Estimation of Static and Dynamic Models of Strategic Interactions

http://www.econ.duke.edu/~ hanhong/ice06chicago.pdf

Han Hong

Department of Economics
Duke University

Presentation at ICE 06, Chicago, July, 2006

This talk is based on the following papers

- Patrick Bajari, Han Hong, John Krainer and Denis Nekipelov.
  Estimating Static Models of Strategic Interactions.
  Working paper, 2005.
- Patrick Bajari and Han Hong
  Semiparametric Estimation of a Dynamic Game.
  Working paper, 2005.
- Patrick Bajari, Han Hong and Stephen Ryan.
  Identification and Estimation of Discrete Games of Complete Information.
  Working paper, 2005.
- Victor Chernozhukov and Han Hong.
  An MCMC arpproach to classical estimation.
  Journal of Econometrics 115 (2003), pp293–346.

- Empirical analysis of games in econometrics and industrial organization.
- Discrete choice model with other agent's actions entering as a right hand side variable.
- Often most straight forward to estimate a game in two steps.
- In a first step, the economist estimates the reduced forms implied by the model.
- In the second step, recover structural utility parameters that rationalize the observed reduced forms.
- Computation of multiple equilibria.

- Early examples, Vuong and Bjorn (1984) and Bresnahan and Reiss (1990,1991).
- Also, Aradillas-Lopez (2005), Ho(2005), Ishii (2005), Pakes, Porter, Ho and Ishii (2005), Rysman (2004), Seim(2004), Sweeting (2004), Tamer (2003) and Ciliberto and Tamer (2005).
- Single agent dynamic models: Rust (1987), Hotz and Miller (1993), Magnac and Thesmar (2003), Heckman and Navarro (2005), among many others.
- Dynamic games: Aguirregabiria and Mira (2003), Bajari, Benkard and Levin (2004), Pakes, Ovstrovsky and Berry (2004), and Pesendorfer and Schmidt-Dengler (2004).

- "Estimating Static Models of Strategic Interactions"
- Players, $i = 1, ..., n$.
- Actions $a_i \in \{0, 1, \ldots, K\}$.
- $A = \{0, 1, \ldots, K\}^n$ and $a = (a_1, ..., a_n)$..
- $s_i \in S_i$: state for player i.
- $S = \Pi_i S_i$ and $s = (s_1, ..., s_n) \in S$.
- s is common knowledge and also observed by econometrician.
- For each agent i, $K + 1$ state variables $\epsilon_i(a_i)$
- $\epsilon_i\left(a_i\right)$: private information to each agent.
- $\epsilon_i = \left(\epsilon_i\left(0\right), \ldots, \epsilon_i\left(K\right)\right)$.
- Density $f(\epsilon_i)$, i.i.d. across $i = 1, \ldots, n$.

- Period utility for player i with action profile a:

$$u_i(a, s, \epsilon_i; \theta) = \Pi_i(a_i, a_{-i}, s; \theta) + \epsilon_i(a_i)$$

- Example: the period profit of firm i for entering the market.
- Generalizes a standard discrete choice model.
- Agents act in isolation in standard discrete choice models.
- Unlike a standard discrete choice model, a_{-i} enters utility.
- Player i's decision rule is a function $a_i = \delta_i(s, \epsilon_i)$.
- Note that ϵ_{-i} does not enter.
- ϵ_{-i} is private information of other players.

- Conditional choice probability $\sigma_i(a_i|s)$ for player i:

$$\sigma_i(a_i = k|s) = \int 1\left\{\delta_i(s, \epsilon_i) = k\right\} f(\epsilon_i) d\epsilon_i.$$

- Choice probability is conditional s: public information.
- Choice specific expected payoff for player i:

$$\Pi_i(a_i, s; \theta) = \sum_{a_{-i}} \Pi_i(a_i, a_{-i}, s; \theta)\sigma_{-i}(a_{-i}|s).$$

- Expected utility from choosing a_i, excluding preference shock.
- The optimal action for player i satisfies:

$$\sigma_i(a_i|s) = Prob\left\{ \begin{array}{c} \epsilon_i|\Pi_i(a_i, s; \theta) + \epsilon_i(a_i) \\ > \Pi_i(a_j, s; \theta) + \epsilon_i(a_j) \text{ for } j \neq i. \end{array} \right\}$$

- $\Pi_i(a_i, a_{-i}, s; \theta)$ is often a linear function, e.g.:

$$\Pi_i(a_i, a_{-i}, s) = \begin{cases} s' \cdot \beta + \delta \sum_{j \neq i} 1\{a_j = 1\} & \text{if } a_i = 1 \\ 0 & \text{if } a_i = 0 \end{cases}$$

- Mean utility from not entering normalized to zero.
- δ measures the influence of j's entry choice on i's profit.
- If firms compete with each other: $\delta < 0$.
- β measure the impact of the state variables on profits.
- $\varepsilon_i(a_i)$ capture shocks to the profitability of entry.
- Often $\varepsilon_i(a_i)$ are assumed to be i.i.d. extreme value distributed:

$$f(\varepsilon_i(k)) = e^{-\varepsilon_i(k)} e^{e^{-\varepsilon_i(k)}}.$$

- Choice specific expected payoff under linearity:

$$\Pi_i(a_i = 1, s; \theta) = s' \cdot \beta + \delta \sum_{j \neq i} \sigma_j(a_j = 1|s).$$

- Choice probability under the extreme value distribution

$$\sigma_i(a_i = 1|s) = \frac{\exp(s' \cdot \beta + \delta \sum_{j \neq i} \sigma_j(a_j = 1|s))}{1 + \exp(s' \cdot \beta + \delta \sum_{j \neq i} \sigma_j(a_j = 1|s))}$$

- Two approaches for estimation:
  - Full information maximum likelihood estimation.
  - Semiparametric two step estimation method.

