function.

- Bidder i's profit from bidding b_i is:

$$\pi_i(b_i; v_i) \equiv (v_i - b_i)F(\phi(b_i))^{N-1}$$

- Bidder i's expected utility is i's surplus $v_i - b_i$, conditional on winning, times the probability that bidder i wins the auction, $F(\phi(b))^{N-1}$.

2.1 Identification and Estimation.

- Let $G(b)$ and $g(b)$ be the distribution and density of the bids, respectively.

- First order conditions can be expressed as:

$$v_i = b_i + \frac{G(b_i)}{g(b_i)(N-1)} \ . \tag{1}$$

- If we could estimate G and g, we could recover v_i.

- Logic is close to using an estimate of MR to recover MC.

- Suppose that the econometrician observed T repetitions of the auction.

- Let $b_{i,t}$ denote the bid that i submits in auction t.

- Since we have multiple repetitions of the same auction, it is possible to estimate G and g.

- Denote these estimates as $\hat{G}(b)$ and $\hat{g}(b)$.

- If we substitute the estimated distributions into equation (1), we can generate an estimate of $\widehat{v}_{i,t}$ of $v_{i,t}$, bidder i's valuation in the tth auction as follows:

$$\widehat{v}_{i,t} = b_{i,t} + \frac{\widehat{G}(b_{i,t})}{\widehat{g}(b_{i,t})(N-1)}$$

$$\widehat{v}_{i,t} = b_{i,t} + \frac{\widehat{G}(b_{i,t})}{\widehat{g}(b_{i,t})(N-1)}. \quad (2)$$

- By applying equation (2) to every bid in our data set, we can generate estimates, $\left\{\widehat{v}_{i,t}\right\}_{i=t,..,N,\ t=1,...,T}$ of the valuations associated with each bid in our data set.

- Then estimate $F(v)$.

To summarize, the estimation procedure involves two steps.

1. First, using non-parametric methods generate estimates $\widehat{G}$ and $\widehat{g}$ of G and g.

2. Given the first stage estimates, apply equation (2) for every observed bid $b_{i,t}$ to generate $\widehat{v}_{i,t}$, an estimate of $v_{i,t}$.

For a technical discussion of this estimator, the interested reader is referred to Guerre, Perrigne and Vuong (2000).

2.2 The Risk-Averse Model.

- CRRA utility function, $U(x) = x^\theta$. In this specification, $1 - \theta$ is the coefficient of relative risk aversion, with $\theta = 1$ corresponding to risk neutrality.

- In this model, the first order condition is:

$$v_i = b_i + \theta \cdot \frac{G(b_i)}{g(b_i)(N-1)}. \tag{3}$$

- Observe that when bidders are risk neutral, that is $\theta = 1$ then this is the risk neutral foc.

2.3 Structural Estimation

- The logic of the estimator is similar to the previous section.

- If the economist knew G and θ, then we could construct a two-step estimator along the lines of the previous section.

- The problem that we face, however, is that θ is not directly observed.

- Let $G(b; N)$ denote the distribution of bids with N bidders.

- Suppose that $F(v)$ is independent of N.

- Let v_α denote the α^{th} percentile of the distribution of valuations.

- Let $b_\alpha(3)$ denote the α^{th} percentile of $G(b;3)$ and let $b_\alpha(6)$ denote the α^{th} percentile of $G(b;6)$.

- By equation (3) it follows that

$$v_\alpha = b_\alpha(3) + \theta \cdot \frac{G(b_\alpha(3);3)}{2g(b_\alpha(3);3)} \quad (4)$$

$$v_\alpha = b_\alpha(6) + \theta \cdot \frac{G(b_\alpha(6);6)}{5g(b_\alpha(6);6)} \quad (5)$$

- By simple algebra, it follows from the equations (4) and (5) that:

$$b_\alpha(3) - b_\alpha(6) = \theta \cdot \left(\frac{G(b_\alpha(6); 6)}{5g(b_\alpha(6); 6)} - \frac{G(b_\alpha(3); 3)}{2g(b_\alpha(3); 3)} \right) \tag{6}$$

- Equation (6) suggests a simple way to estimate θ.

- If we knew the distribution of bids in the 3 and 6 bidder experiments, given α, all of the terms on the left and right hand in this equation would be directly observable except for θ.

- By evaluating (6) at a large number of percentiles, we could then estimate θ using regression.

- Given an estimate $\widehat{\theta}$ of θ, we can then estimate the valuations v_i by evaluating the empirical analogue of equation (3) as in the previous section.

To summarize, we generate estimates $\widehat{v}_{i,t}$ of $v_{i,t}$ as follows:

1. Generate non-parametrically estimates $\widehat{G}(b; N)$ and $\widehat{g}(b; N)$ of $g(b; N, e)$ and $G(b; N, e)$.

2. Generate an estimate $\widehat{\theta}$ of θ by running the following regression, using a finite number of percentiles α:

$$\widehat{b}_\alpha(3) - \widehat{b}_\alpha(6) = \theta \cdot \left(\frac{\widehat{G}(\widehat{b}_\alpha(6); 6)}{5\widehat{g}(\widehat{b}_\alpha(6); 6)} - \frac{\widehat{G}(\widehat{b}_\alpha(3); 3)}{2\widehat{g}(\widehat{b}_\alpha(3); 3)} \right) + \varepsilon_\alpha \tag{7}$$

3. Given $\widehat{\theta}, \widehat{G}(b; N)$ and $\widehat{g}(b; N)$ use the empirical analogue of (3) to generate an estimate $\widehat{v}_{i,t}$ of $v_{i,t}$.

$$\widehat{v}_{i,t} = b_{i,t} + \widehat{\theta} \cdot \frac{\widehat{G}(b_i)}{\widehat{g}(b_i)(N-1)} \tag{8}$$

- Experimental first-price auctions with 3 and 6 bidders.

- Estimate valuations using nonparametric techniques above and then ask if you got the right answer.

- Plot the estimated versus actual distribution of valuations.

- Truth is uniform [0,30].

- Observe all bids, impliment estimator.

- Compare estimated valuations to experimental valuations.

Figure 2: Histograms of Estimated and Actual Valuations, Risk Neutral Model.