- Choice probability equations define fixed point mappings for

$$\sigma_j(a_j = 1|s).$$

- For each $\theta = (\beta, \delta)$, solve for

$$\sigma_j(a_j = 1|s; \theta)$$

  for all $j = 1, \ldots, n$.
- Maximize likelihood function:

$$L(\theta) = \prod_{t=1}^{T} \prod_{i=1}^{n} \left(\sigma_i \left(a_i = 1|s; \theta\right)\right)^{1\{a_{i,t}=1\}} \left(1 - \sigma_i \left(a_i = 1|s; \theta\right)\right)^{1\{a_{i,t}=0\}}$$

- At the true parameter θ, $\sigma_j(a_j = 1|s;\theta))$ can be recovered from the data nonparametrically:

$$\hat{\sigma}_j(a_j = 1|s).$$

- Pseudo likelihood easy to maximize: (multinomial) logit

$$\prod_{t=1}^{T}\prod_{i=1}^{n}\left(\frac{\exp(s'\cdot\beta+\delta\sum_{j\neq i}\hat{\sigma}_j(a_j=1|s))}{1+\exp(s'\cdot\beta+\delta\sum_{j\neq i}\hat{\sigma}_j(a_j=1|s))}\right)^{1\{a_{i,t}=1\}}$$
$$\left(1-\frac{\exp(s'\cdot\beta+\delta\sum_{j\neq i}\hat{\sigma}_j(a_j=1|s))}{1+\exp(s'\cdot\beta+\delta\sum_{j\neq i}\hat{\sigma}_j(a_j=1|s))}\right)^{1\{a_{i,t}=0\}}$$

- Both the first stage estimates $\widehat{\sigma}_i(a_i = 1|s)$ and the term $s'\cdot\beta$ depend on the vector of state variables s.
- Colinearity and identification: Need a covariate that enters the first stage, but not the second stage.

A1 Assume that the error terms $\epsilon_i(a_i)$ are distributed i.i.d. across actions a_i and agents i, and come from a known parametric family.

- Not possible to allow nonparametric mean utility and error terms at once, even in simple single agent problems (e.g. a probit).
- In Bajari, Hong and Ryan (2005)- even a single agent model is not identified without an independence assumption.
- Well known that $\Pi_i\left(0, s\right)$ are not identified.
- $\sigma_i\left(a_i|s\right)$ only functions of $\Pi_i\left(a_i, s\right) - \Pi_i\left(0, s\right)$.
- Suppose $\epsilon_i(a_i)$ is extreme value,

$$\sigma_i(a_i|s) = \frac{\exp(\Pi_i(a_i, s) - \Pi_i(0, s))}{\sum_{k=0}^{K} \exp(\Pi_i(k, s) - \Pi_i(0, s))}$$

A2 For all i and all a_{-i} and s, $\Pi_i(a_i = 0, a_{-i}, s) = 0$.

- Can only learn choice specific value functions up to a first difference. Need normalization
- Similar to "outside good" assumption in single agent model.
- Entry: the utility from not entering is normalized to zero.

- Hotz and Miller (1993) inversion, for any k, k':

$$\log\left(\sigma_i(k|s)\right) - \log\left(\sigma_i(k'|s)\right) = \Pi_i(k,s) - \Pi_i(k',s).$$

- More generally let $\Gamma : \{0, ..., K\} \times S \rightarrow [0,1]$:

$$(\sigma_i(0|s), ..., \sigma_i(K|s)) = \Gamma_i\left(\Pi_i(1,s) - \Pi_i(0,s), ..., \Pi_i(K,s) - \Pi_i(0,s)\right)$$

- And the inverse Γ^{-1}:

$$(\Pi_i(1,s) - \Pi_i(0,s), ..., \Pi_i(K,s) - \Pi_i(0,s)) = \Gamma_i^{-1}\left(\sigma_i(0|s), ..., \sigma_i(K|s)\right)$$

- Invert equilibrium choice probabilities to nonparametrically recover $\Pi_i(1,s) - \Pi_i(0,s), ..., \Pi_i(K,s) - \Pi_i(0,s)$.
- $\Pi_i(a_i,s)$ is known by our inversion and probabilites σ_i can be observed by econometrician.

- Next step: how to recover $\Pi_i(a_i, a_{-i}, s)$ from $\Pi_i(a_i, s)$.
- Requires inversion of the following system:

$$\Pi_i(a_i, s) = \sum_{a_{-i}} \sigma_{-i}(a_{-i}|s)\Pi_i(a_i, a_{-i}, s),$$

$$\forall i = 1, \ldots, n, a_i = 1, \ldots, K..$$

- Given s, $n \times K \times (K+1)^{n-1}$ unknowns utilities of all agents.
- Only $n \times (K)$ known expected utilities.
- Obvious solution: impose exclusion restrictions.

- Partition $s = (s_i, s_{-i})$, and suppose

$$\Pi_i(a_i, a_{-i}, s) = \Pi_i(a_i, a_{-i}, s_i)$$

depends only on the subvector s_i.

$$\Pi_i(a_i, s_{-i}, s_i) = \sum_{a_{-i}} \sigma_{-i}(a_{-i}|s_{-i}, s_i)\Pi_i(a_i, a_{-i}, s_i).$$

- Identification: Given each s_i, the second moment matrix of the "regressors" $\sigma_{-i}(a_{-i}|s_{-i}, s_i)$,

$$E\sigma_{-i}(a_{-i}|s_{-i}, s_i)\sigma_{-i}(a_{-i}|s_{-i}, s_i)'$$

is nonsingular.
- Needs at least $(K+1)^{n-1}$ points in the support of the conditional distribution of s_{-i} given s_i.

- Step 1: Estimation of choice probabilities (e.g. sieve)

$$\widehat{\sigma}_i(k|s) = \sum_{t=1}^{T} 1(a_{it} = k) q^{\kappa(T)}(s_t)(Q_T' Q_T)^{-1} q^{\kappa(T)}(s).$$

where

$$Q_T = (q^{\kappa(T)}(s_1), \ldots, q^{\kappa(T)}(s_T),$$

and

$$q^{\kappa(T)}(s) = (q_1(s), \ldots, q_{\kappa(T)}(s)).$$

- Step 2: Inversion of expected utilities

$$\left(\widehat{\Pi}_i(1, s_t) - \widehat{\Pi}_i(0, s_t), \ldots \widehat{\Pi}_i(K, s_t) - \widehat{\Pi}_i(0, s_t)\right) = \Gamma_i^{-1} \left(\widehat{\sigma}_i(0|s_t), \ldots, \widehat{\sigma}_i(K|s_t)\right)$$

In the logit model

$$\hat{\Pi}_i(k, s_t) - \hat{\Pi}_i(0, s_t) = \log\left(\widehat{\sigma}_i(k|s_t)\right) - \log\left(\widehat{\sigma}_i(0|s_t)\right)$$

- Step 3: Recovering structural parameters.
- Infer $\Pi_i(a_i, a_{-i}, s_i)$ from $\hat{\Pi}_i(k, s)$.
- Use empirical analog of

$$\Pi_i(a_i, s_{-i}, s_i) = \sum_{a_{-i}} \sigma_{-i}(a_{-i}|s_{-i}, s_i)\Pi_i(a_i, a_{-i}, s_i).$$

- For given s_i and given $a = (a_i, a_{-i})$, run local regression

$$\sum_{t=1}^{T} \left(\hat{\Pi}_i(a_i, s_{-it}, s_i) - \sum_{a_{-i}} \hat{\sigma}_{-i}(a_{-i}|s_{-it}, s_i)\, \Pi_i(a_i, a_{-i}, s_i) \right)^2 w(t, s_i).$$

  Nonparametric weights can take different forms, e.g.

$$w(t, s_i) = k\left(\frac{s_{it} - s_i}{h}\right) / \sum_{\tau=1}^{T} k\left(\frac{s_{i\tau} - s_i}{h}\right).$$

- Asymptotically identified.

- Linear model of deterministic utility

$$\Pi_i\left(a_i, a_{-i}, s_i\right) = \Phi_i\left(a_i, a_{-i}, s_i\right)' \theta.$$

- Choice specific value function

$$\Pi_i\left(a_i, s\right) = E\left[\Pi_i\left(a_i, a_{-i}, s_i\right) | s, a_i\right] = \Phi_i\left(a_i, s\right)' \theta.$$

where

$$\Phi_i\left(a_i, s\right) = \sum_{a_{-i}} \Phi_i\left(a_i, a_{-i}, s_i\right) \prod_{j \neq i} \sigma(a_j = k_j | s).$$

- Estimated parameter choice probabilities:

$$\sigma_i\left(a_i | s, \hat{\Phi}, \theta\right) = \frac{\exp\left(\hat{\Phi}_i\left(a_i, s\right)' \theta\right)}{1 + \sum_{k=1}^{K} \exp\left(\hat{\Phi}_i\left(k, s\right)' \theta\right)}$$

- Semiparametric method of moment estimator

$$\frac{1}{T} \sum_{t=1}^{T} \hat{A}\left(s_t\right)\left(y_t - \sigma\left(s_t, \hat{\Phi}, \hat{\theta}\right)\right) = 0.$$

where

$$\sigma\left(s_t, \hat{\Phi}, \theta\right) = \left(\sigma_i\left(k | s_t, \hat{\Phi}, \theta\right), \forall k = 1, \ldots, K, \forall i = 1, \ldots, n\right)$$

- Nonparametric vs semiparametric methods
- Typically $\widehat{\sigma}_i(k|s)$ will converge to the true $\sigma_i(k|s)$ at a nonparametric rate which is slower than $T^{1/2}$.
- Alternative estimators: kernel smoothing or local polynomial
- Nonparametric method is more robust against misspecification.
- May suffer from curse of dimensionality.
- Semiparametric method more practical in applications.
- θ can be estimated at parametric rates despite first stage nonparametric estimation.
- Pseudo MLE can be used in place of Method of moments.

Linear probability model

- Binary choice model $K = 1$:

$$\Pi_i(0, a_{-i}, s) = \epsilon_i(0) \equiv 0, \quad , \Phi_i(1, a_{-i}, s; \theta) = \Phi_i^1(s_i)' \beta + \Phi_i^1(a_{-i}, s)' \gamma$$

- Action 1 is chosen if and only if

$$\Phi_i^1(s_i)' \beta + E\left[\Phi_i^2(a_{-i}, s) | s\right]' \gamma + \epsilon_i(1) > 0.$$

- Assume uniform distribution of $\epsilon_i(1)$:

$$a_i = \Phi_i^1(s_i)' \beta + E\left[\Phi_i^2(a_{-i}, s) | s\right]' \gamma + \eta_{it}(1),$$

where $E(\eta_{it}(1) | s_t) = 0$. Alternative representation:

$$a_i = \Phi_i^1(s_i)' \beta + \Phi_i^2(a_{-i}, s)' \gamma + E\left[\Phi_i^2(a_{-i}, s) | s\right]' \gamma - \Phi_i^2(a_{-i}, s)' \gamma + \eta_{it}(1),$$

- Valid instruments, function of s_t are mean independent of

$$E\left[\Phi_i^2(a_{-i}, s) | s\right]' \gamma - \Phi_i^2(a_{-i}, s)' \gamma + \eta_{it}(1),$$

- Can be estimated by 2SLS (ivreg in Stata)

- Smooth function of state variables s: $\alpha(a_i, s)$.
- Estimatable belief of players only a function of s.
- Nonparametric Identification:

$$\Pi_i(a_i, a_{-i}, s; \theta) = \alpha(a_i, s) + \tilde{\Pi}_i(a_i, a_{-i}, s_i; \theta).$$

- Expected utility nonparametric identified:

$$\Pi_i(a_i, s) = \quad \alpha(a_i, s) + \textstyle\sum_{a_{-i}} \sigma_{-i}(a_{-i}|s)\tilde{\Pi}_i(a_i, a_{-i}, s_i).$$

- Differencing between pairs of players

$$\Pi_i(k, s) - \Pi_j(k, s) = \sum_{a_{-i}} \sigma_{-i}(a_{-i}|s)\, \tilde{\Pi}_i(k, a_{-i}, s_i) - \sum_{a_{-j}} \sigma_{-j}(a_{-j}|s)\, \tilde{\Pi}_j(k, a_{-j}, s_j)$$

- Given s_i and s_j, use variations in $\sigma_{-i}(a_{-i}|s)$ and $\sigma_{-j}(a_{-j}|s)$ to identify $\tilde{\Pi}_i(k, a_{-i}, s_i)$ and $\tilde{\Pi}_j(k, a_{-j}, s_j)$.
- Under symmetry:

$$\Pi_i(k, s) - \Pi_j(k, s) = \sum_{a_{-ij}} \sigma_{-ij}(a_{-ij}|s) \left[\tilde{\Pi}_{ij}(k, a_{-ij}, s_i) - \tilde{\Pi}_{ij}(k, a_{-ij}, s_j)\right].$$