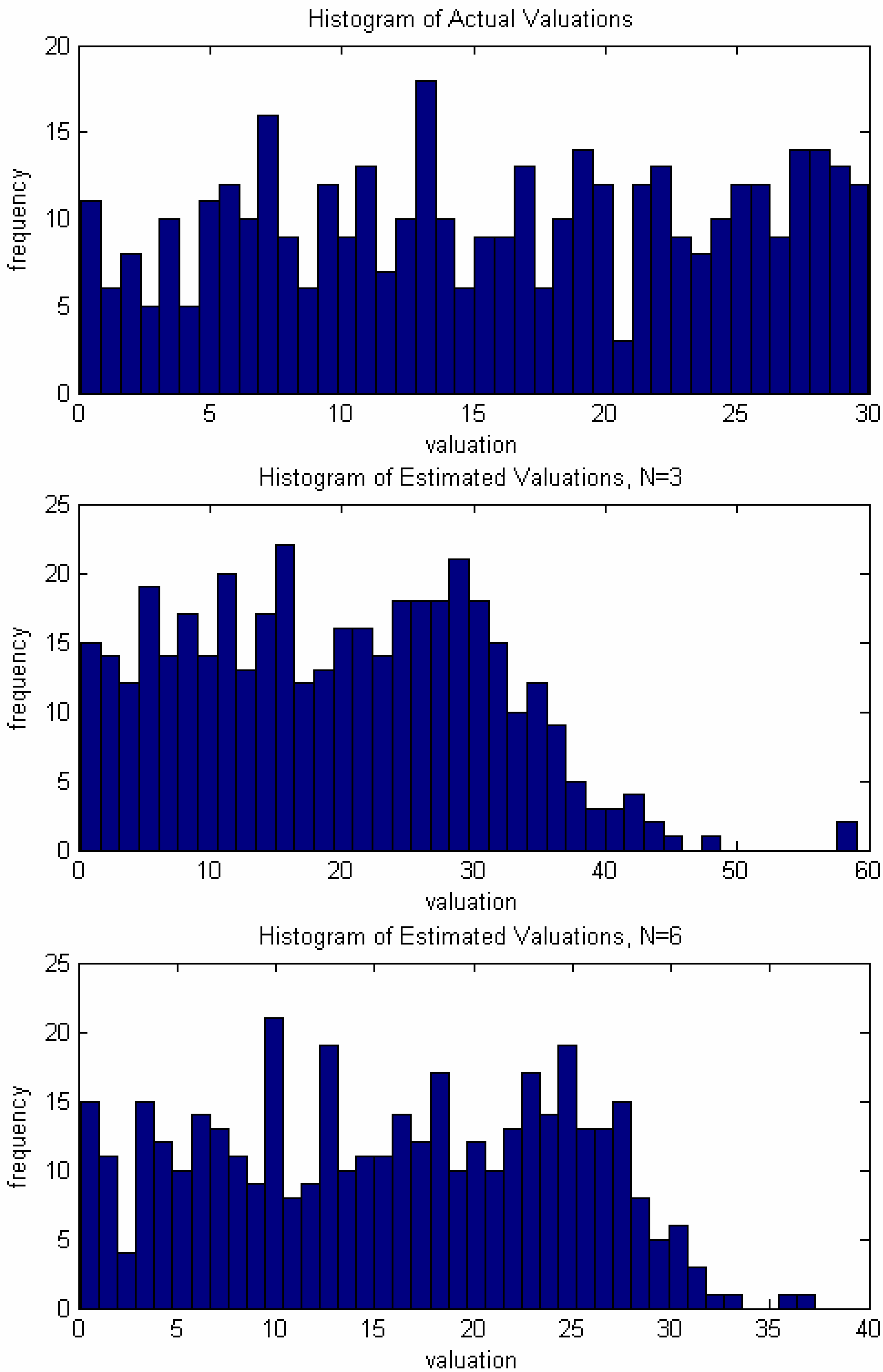
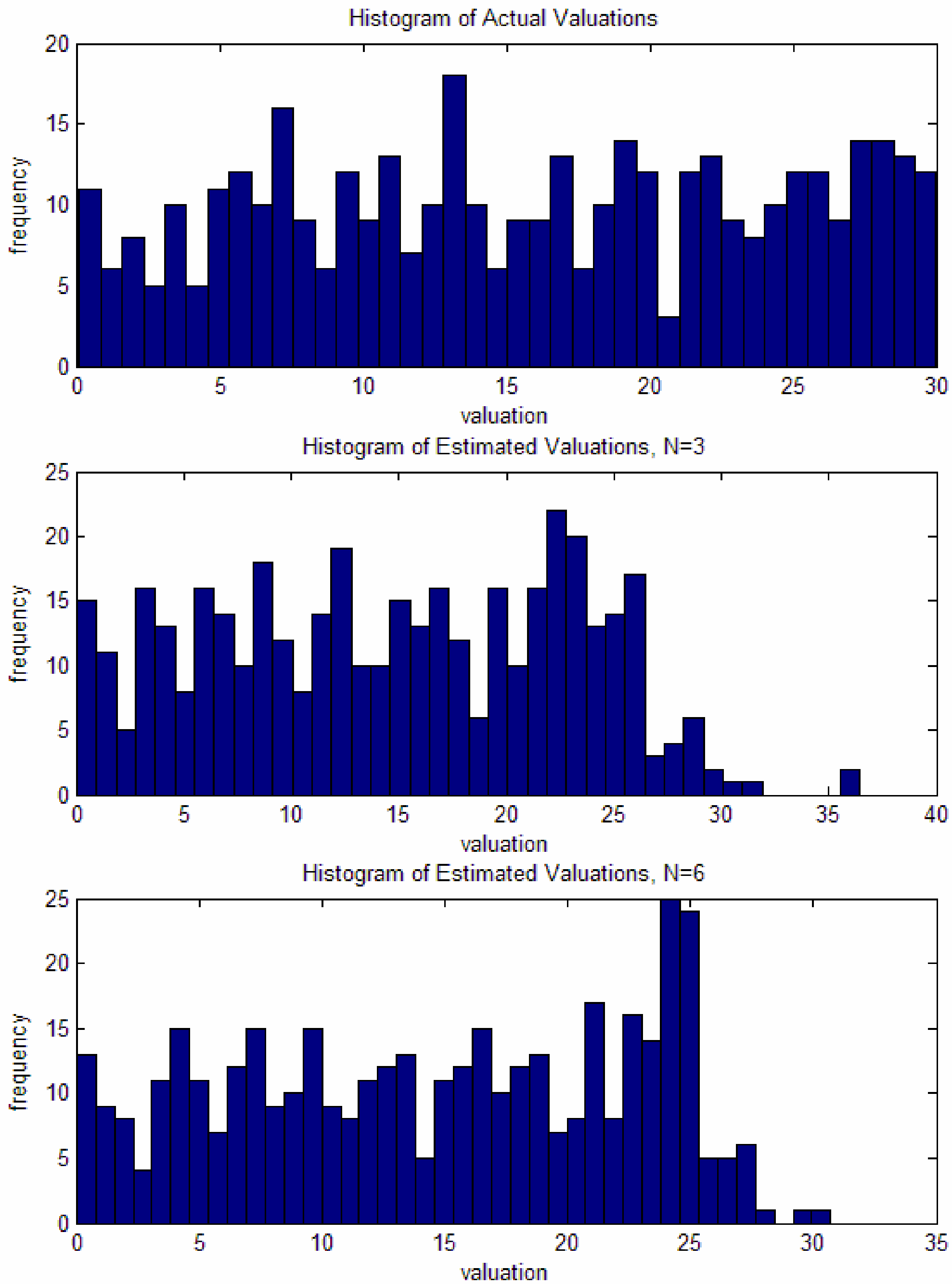


Figure 3: Histograms of Estimated and Actual Valuations, Risk Aversion Model.



3 Demand Estimation.

- The second example we consider is a nonparametric demand estimation problem.

- The demand system comes from Bajari and Benkard (2001), but Petrin and Train (2004) use a similar logic.

- Consider a model like BLP, only drop the parametric assumptions about the distribution of random coefficients.

- In the I.O. literature, many researchers, Berry (1994), BLP (1995), Nevo (2001), Petrin(2002)... have studied models of the form:

$$u_{ij} = \sum_k \beta_{i,k} \log(x_{j,k}) + \xi_j - \alpha_i \log(y_i - p_j) + \varepsilon_{ij}$$
$$(\beta_i, \alpha_i) \sim F(d_i; \theta)$$
$$E(\xi_j | x_j) = 0$$

- $x_{j,k}$ is a vector of characteristics for product j.

- ξ_j is a scalar characteristic observed to the consumer, but not the economist.

- ε_{ij} is a household level idiosyncratic taste shock (e.g. logit, probit, GEV...).

- (β_i, α_i) are random coefficients from a parametric distribution which depend on a vector of household level demographics d_i.

- The model that we will study will be of the form

$$u_{ij} = \sum_k \beta_{i,k} \log(x_{j,k}) + \alpha_i \log(\xi_j) + \log(y_i - p_j)$$
$$(\beta_i, \alpha_i) \sim F(d_i)$$
$$p_j = \mathbf{p}(x_{j,k}, \xi_j)$$
$$\xi_j \perp x_j$$

- $(\beta_i, \alpha_i) \sim F(d_i)$ will be estimated nonparametrically.

- No ε_{ij} (the model above will typically be just identified)

- ξ_j independent from $x_{j,k}$, instead of mean independence

- We assume that a hedonic price function $p_j = \mathbf{p}(x_{j,k}, \xi_j)$ exists.

- We estimate this function using methods based on Matzkin (1999).

- Suppose that product characteristics are continuous, then the first order conditions can be written as:

$$
\begin{aligned}
\beta_{i,k} &= \frac{x_{j,k}}{y_i - p_j} \frac{\partial \mathbf{p}}{\partial x_{j,k}} \\
\alpha_i &= \frac{\xi_j}{y_i - p_j} \frac{\partial \mathbf{p}}{\partial \xi_j}
\end{aligned}
$$

- Given an estimate of $\mathbf{p}$ and ξ_j, we can recover household level taste parameters.

- Next, we can estimate the following equation to recover the joint distribution of tastes and demographics

$$
\begin{aligned}
\beta_{i,k} &= f_k(d_i) + \eta_i \\
E(\eta_i | d_i) &= 0
\end{aligned}
$$

- The term $f_k(d_i)$ is the part of tastes that can be explained by demographics.

- The scalar η_i is a household level taste shock.

- Our approach is computationally simple, drops parametric assumptions and reduces the influence of ε_{ij} on estimates.

- However, not all products will be strong gross substitutes if number of characteristics is much smaller than number of products.

- Also, ξ_j is independent of x_j, not just mean independent.

- Possible to characterize the joint distribution of tastes and demographics much more generally.

4 Summary of Steps.

4.1 First Step: Estimating the Hedonic

- Fix product j^* and suppose that locally the hedonic satisfies:

$$p_j = \alpha_{0,j^*} + \alpha_{1,j^*} \log(x_{j^*,1}) + ... + \alpha_{K,j^*} \log(x_{j^*,K}) + \xi_j$$

- In local linear regression, you estimate a separate set of coefficients of each x_{j^*}

- Run weighted least squares to estimate α_{j^*}

$$\alpha_{j^*} = \arg\min_{\alpha} (\mathbf{p} - \mathbf{X}\alpha)' \mathbf{W} (\mathbf{p} - \mathbf{X}\alpha) \tag{9}$$

$$\mathbf{p} = \left[p_j \right], \ \mathbf{X} = \left[x_j \right], \mathbf{W} = diag\{K_h(x_j - x_{j^*})\} \tag{10}$$

- This gives you a set of implicit prices for each observed x in the sample.

- True for both discrete and continuous product characteristics.

- The unobserved product characteristic is recovered as a residual.

$$\xi_{j^*} = p_{j^*} - \Big(\alpha_{0,j^*} + \alpha_{1,j^*} \log(x_{j^*,1}) + ... + \alpha_{K,j^*} \log(x_{j^*,K}) \Big) \tag{11}$$

4.2 Second Step: Applying the First Order Conditions.

- After estimating the implicit prices, we next estimate the preferences for continuous characteristics.

- Suppose that utility takes the following form:

$$u_{ij} = \sum_k \beta_{i,k} \log(x_{j,k}) + \alpha_i \log(\xi_j) + c$$

- Using estimates of implicit prices obtained from the first step and the observed choice of $x_{j^*,k}$, an estimation $\widehat{\beta}_{i,k}$ of $\beta_{i,k}$ can be recovered as follows:

$$\widehat{\beta}_{i,k} = x_{j^*,k} \frac{\partial \widehat{\mathbf{p}}_m(x_{j^*}, \xi_{j^*})}{\partial x_{j,k}} \tag{12}$$

- Apply the FOC's for all product characteristics and for all persons.

- This gives you an estimate of the population distribution of tastes.

4.3 Third Step: Modeling the Joint Distribution of Tastes and Demographics.

- Finally, you may want to regress the taste parameters on demographics.

- This allows you to recover joint distribution of tastes and demographics.

$$\beta_{i,k} = \theta_{0,k} + \sum_{s} \theta_{k,s} d_{i,s} + \eta_{i,k} \tag{13}$$

- We then simply estimate (13) using regression. The regression that we run is:

$$\widehat{\beta}_{i,k} = \theta_{0,k} + \sum_{s} \theta_{k,s} d_{i,s} + \eta_{i,k} \tag{14}$$

- In equation (14), we have simply substituted our estimate of $\widehat{\beta}_{i,k}$ from the second stage into equation (13).

- The residuals can be interpreted as household specific taste shocks.

- Handling dichotomous characteristics can be formulated as estimating a probit/logit.

- See Bajari and Kahn (2004).

- If you assume the shocks are normal, then with linear functional form for utility, buying a discrete characteristic involves assuming that tastes are above a threshhold.

5 Application: Demand for Personal Computers.

- Data: PC Data Retail Hardware Monthly Report.

- Includes: quantity, average price, long list of machine characteristics for 29 months of data.

- 75% of U.S. retail computer sales.

- We examine Dec. 1999 data in the paper for expositional simplicity.

- Final data set 695 machines, 19 characteristics (5 operating system dummies, CPU benchmark, MMX, RAM, hard drive capacity, CDROM, DVD, modem, modem speed, NIC, monitory dummry, monitor size, zip drive, desktop (versus tower), and refurbished.

- Assume that the price function is additive in all but 3 characteristics: CPU benchmark, RAM, hard drive capacity.