- Parametric mean utilities

$$\tilde{\Pi}_i\left(a_i, a_{-i}, s_i\right) = \Phi_i\left(a_i, a_{-i}, s_i\right)' \theta.$$

- Conditional logit estimation, $\log L$

$$\sum_{t=1}^{T}\left(\log\left[\exp\left(\theta' \sum_{i=1}^{n} a_{it}\hat{\Phi}_i\left(1, s_{it}\right)\right)\right] - \log\left[\sum_{d_t \in B_t} \exp\left(\theta' \sum_{i=1}^{n} d_{it}\hat{\Phi}_i\left(1, s_{it}\right)\right)\right]\right)$$

where

$$B_t \equiv \{d_t : \sum_{i=1}^{n} d_{it} = \sum_{i=1}^{n} a_{it}\}$$

- Conditional likelihood given the number of entrants.
- Rank estimation

$$\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j\neq i} 1\left(a_{it} > a_{jt}\right)\rho_-\left(\left(\hat{\Phi}_i\left(1, s_{it}\right) - \hat{\Phi}_j\left(1, s_{jt}\right)\right)'\theta\right).$$

- Examples of penalty functions

$$\rho_-\left(x\right) = 1\left(x < 0\right) \quad \text{or} \quad \rho_-\left(x\right) = 1\left(x < 0\right)x^2.$$

- Known distribution for the error term $F(\epsilon_i)$
- Known mean utility functions $\Pi_i(a_i, a_{-i}, \theta)$.
- Conditional choice probability fixed point mappings:

$$\sigma_i(a_i|s) = \Gamma_i\left(\sum_{-i} \sigma_{a_{-i}}(a_{-i}|s)\left[\Pi_i(k, a_{-i}, s; \theta) - \Pi_i(0, a_{-i}, s; \theta)\right], \forall k\right).$$

- Given linear mean utility, fixed point mappings:

$$\sigma_i(a_i|s) = \Gamma_i\left(\sum_{a_{-i}} \sigma_{-i}(a_{-i}|s)\Phi_i(a_i, a_{-i}, s)'\theta,\ a_i = 1, \ldots, K\right),\ i = 1, \ldots, n.$$

- $K \times n$ equations and $K \times n$ unknown variables

$$\sigma_i(a_i|s), \forall a_i = 1, \ldots, K, i = 1, \ldots, n,$$

- Possibility of multiple solutions.

- Find multiple and possibly all solutions
- The fixed point system, for $\sigma = \sigma(s)$:

$$\sigma - \Gamma(\sigma) = 0,$$

- Homotopy: linear mapping between two spaces of functions

$$H(\sigma, \tau) = \tau G(\sigma) + (1 - \tau)(\sigma - \Gamma(\sigma)), \quad \tau \in [0, 1],$$

- $H(\sigma, \tau)$ and $G(\sigma)$: vectors of $n \times K$ component functions

$$H_{i,a_i}(\sigma, \tau) \quad \text{and} \quad G_{i,a_i}(\sigma) \quad i = 1, \ldots, n \quad a_i = 1, \ldots, K.$$

- Initial system: $\tau = 1$, $H(\sigma, 0) = G(\sigma)$.
- End system: $\tau = 0$, $H(\sigma, 0) = \Gamma(\sigma)$.

- Typically initial system $G(\sigma)$ easy to solve.
- The solution path $\sigma(\tau)$:

$$H(\sigma(\tau), \tau) = 0.$$

- Differentiating this homotopy with respect to τ:

$$\frac{d}{\partial \tau} H(\sigma(\tau), \tau) = \frac{\partial H}{\partial \tau} + \frac{\partial H}{\partial \sigma} \cdot \frac{\partial \sigma}{\partial \tau}.$$

- Trace the path $\tau = 1$ to $\tau = 0$.
- Algorithms for numerically tracing this differential equation system.

Condition 1 (Regularity) *Let $\nabla(\tau)$ denote the Jacobian of the Homotopy functions with respect to σ at the solution path $\sigma(\tau)$:*

$$\nabla(\tau) = \frac{\partial}{\partial \sigma} Re\{H(\sigma, \tau)\}\Big|_{\sigma=\sigma(\tau)},$$

where $Re\{H(\sigma, \tau)\}$ denotes the real component of the homotopy functions. The jacobian $\sigma(\tau)$ has full rank for almost all τ.

Condition 2 (Path Finiteness) *Define $H^{-1}(\tau)$ to be the set of solutions $\sigma(\tau)$ to the homotopy system at τ. $H^{-1}(\tau)$ is bounded for all $0 < \tau \leq 1$. In other words, for all $\tau > 0$.*

$$\lim_{||\sigma|| \to \infty} H(\sigma, \tau) \neq 0.$$

- $\Gamma(\cdot)$ a linear function in a linear probability model.
- May not have multiple equilibria if argument to $\Gamma(\cdot)$ is linear in choice probabilities.
- For example if profit depends on number of competitors.
- But nonlinear interactions of choice probabilities also possible.
- Functional form of $\Gamma(\cdot)$ depends on distribution of error terms.
- In general, $\Gamma(\sigma)$ nonlinear function of a polynomial index of choice probabilities.
- Polynomial error distributions difficult to justify.
- Choice of initial system:

$$G_{i,a_i}(\sigma) = \sigma_i(a_i)^{q_{i,a_i}} - 1 = 0, \quad i = 1, \ldots, n, \quad a_i = 1, \ldots, K.$$

Theorem 2 *Define the sets* $H^{-1} = \{(\sigma_r, \sigma_i, \tau) \mid H(\sigma_r, \sigma_i, \tau) = 0\}$ *amd*

$$H^{-1}(\tau) = \{(\sigma_r, \sigma_i) \mid H(\sigma, \tau) = 0\} \quad \text{for} \quad \sigma_r \in \mathbb{R}^{nK}, \quad \text{and} \quad \sigma_i \in \mathbb{R}^{nK}.$$