- Estimate non-separable regression model.

- Preference log linear in continuous characteristics, linear in discrete characteristics and price.

- Willingness to pay is not normally distributed and independent.

- Let consumer choose from 24 largest products (accounting for 72% of sales).

- Median elasticity of residual demand is -4 to -72 w/ median of -11.

- With all 695 product median elasticity is -100 (but search costs may be unrealistically high to locate all these products).

D Tables and Graphs

Table 1: Summary Statistics

Variable	Mean	S.D.	Min	Max	OLS Coeff
CPU Bench	1354.5	362.3	516	2544	0.836
RAMMB	74.0	35.1	16	256	3.010
HDMB	9276.8	4850.3	2100	40000	0.008
MMX	0.64	0.48	0	1	-56.971
SCSI	0.01	0.08	0	1	310.747
CDROM	0.67	0.47	0	1	26.478
DVD	0.14	0.35	0	1	32.213
NIC	0.36	0.48	0	1	9.481
Monitor?	0.31	0.46	0	1	29.625
Mon.Size	0.75	3.27	0	15	22.822
ZIP	0.05	0.22	0	1	20.440
DT	0.17	0.37	0	1	25.611
Refurb.	0.09	0.28	0	1	-144.314
No Modem	0.55	0.50	0	1	145.169
Win NT 4.0	0.02	0.14	0	1	-106.374
Win NT	0.17	0.37	0	1	22.567
Win 98	0.58	0.49	0	1	-59.590
Win 95	0.16	0.37	0	1	-42.058
Constant	(Win3.1 omitted)				-590.2
R^2					0.79
N					695

Table 2: Distribution of Standard Errors for Estimates of ξ

	Asymptotic	Bootstrap (1000 samples)
Quantiles:		
Min	0.002	0.002
0.30	0.007	0.007
0.50	0.009	0.009
0.95	0.018	0.022
0.99	0.052	0.091
Max	0.104	0.121
Average	0.010	0.012
N	695	695

Table 3: Correlation Matrix of Taste Coefficients for a Subset of Characteristics

	CPU	RAM	HD	SCSI	ξ
CPU	1.000	0.510	0.357	0.694	0.418
RAM	0.510	1.000	0.533	0.511	0.477
HDM	0.357	0.533	1.000	0.527	0.351
SCSI	0.694	0.511	0.527	1.000	0.393
ξ	0.418	0.477	0.351	0.393	1.000

Table 4: Top Five Products in 12/99

Brand/Model	CPU (Benchmark)	RAM	H.D.	Price	Sales
Hewlett Packard Pavilion 6535	Intel Celeron/466MHZ (1281)	64MB	8.4GB	590	71199
Compaq Presario 5441	AMD A6-2/475MHZ (1076)	64MB	8.0GB	540	54449
Compaq Presario 5461	AMD A6-2/500MHZ (1115)	64MB	10.0GB	727	43029
E-Machines eTower 433	Celeron/433 (1167)	32MB	4.3GB	471	40399
Hewlett Packard Pavilion 6545C	Celeron/500 (1398)	64MB	13.0GB	858	34198

Table 5: Matrix of Cross Price Elasticities for Top Five Products

	HP6535	Compaq5441	Compaq5461	E-Machines	HP6545C
HP6535	-4.14	0.12	0.00	0.43	0.28
Compaq5441	0.17	-5.95	2.98	0.73	0.55
Compaq5461	0.00	2.80	-8.00	0.85	0.11
E-Machines	0.61	0.69	0.91	-10.65	0.66
HP6545C	0.70	0.86	0.18	1.02	-4.46

Figure 1: Global Identification of Indifference Curves

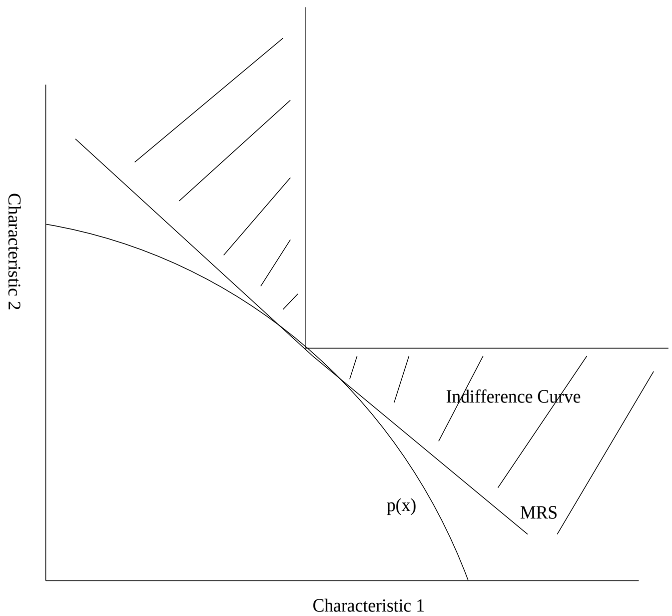


Figure 2: CPU Benchmark Willingness-To-Pay

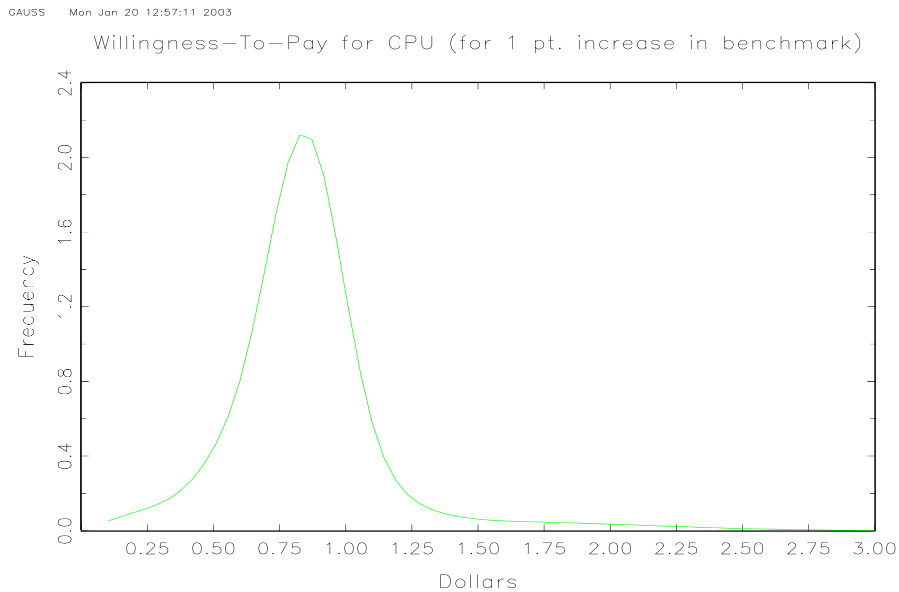


Figure 3: RAM Willingness-To-Pay

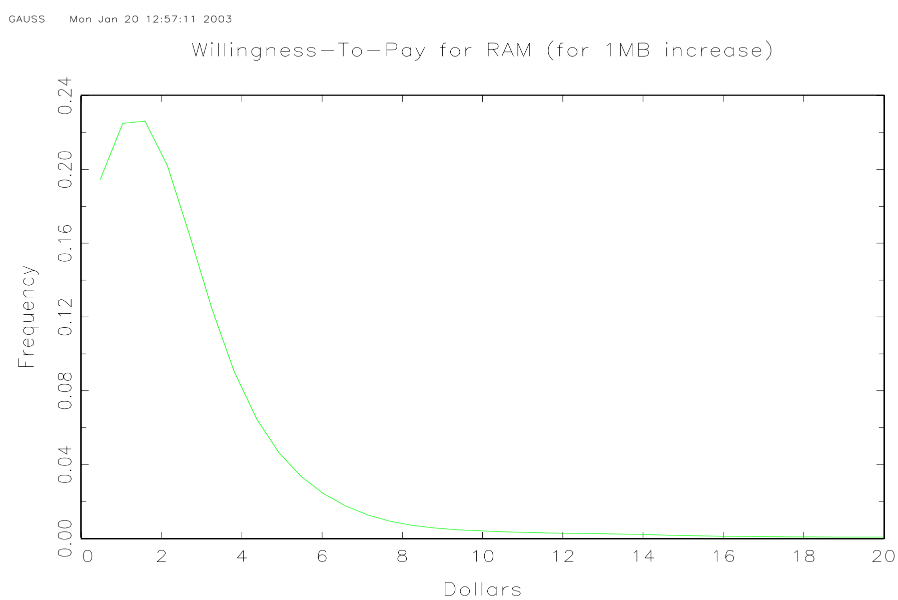


Figure 4: Hard Drive Willingness-To-Pay

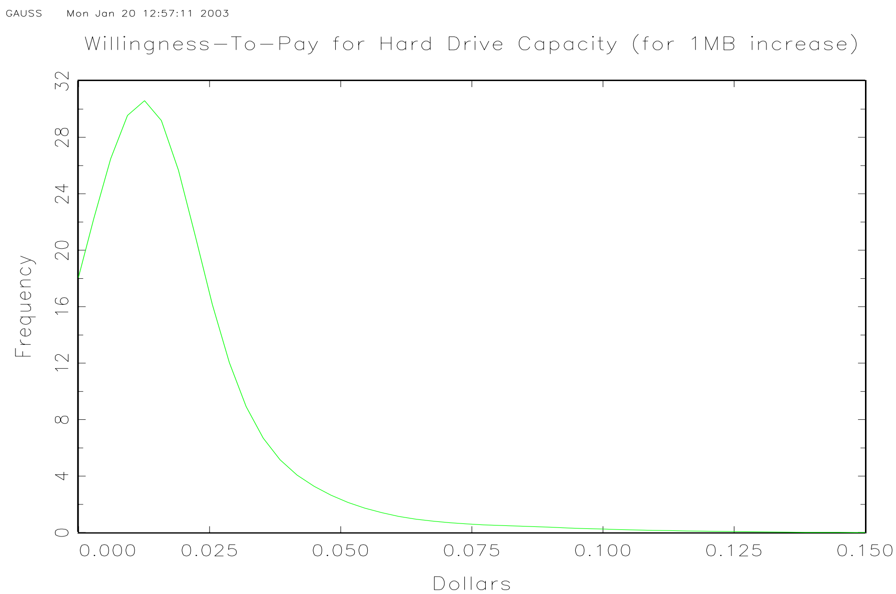


Figure 5: SCSI Willingness-To-Pay

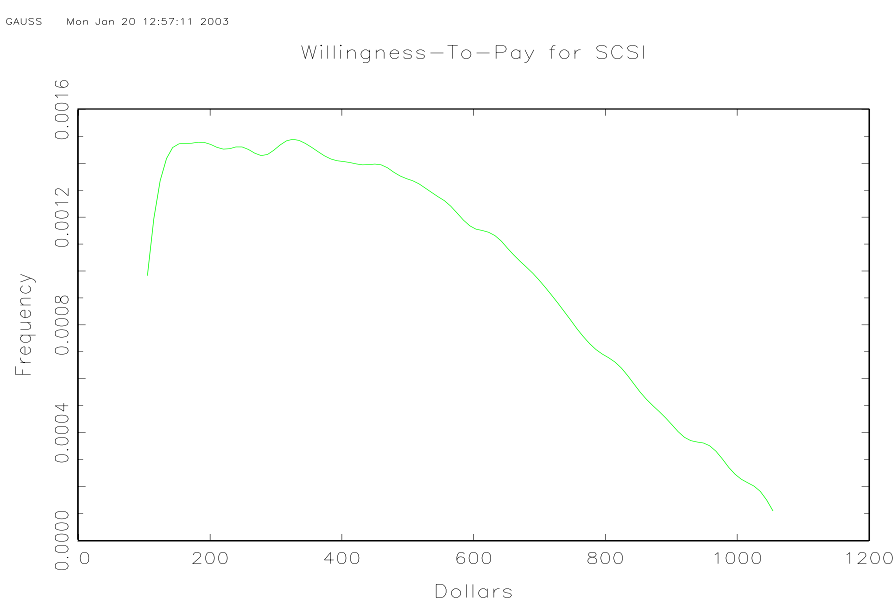


Figure 6: ξ Willingness-To-Pay

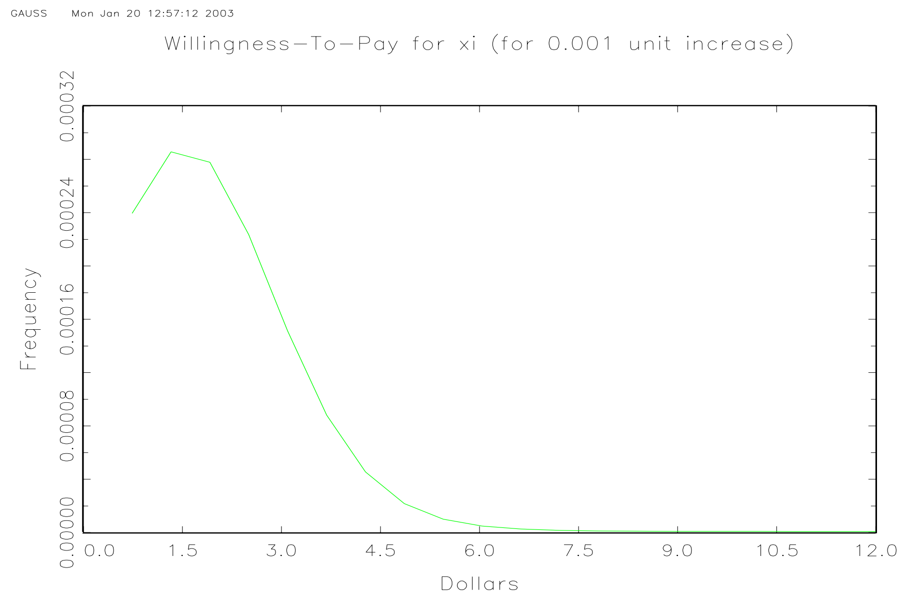


Figure 7: HP6535 Demand Curve

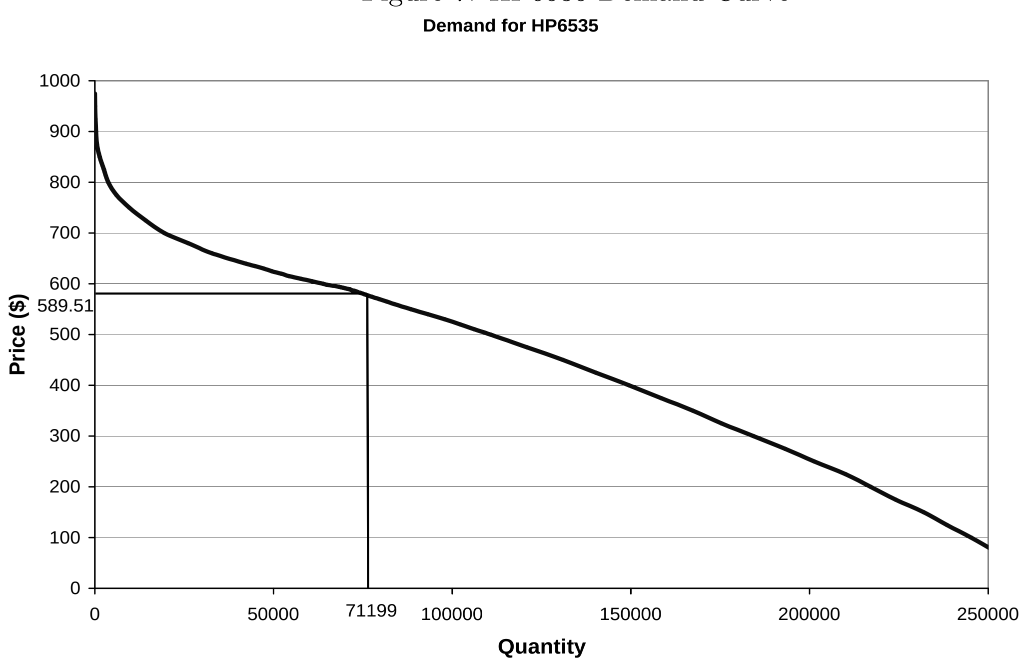
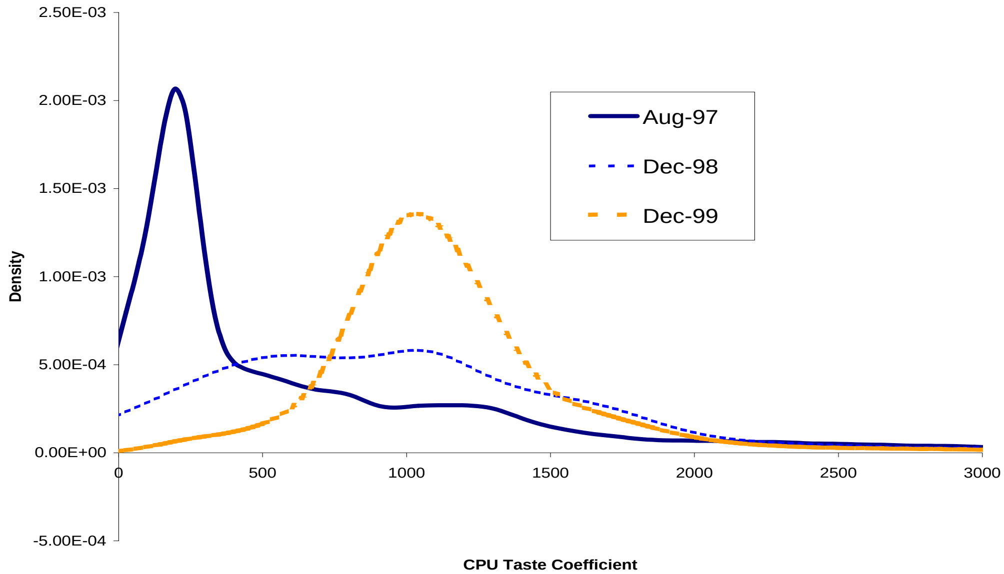


Figure 8: Taste for CPU Over Time



5 Dynamic Games.

- Final example is dynamic games with discrete and/or continuous choices.

- Examples include Aguerraguberia and Mira (2003), Pesendorfer and Schmidt-Dengler (2003), Berry, Ovstrovsky and Pakes (2003), and Bajari, Benkard and Levin (2003).

- The idea behind the estimation is that in a first step, you nonparametrically estimate agents policy functions.

- In a second step, you recover the parameters of the period return function consistent with the observed policies.

Notation.

- Assume discrete state space and discrete action space (for convenience only).

- Agents: $i = 1, ..., N$

- Time: $t = 1, ..., \infty$

- States: $\mathbf{s}_t \in S \subset R^G$, commonly known.

- Actions: $a_{it} \in A_i$, simultaneously chosen.