Also define $\wp_\epsilon = \cup_{i,a_i}\{|\sigma_{r,i,a_i}| \leq \epsilon\}$ *to be the area around the imaginary axis.*

1) *The set* $H^{-1} \cap \{\mathbb{R}^{2nK} \setminus \wp_\epsilon \times [0,1]\}$ *consists of closed disjoint paths.*
2) *For any* $\tau \in (0,1]$ *there exists a bounded set such that* $H^{-1}(\tau) \cap \mathbb{R}^{2nK} \setminus \wp_\epsilon$ *is in that set.*
3) *For* $(\sigma_r, \sigma_i, \tau) \in H^{-1} \cap \{\mathbb{R}^{2nK} \setminus \wp_\epsilon \times [0,1]\}$ *the homotopy system allows parametrization* $H(\sigma_r(s), \sigma_i(s), \tau(s)) = 0$*. Moreover,* $\tau(s)$ *is a monotone function.*

Theorem 2 *For given* τ *one can pick the power* q_{i,a_i} *of the initial function (1) such that the homotopy system is regular and path finite outside any sequence of converging polyhedra* $\wp_\epsilon$*,* $\epsilon \longrightarrow 0$*.*

- Entry game with a small number of players
- Multiple equilibria computation, not about identification.
- Payoff to player i a linear function of the indicator of the rival's entry $(a_i = 1)$, market covariates and random term:

$$U_i\left(1, a_{-i}\right) = \theta_1 - \theta_2 \left(\sum_{j \neq i} 1\left(a_j = 1\right) \right) + \theta_3 x_1 + \theta_4 x_2 + \epsilon_i\left(a\right),$$
$$i = 1, \ldots, n.$$

- Symmetric model, ex-ante probability of entry:

$$P_i = \frac{e^{\theta_1 - \theta_2\left(\sum_{j \neq i} P_j\right) + \theta_3 x_1 + \theta_4 x_2}}{1 + e^{\theta_1 - \theta_2\left(\sum_{j \neq i} P_j\right) + \theta_3 x_1 + \theta_4 x_2}}$$

Table: Characteristics of the parameters

Parameter	Mean	Variance	Distribution
θ_1	2.45	1	Normal
θ_2	5.0	1	Normal
θ_3	1.0	1	Normal
θ_4	-1.0	1	Normal
x_1	1.0	0.33	Uniform
x_2	1.0	0.33	Uniform

Table: Results of Monte-Carlo Simulations.

Parameter	Mean	Std Dev	Max	Min
$n = 3$				
# of equilibria	1.592	1.175	7	1
P_1	0.366	0.362	0.998	0
P_2	0.360	0.367	0.995	0
P_3	0.363	0.348	0.993	0.003
$n = 4$				
# of equilibria	1.292	0.777	5	1
P_1	0.278	0.328	0.981	0.001
P_2	0.246	0.320	0.981	0.003
P_3	0.276	0.338	0.999	0.001
P_4	0.280	0.338	0.987	0.002
$n = 5$				
# of equilibria	1.106	0.505	5	1
P_1	0.104	0.201	0.964	0
P_2	0.138	0.252	0.975	0
P_3	0.315	0.338	0.992	0
P_4	0.356	0.385	0.983	0
P_5	0.319	0.344	0.982	0

- Build on previous work by Bajari and Krainer.
- Study the set of recommendations on a stock can be viewed as the outcome of a game.
- Payoffs may depend on recommendations made by peers since they are benchmarked against peer recommendations.
- Also, inherent indeterminacy in system of rankings.
- Focus on interesting behavior during run-up and subsequent collapse of the NASDAQ in 2000.
- Previous studies: least recommended stocks earned an average abnormal return of 13% in 2000-2001.
- Most highly recommended stocks earned average abnormal returns of -7%.

Four factors that could have influenced recommendations.

1. Recommendations must depend on public info about the future profitability of a firm (stock and time effects).
2. Second, analysts may have heterogeneous forecasts (merge earnings forecasts with recommendations).
3. Third, conflicts of interest (dummy variable for investment banking business in quarters before and after recommendation was made).
4. Finally, recommendations of other analysts matter.

Variables in data set

- RELATION-A dummy variable that is one if the analyst's brokerage engages in investment banking business with the company to which the recommendation applies.
- SPITDUM-A dummy variable that is equal to one after the quarter starting in June of 2001. Based on a comprehensive search of Wall Street Journal articles, this is when Elliot Spitzer began making very public criticisms of industry practices.
- IBANK-A dummy variable that is equal to one if the brokerage does any investment banking business with stocks in the NASDAQ 100.
- SBANK-the share of analysts that issued recommendations for a particular stock during a particular quarter where IBANK was one.

Table 1: Recommendation Variables.

Recommendation	*Numerical Value Recorded by I/B/E/S*
Strong Buy	1
Buy	2
Hold	3
Underperform	4
Sell	5

Table 2: Summary Statistics.

Variable	*Mean*	*Std.*	*Min.*	*Max.*	*Nobs*
Recommendation	2.210	0.9168	1	5	12719
Relation	0.0350	0.1839	0	1	12719
Ibank	0.8155	0.3878	0	1	12719
Earnings	0.1111	0.2439	-3.010	1.720	12719

Table 3: Tabulation of Recommendations by Quarter.

Variable/Time Period	*Q1 1998*	*Q1 2000*	*Q2 2003*
% Recs. Equal to 1	30.51	46.73	11.65
% Recs. Equal to 2	30.51	41.46	18.12
% Recs. Equal to 3	37.62	11.81	53.07
% Recs. Equal to 4	1.02	0.00	12.62
% Recs. Equal to 5	0.34	0.00	4.53

- Utility is modeled as an ordered logit.
- The strategic interaction is a best response to the expected recommendation of other analysts.
- Two sources of identification.
- The first is characteristics of other firms (e.g. whether or not they are IBANKS).
- The second is Elliot Spitzer.

- Reduced form ordered logit regression.
- High correlation between quarterly effects and market indexes.
- Conflict of Interest: coef on RELATION
- IBANK: general investment banking firms generally more conversative.
- Companies select banking firms favorable to them.
- Peer effects in the interaction model.
- Conformable peer effects important (1.8-2.3 IVBELIEF).
- Explains most of variation in the data.
- Parametric vs semiparametric first stage: similar results.

Table 6: Regression of Dummies on Market Indexes.

Variable	*Coefficient*	*Coefficient*
Constant	.8208896 (3.965)	0.6270 (4.3)
Nasdaq Index	-.0003467 (-4.960)	-
QQQ Price	-	-.007 (-4.7)
Nobs	21	18
R^2	0.48	0.6830

Table 7: Ordered Logit Estimates of the Effect of Conflicts of Interest.

Variable	*Coef.*	*Coef.*	*Coef.*	*Coef.*
RELATION	-.4675915	-.1662009	-.0592953	-.0789189
	(-7.20)	(-2.44)	(-0.80)	(-1.06)
IBANK	-	-	-	.3046066
				(4.63)
Log Likelihood	-16136.579	-15297.042	-14837.213	-14826.47
Pseudo-R^2	0.0016	0.0536	0.0820	0.0827
Fixed Effects	none	quarterly	quarterly, stock	quarterly, stock

In the ordered logit model, the dependent variable is the analyst's recommendation as coded by IBES. This takes on discrete values from one to five. In the table above, t-statistics are included in parentheses. We do not report ancillary parameters, such as the cut values and values of the fixed effects.

Table 8: Ordered Logit Estimates including Peer Effects (Parametric First Stage)

Variable	*Coef.*	*Coef.*	*Coef.*	*Coef.*
IVBELIEF	2.288576	2.288961	1.800282	1.96927
	(49.996)	(42.397)	(1.263)	(1.030)
RELATION	-	-	-	.0156405
				(0.16)
%DEV	-	-	-	.0048048
				(0.46)
Log Likelihood	-14842.233	-14837.558	-14836.574	-14836.453
Pseudo-R^2	0.0817	0.0820	0.0820	0.0820
Fixed Effects	none	stock	quarterly, stock	quarterly, stock

In the ordered logit model, the dependent variable is the analyst's recommendation as coded by IBES. This takes on discrete values from one to five. In the table above, t-statistics are included in parentheses (the t-statistic for the variable IVBELIEF is corrected using bootstrap). IVBELIEF is constructed from fitted values of first stage regression of average recommendation on covariates listed in section 6.4.3. Most of the quarterly and stock fixed effects are significant in the specifications that we study.

Table 9: Ordered Logit Estimates including Peer Effects (Semiparametric First Stage)

Variable	*Coef.*	*Coef.*	*Coef.*	*Coef.*
IVBELIEF	2.288268	2.28852	1.845005	1.914881
	(44.585)	(39.301)	(2.023)	(1.535)
RELATION	-	-	-	.0133295
				(0.16)
%DEV	-	-	-	.0041601
				(0.40)
Log Likelihood	-14841.411	-14836.739	-14835.71	-14835.616
Pseudo-R^2	0.0817	0.0820	0.0821	0.0821
Fixed Effects	none	stock	quarterly, stock	quarterly, stock

In the ordered logit model, the dependent variable is the analyst's recommendation as coded by IBES. This takes on discrete values from one to five. In the table above, t-statistics are included in parentheses (the t-statistic for the variable IVBELIEF is corrected using bootstrap). IVBELIEF is constructed from Semiparametric sieve estimator of average recommendation on covariates listed in section 6.4.3. Most of the quarterly and stock fixed effects are significant in the specifications that we study.

- Focus on two player case.
- Quarter 9 and quarter 21.
- Average stock dummies and %DEV.
- Findings: two equilibria pre-Spitzer, one post-Spitzer.

Table 6: Computed choice probabilities for choices 1-4.

$RELATION = 0$	(0.74876E-04; 0.37660E-03; 0.66071E-02; 0.17543E-01)
Quarter 9	(0.21574E-01; 0.95336E-01; 0.55826E-0; 0.20545E-0)
$RELATION = 1$	(0.16503E-03; 0.82499E-03; 0.14331E-01; 0.37016E-01)
Quarter 9	(0.21124E-01; 0.93580E-01; 0.55573E-0; 0.207916E-0)
$RELATION = 0$	(0.12220E-03; 0.61102E-03; 0.10656E-01; 0.27896E-01)
Quarter 21	
$RELATION = 1$	(0.12005E-03; 0.60027E-03; 0.10471E-01; 0.27429E-01)
Quarter 21	

- Players are forward looking.
- Infinite Horizon, Stationary, Markov Transition
- Now players maximize expected discounted utility using discount factor β.

$$W_i(s,\epsilon_i;\sigma) = \max_{a_i \in A_i} \Big\{ \Pi_i(a_i,s) + \epsilon_i\,(a_i) + \beta \int \sum_{a_{-i}} W_i(s',\epsilon_i';\sigma) g(s'|s,a_i,a_{-i}) \sigma_{-i}(a_{-i}|s) f(\epsilon_i') d\epsilon_i' \Big\}$$

- Definition: A Markov Perfect Equilibrium is a collection of $\delta_i(s,\epsilon_i)$, $i = 1,...,n$ such that for all i, all s and all ϵ_i, $\delta_i(s,\epsilon_i)$ maximizes $W_i(s,\epsilon_i;\sigma_i,\sigma_{-i})$.

- Conditional independence:
  - ϵ distributed i.i.d. over time.
  - State variables evolve according to $g\left(s'|s, a_i, a_{-i}\right)$.
- Define choice specific value function

$$V_i(a_i, s) = \Pi_i(a_i, s) + \beta E\left[V_i(s')|s, a_i\right].$$

- Players chooose a_i to maximize $V_i\left(a_i, s\right) + \epsilon_i\left(a_i\right)$,
- Ex ante value function (Social surplus function)

$$\begin{aligned}
V_i\left(s\right) =& E_{\epsilon_i} \max_{a_i} \left[V_i(a_i, s) + \epsilon_i\left(a_i\right)\right] \\
=& G\left(V_i(a_i, s), \forall a_i = 0, \ldots, K\right) \\
=& G\left(V_i(a_i, s) - V_i\left(0, s\right), \forall a_i = 1, \ldots, K\right) + V_i\left(0, s\right)
\end{aligned}$$