- Transitions: $P(\mathbf{s}_{t+1} | \mathbf{a}_t, \mathbf{s}_t)$.

- Discount Factor: β (known to econometrician).

Objective Function: Agent maximizes EDV,

$$\mathbb{E}_t \sum_{t=0}^{\infty} \beta^t u_i(\mathbf{a}_t, \mathbf{s}_t). \tag{1}$$

Equilibrium.

- Concept: Markov Perfect Equilibrium [MPE]

- Strategies: $\sigma_i : S \to A_i$.

- Recursive Formulation:

  $$V_i(\mathrm{s}|\boldsymbol{\sigma}) = u_i(\boldsymbol{\sigma}(\mathrm{s}), \mathrm{s}) + \beta \int V_i(\mathrm{s}'|\boldsymbol{\sigma}) dP(\mathrm{s}'|\sigma(\mathrm{s}), \mathrm{s})$$

- A MPE is given by a Markov profile, σ, such that for all i, s, σ_i'

$$V_i(\mathrm{s}|\sigma_i, \boldsymbol{\sigma}_{-i}) \geq V_i(\mathrm{s}|\sigma_i', \boldsymbol{\sigma}_{-i})$$

First Step.

- Estimate policy functions,

$$\sigma_i : S \to A_i$$

- and state transition function,

$$P : S \times A \to \Delta(S).$$

- Often will also estimate "static" parts of period return.

Examples:

- Production functions, (Olley-Pakes)
- Investment policies, (nonparametric)
- Entry/Exit policies, (nonparametric)
- Labor Supply,
- Static supply-demand system (BLP)
- State transitions: (parametric/nonparametric).

Second Step.

- Idea: Find the set of parameters that rationalize the data.

- I.e., conditional on P and σ, find the set of parameters that satisfy the requirements for equilibrium.

- Optimality Inequalities: For all i, σ_i', and initial state, s_0, it must be that

$$\mathbb{E}_{\sigma_i,\boldsymbol{\sigma}_{-i}|\mathbf{s}_0} \sum_{t=0}^{\infty} \beta^t u_i(\mathbf{a}_t, \mathbf{s}_t) \geq \mathbb{E}_{\sigma_i',\boldsymbol{\sigma}_{-i}|\mathbf{s}_0} \sum_{t=0}^{\infty} \beta^t u_i(\mathbf{a}_t, \mathbf{s}_t), \tag{3}$$

- The system of inequalities, (3), contains all information available from the definition of equilibrium.

- Assume: period return function is linear in the parameters
  (stronger than needed),

$$u_i(\mathbf{a}, \mathrm{s}; \boldsymbol{\theta}) = \boldsymbol{\Phi}_i(\mathbf{a}, \mathrm{s}) \cdot \boldsymbol{\theta}. \tag{4}$$

- Let

$$W(\mathrm{s}_0; \sigma_i, \boldsymbol{\sigma}_{-i}) = \mathbb{E}_{\sigma_i, \boldsymbol{\sigma}_{-i} | \mathrm{s}_0} \sum_{t=0}^{\infty} \beta^t \boldsymbol{\Phi}_i(\mathbf{a}_t, \mathrm{s}_t).$$

- Then the system (3) can be written as,

$$W(\mathrm{s}_0; \sigma_i, \boldsymbol{\sigma}_{-i}) \cdot \boldsymbol{\theta} \geq W(\mathrm{s}_0; \sigma_i', \boldsymbol{\sigma}_{-i}) \cdot \boldsymbol{\theta}, \tag{5}$$

- for all i, σ_i', s_0.

Identified Case.

- Let observed policy function be $\sigma = (\sigma_1, ..., \sigma_N)$.

- To simplify notation, abstract away from estimation error (easy to fix using standard theory of two-step estimators).

- Consider a finite set of alternative policies, σ_i' that agent i could have chosen, but did not.

- Define:

$$g(x,\theta) = \left[\widehat{W}(\mathrm{s};\sigma_i,\boldsymbol{\sigma}_{-i}) - \widehat{W}(\mathrm{s};\sigma_i',\boldsymbol{\sigma}_{-i})\right] \cdot \boldsymbol{\theta}$$

- Let n_I be the number of alternative (but not chosen) revealed preference inequalities that we consider.

- Abusing notation, let $g(X_k, \theta)$ denote the function g evaluated at a particular revealed preference inequality.

- Minimize:

$$Q_n(\theta) = \frac{1}{n} \sum_i 1 \{g(X_k, \theta) < 0\} \, g(X_k, \theta)^2$$

Comments:

- Two-step estimator is constructed by evaluating Q_n usiing first stage policy function estimates.

- Variances in denominator are known because they can be estimated very precisely using the simulation draws.

- Computationally light because simulation only needs to be done once, prior to maximizing likelihood.

- All second stage error comes from simulation.

Application to Dynamic Discrete Choice.

- Agent chooses single action, a, out of finite set, A.

- States include variables, s, observed by econometrician, and unobserved states, ϵ, representing stochastic shocks to preferences.

- Period utility received from choice j is,

$$u_{ij} = u(a_j, s; \theta) + \epsilon(a_j)$$

- ϵ is iid over time and individuals, and $N(0, V_\epsilon)$.

First Stage.

- Policy function is,

$$P(a_j|s) = Pr(v(a_j,s)+\varepsilon(a_j) \geq v(a_k,s)+\varepsilon(a_k), \text{for all } k) \tag{10}$$

- where $v(a_j, s)$ satisfies

$$v(a_j,s) = u(a_j,s)+ \\ \beta \int \int \max_{j'} \left[v(a_{j'}, s') + \varepsilon(a_{j'}) \right] P(d\varepsilon_{j'}) P(ds'|s, a = a_j). \tag{11}$$

- The policy function can be estimated by running a probit on a flexible function, $f^{(M)}(a, s; \theta)$. This also provides an estimate of V_ϵ. The estimated policy function is,

$$\sigma_i(s, \varepsilon(a)) = a_j, \text{ if} \\ \widehat{f}^{(M)}(a_j, s) + \varepsilon(a_j) \geq \widehat{f}^{(M)}(a_k, s) + \varepsilon(a_k) \text{ for all } k.$$

- Alternative policies are represented by alternative functions, $\widehat{f}^{(M)}$, e.g.,

$$\widehat{f}^{(M)} + \mu(a_j, s),$$

- where in the estimation $\mu(a_j, s)$ is iid from some distribution.

- Period Return Function:

$$u_{ij}(a_j, s; \theta) = \mathbf{\Phi}(a_j, s) \cdot \boldsymbol{\theta} + \varepsilon(a_j).$$

- Optimality Inequalities.

- For all s, σ_i'

$$\left[\widehat{\mathbb{E}}_{\sigma_i,s} \sum_{t=0}^{\infty} \beta^t \mathbf{\Phi}(a_j, s) - \widehat{\mathbb{E}}_{\sigma_i',s} \sum_{t=0}^{\infty} \beta^t \mathbf{\Phi}(a_j, s)\right] \cdot \boldsymbol{\theta} +$$
$$\left[\widehat{\mathbb{E}}_{\sigma_i,s} \sum_{t=0}^{\infty} \beta^t \varepsilon(a_t) - \widehat{\mathbb{E}}_{\sigma_i',s} \sum_{t=0}^{\infty} \beta^t \varepsilon(a_t)\right] \geq 0 \quad (2)$$

Sampling Inequalities:

- Randomly draw n_I starting states, s_0 from S.
- Randomly draw n_I alternative policies, σ_i'.
- Simulate paths of states by drawing unobserved states and executing policy functions. Use these simulated paths to simulate terms in brackets for each of the n_I inequalities sampled.

Simple Example, a la Rust '87

- State is age: $a_t \in \{1, 2, 3, 4, 5\}$

- Control is "replacement": $i_t \in \{0, 1\}$

- Machines start out at age 1. If machine is replaced, then it goes back to being age 1. If not, then it ages by one year (at age 5 it stays 5).

- Period payoff function is,

$$u(i_t, a_t) = \begin{cases} -\theta a_t + \epsilon_0 & \text{if } i_t = 0 \\ -R + \epsilon_1 & \text{if } i_t = 1 \end{cases}$$

  where ϵ_0 and ϵ_1 are logit errors and θ and R are parameters.

- True Values: $\theta = 1$, $R = 4$.