- When the error terms are extreme value distributed

$$
\begin{aligned}
V_i(s) &= \log \sum_{k=0}^{K} \exp\left(V_i(k,s)\right) \\
&= \log \sum_{k=0}^{K} \exp\left(V_i(k,s) - V_i(0,s)\right) + V_i(0,s).
\end{aligned}
$$

- Relationship between $\Pi_i(a_i, s)$ and $V_i(a_i, s)$:

$$
\begin{aligned}
V_i(a_i,s) =& \Pi_i(a_i,s) + \beta E\left[G\left(V_i(a_i,s'), \forall a_i = 0,\ldots,K\right)|s,a_i\right] \\
=& \Pi_i(a_i,s) + \beta E\left[G\left(V_i(k,s') - V_i\left(0,s'\right), \forall k = 1,\ldots,K\right)|s,a_i\right] \\
& + \beta E\left[V_i\left(0,s'\right)|s,a_i\right]
\end{aligned}
$$

- With extreme value distributed error terms

$$
\begin{aligned}
V_i(a_i,s) =& \Pi_i(a_i,s) + \beta E\left[\log \sum_{k=0}^{K} \exp\left(V_i\left(k,s'\right) - V_i\left(0,s'\right)\right)|s,a_i\right] \\
& + \beta E\left[V_i\left(0,s'\right)|s,a_i\right]
\end{aligned}
$$

- Hotz and Miller (1993): one to one mapping between $\sigma_i\left(a_i|s\right)$ and differences in choice specific value functions:

$$\left(V_i(1,s)-V_i(0,s),\dots V_i(K,s)-V_i(0,s)\right)=\Omega_i\left(\sigma_i(0|s),\dots,\sigma_i(K|s)\right)$$

- Example: i.i.d extreme value $f(\epsilon_i)$:

$$\sigma_i(a_i|s)=\frac{\exp(V_i(a_i,s)-V_i(0,s))}{\sum_{k=0}^{K}\exp(V_i(k,s)-V_i(0,s))}$$

- Inverse mapping:

$$\log\left(\sigma_i(k|s)\right)-\log\left(\sigma_i(0|s)\right)=V_i(k,s)-V_i(0,s)$$

- Since we can recover $V_i\left(k,s\right)-V_i\left(0,s\right)$, we only need to know $V_i\left(0,s\right)$ to recover $V_i\left(k,s\right),\forall k$.
- If we know $V_i\left(0,s\right)$, $V_i\left(a_i,s\right)$ and $\Pi_i\left(a_i,s\right)$ is one to one.

- Identify $V_i(0,s)$ first. Set $a_i = 0$:

$$V_i(0,s) = \Pi_i(0,s) + \beta E\left[\log \sum_{k=0}^{K} \exp\left(V_i(k,s') - V_i(0,s')\right)|s,0\right] + \beta E\left[V_i(0,s')\,|s,0\right]$$

- This is a single contraction mapping unique fixed point iteration.
- Add $V_i(0,s)$ to $V_i(k,s) - V_i(0,s)$ to identify all $V_i(k,s)$.
- Then all $\Pi_i(k,s)$ calculated from $V_i(k,s)$ through

$$\Pi_i(k,s) = V_i(k,s) - \beta E\left[V_i(s')\,|s,k\right].$$

- Why normalize $\Pi_i(0, s) = 0$?
- Why not $V_i(0, s) = 0$?
- If a firm stays out of the market in period t, current profit 0, but option value of future entry might depend on market size, number of other firms, etc.
- These state variables might evolve stochastically.
- Rest of the identification arguments: identical to the static model.

- Nonparametric and Semiparametric Estimation
- Hotz-Miller inversion recovers $V_i(k,s) - V_i(0,s)$ instead of $\Pi_i(k,s) - \Pi_i(0,s)$.
- Nonparametrically compute $V_i(0,s)$ using

$$\begin{aligned}\hat{V}_i(0,s) =& \beta \hat{E}\left[\log \sum_{k=0}^{K} \exp\left(\hat{V}_i(k,s') - \hat{V}_i(0,s')\right)|s,0\right] \\ &+ \beta \hat{E}\left[\hat{V}_i(0,s')\,|s,0\right]\end{aligned}$$

- Obtain and $\hat{V}_i(k,s)$ and forward compute $\hat{\Pi}_i(k,s)$.
- The rest is identical to the static model.

- In semiparametric models, $\hat{\theta}$ converges at a $T^{1/2}$ rate and has normal asymptotics.
- Apply the results of Newey (1994)-derive appropriate "influence functions".
- The asymptotic distribution is invariant to the choice of method used to estimate the first stage.
- With proper weighting function (need to estimate nonparametrically), can achieve the same efficiency as full information maximum likelihoood.
- These results hold for both static and dynamic models.

- Static and dynamic interaction models.
- Incomplete information assumption.
- Nonparametric identification.
- Nonparametric and Semiparametric estimation.
- Need for computation of multiple equilibria.
- Equilibria computation important for model simulation.
- Extension: parametric estimation method that allows for multiple equilibria.

Patrick Bajari, Han Hong and Stephen Ryan.
Identification and Estimation of Discrete Games of Complete Information.
Working paper, 2005.

- Preference shocks common knowledge- much harder.
- Important previous work includes original works of Bresnahan and Reiss (1990,1991), Berry (1992), and recent papers by Tamer (2002), and Ciliberto and Tamer (2003).
- Issues with existence and multiplicity of equilibrium.

- Work with general games instead of entry games.
- Allow for mixed strategies.
- Equilibrium selection is part of model.
- Exploit computational tools in McKelvy and McLennan (1994) to find all equilibrium.
- Reduce computational burden in structural models using importance sampling.
- Study nonparametric identification.

- Previous papers establish negative results on identification.
- Since our model is more general (allowing for selection of equilibrium) we also get negative results.
- Positive results with exclusion restrictions.
- Find variables that influence selection equation (which equilibrium is played) but not treatment (latent utility).
- Laws, regulations, previous plays may generate exclusion restrictions.

- Simultaneous move game of complete information (normal form game).
- There are $i = 1, ..., N$ players with a finite set of actions A_i.
- Let $A = \prod_i A_i$.
- Utility is $u_i : A \rightarrow R$, where R is the real line.
- Let π_i denote a mixed strategy over A_i.
- A Nash equilibrium is a set of best responses.