A Tables and Figures

Table 1: DDC Monte Carlo, 500 Monte Carlo runs, 25 subsamples of size $n/2$

	Mean	SE(Real)	5%(Real)	95%(Real)	SE(Subsampling)
$n = 400$, $n_I = 200$, $n_s = 1000$					
θ	1.00	0.14	0.79	1.24	0.10
R	4.02	0.53	3.24	4.96	0.39
$n = 200$, $n_I = 200$, $n_s = 500$					
θ	0.99	0.18	0.72	1.37	0.17
R	4.00	0.78	2.94	5.95	0.86
$n = 100$, $n_I = 200$, $n_s = 250$					
θ	0.94	0.32	0.47	1.48	0.35
R	3.75	1.26	1.92	5.70	1.15
$n = 50$, $n_I = 200$, $n_s = 150$					
θ	0.89	0.54	0.11	2.03	0.47
R	3.57	2.35	0.60	8.16	2.27

Dynamic Oligopoly w/ Investment

- Like Pakes and McGuire or Ericson and Pakes.

- Period return function

$$\pi_i(s, a, I) = \tilde{\pi}_i(s, a) - C(I_i) \qquad (13)$$

- s - vector of states, (some may be unobserved).

- a - actions that do not influence state transitions.

- I - actions that do influence state transitions.

- $C(I)$ - cost of investment function.

- Ψ - scrap value of firm.

- $F(x^e)$ - distribution of privately known entry cost.

(iii)
- s is capital, and an unobserved productivity shock.
- a is quantity.
- I is investment that increases capital.

First Stage.

- Estimate $\tilde{\pi}_i(s, a)$ and $a_i(s)$ using standard techniques.
  (BLP/Olley and Pakes/etc.)

- Estimate $P(s'|s, I)$
  (parametric or nonparametric – depends on model).

- Estimate investment function ($I_i(s)$), exit function ($\chi(s)$), and entry probabilities ($\chi^e(s)$).
  (typically nonparametric)

- Unobserved states must be recovered in first stage.

Second Stage.

- Assume that cost of investment function is linear in parameters:

$$C(I) = c \cdot \mathbf{\Phi}(I).$$

- For every initial state, s_0, and every alternative investment policy, $\sigma'(\mathbf{s}) = (I'(\mathbf{s}), \chi'(\mathbf{s}))$,

$$\left[\widehat{\mathbb{E}}_{\sigma_i,\sigma_{-i}} \sum_{t=0}^{\infty} \beta^t \tilde{\pi}_i(\mathbf{a}_t, \mathbf{s}_t) - \widehat{\mathbb{E}}_{\sigma_i',\sigma_{-i}} \sum_{t=0}^{\infty} \beta^t \tilde{\pi}_i(\mathbf{a}_t, \mathbf{s}_t)\right]$$

$$+c \cdot \left[\widehat{\mathbb{E}}_{\sigma_i,\sigma_{-i}} \sum_{t=0}^{\infty} \beta^t \mathbf{\Phi}(I_{it}) - \widehat{\mathbb{E}}_{\sigma_i',\sigma_{-i}} \sum_{t=0}^{\infty} \beta^t \mathbf{\Phi}(I_{it})\right]$$

$$+\mathbf{\Psi} \left[\begin{array}{c} \widehat{\mathbb{E}}_{\sigma_i,\sigma_{-i}} \sum_{t=0}^{\infty} \beta^t \{\chi(\mathbf{s}_t) = 1\} \\ -\widehat{\mathbb{E}}_{\sigma_i',\sigma_{-i}} \sum_{t=0}^{\infty} \beta^t \{\chi'(\mathbf{s}_t) = 1\} \end{array}\right] \geqslant 0$$

- Extremum estimator to estimate c and Ψ.

- Also straightforward to estimate sunk cost of entry distribution (parametrically or nonparametrically) – see paper for details.

Table 2: Dynamic Oligopoly Monte Carlo Parameters

Parameter	Value	Parameter	Value
Demand:		Investment Cost:	
α	1.5	c_1	1
γ	0.1		
M	5	Marginal Cost:	
y	6	mc	3
Investment Evolution		Entry Cost Distribution	
δ	0.7	x^l	7
a	1.25	x^h	11
Discount Factor		Scrap Value:	
β	0.925	Ψ	6

Table 3: $n_s = 2000$, 400 Monte Carlo runs, 20 subsamples of size $n/2$

	Mean	SE(Real)	5%(Real)	95%(Real)	SE(Subsampling)
$n = 400$, $n_I = 500$					
c_1	1.01	0.05	0.91	1.10	0.03
Ψ	5.38	0.43	4.70	6.06	0.39
$n = 200$, $n_I = 500$					
c_1	1.01	0.08	0.89	1.14	0.05
Ψ	5.32	0.56	4.45	6.33	0.53
$n = 100$, $n_I = 300$					
c_1	1.01	0.10	0.84	1.17	0.06
Ψ	5.30	0.72	4.15	6.48	0.72

Figure 1: Entry Cost Distribution for $n = 400$

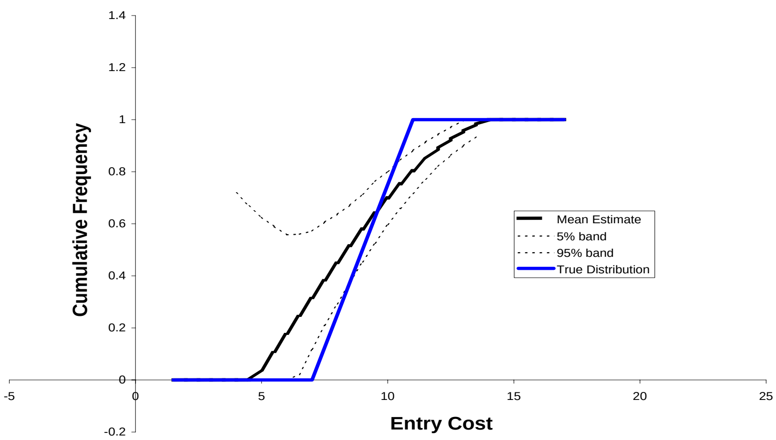


Figure 2: Entry Cost Distribution for $n = 200$

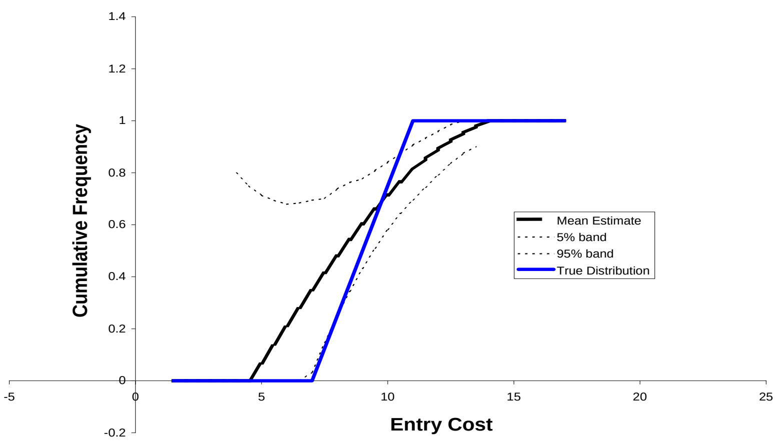


Figure 3: Entry Cost Distribution for $n = 100$