- Following Bresnahan and Reiss (1990,1991), utility is:

$$u_i(a) = f_i(x, a; \theta) + \varepsilon_i(a).$$

- Mean utility, $f_i(x, a; \theta_1)$
- a, the vector of actions, covariates x, and a parameters θ.
- $\varepsilon_i(a)$ preference shocks.
- $\varepsilon_i(a) \sim g(\varepsilon|\theta)$ iid.
- Standard random utility model, except utility depends on actions of others.
- $E(u)$ set of Nash equilibrium.
- $\lambda(\pi; E(u), \beta)$ is probability of equilibrium, $\pi \in E(u)$ given parameters β.
- $\lambda(\pi; E(u), \beta)$ corresponds to a finite vector of probabilities.

- Example of λ. Theorists have suggested that an equilibrium may be more likely to be played if it:
  - Satisfies a particular refinement concept (e.g. trembling hand perfection).
  - The equilibrium is in pure strategies.
  - The equilibrium is risk dominant.
- Create dummy variable for whether a given equilibrium, $\pi \in E(u)$ satisfies criteria 1-3 above.
- Let $x(\pi, u)$ denote the vector of covariates that we generate in this fashion.
- Then a straightforward way to model λ is:

$$\lambda(\pi; E(u), \beta) = \frac{\exp(\beta \cdot x(\pi, u))}{\sum_{\pi' \in E(u)} \exp(\beta \cdot x(\pi', u))}$$

- Computing the set $E(u)$, all of the equilibrium to a normal form game, is a well understood problem.
- McKelvy and McLennan (1996) survey the available algorithms in detail.
- Software package Gambit.
- $P(a|x,\theta,\beta)$ is probability of a given x, θ and β

$$P(a|x,\theta,\beta) = \int \left\{ \sum_{\pi \in E(u(x,\theta,\varepsilon))} \lambda(\pi; u(x,\theta_1,\varepsilon),\beta) \left(\prod_{i=1}^{N} \pi(a_i) \right) \right\} g(\varepsilon|\theta_2) d\varepsilon$$

- Computation of the above integral is facilitated by the importance sampling procedure. (cf. Ackerberg (2003))

- Often $g(\varepsilon|\theta_2)$ is a simple parametric distribution (e.g. normal). For instance, suppose it is normal and let $\phi(\cdot|\mu, \sigma)$ denote the normal density.
- Then, the density $h(u|\theta, x)$ for the vNM utilities u is:

$$h(u|\theta, x) = \prod_{i} \prod_{a \in A} \phi(\varepsilon_i(a); f_i(\theta, x, \theta) + \mu, \sigma)$$

where for all i and all a, $\varepsilon_i(a) = f_i(x, a; \theta_1) - u_i(a)$

- Evaluating $h(u|\theta, x)$ is not difficult.
- Draw $s = 1, ..., S$ vectors of vNM utilities, $u^{(s)} = (u_1^{(s)}, ..., u_N^{(s)})$ from an importance density $q(u)$.

- We can then simulate $P(a|x,\theta,\beta)$ as follows:

$$\widehat{P}(a|x,\theta,\beta) = \sum_{s=1}^{S} \left\{ \sum_{\pi \in E(u)} \lambda(\pi; E(u^{(s)}), \beta) \left(\prod_{i=1}^{N} \pi(a_i) \right) \right\} \frac{h(u^{(s)}|\theta,x)}{q(u^{(s)})}$$

- Precompute $E(u^{(s)})$ for a large number of randomly drawn games $s = 1, ..., S$.
- Evaluating $\widehat{P}(a|x,\theta,\beta)$ at new parameters does not require recomputing $E(u^{(s)})$ for new $s = 1, ..., S$!
- Evaluating simulation estimator of $\widehat{P}(a|x,\theta,\beta)$ of $P(a|x,\theta,\beta)$ only requires "reweighting" of the equilibrium by new λ and $\frac{h(u^{(s)}|\theta,x)}{q(u^{(s)})}$.
- This is a feasible computation.

- Normally, the computational expense of structural estimation comes from recomputing the equilibrium many times.
- This can save on the computational time by orders of magnitude.
- Given $\widehat{P}(a|x, \theta, \beta)$ we can simulate the likelihood function or simulate the moments.
- The asymptotics are standard.
- Simulated method of moments is unbiased and might have lower number of simulation draws required to converge to the asymptotic distribution at $T^{1/2}$ rate.
- The model we propose is parametric.
- We want to see if identification hinges on parametric assumptions.
- Main result: To identify with weak functional form assumptions, need exclusion restrictions.

- Three firms who must decide whether or not to enter a single market.
- Two actions, $a_i = 0$ if they do not enter and $a_i = 1$ if they do.
- The payoffs from entering the market:

$$u_i(a_i = 1) = \theta_1 - \theta_2(\sum_{j \neq i} 1(a_j = 1)) + \theta_3 x_1 + \theta_4 x_2 + \varepsilon_i(a),$$

- Payoff staying out of market

$$u_i(a_i = 0) = \varepsilon_i(a).$$

- Logit probability of selecting a given equilibrium:

$$\lambda(\pi_i; E(u), \beta) = \frac{\exp(\beta_1 \text{MIXED}_i)}{\sum_{\pi' \in E(u)} \exp(\beta_1 \text{MIXED}_i)},$$

  MIXED_i is a dummy variable indicating mixed strategy equilibrium.
- True parameters $(\theta_1, \theta_2, \theta_3, \theta_4, \beta_1)$ are (5, 1.5, 1.0, -1.0, 1.0).

Table 1: Monte Carlo Results.

Parameter	*Mean*	*Median*	*IQR*	*Confidence Interval*	*Std. Dev*
T=25					
θ_1	5.0342	5.0404	0.21418	[4.5684, 5.5686]	0.26308
θ_2	1.4877	1.4675	0.23879	[1.0771, 1.9592]	0.26361
θ_3	1.1283	1.0394	0.19277	[0.72062, 1.9593]	0.36912
θ_4	-1.0711	-1.0244	0.04973	[-1.2794, -0.97999]	0.19687
β	1.0330	0.99922	0.01356	[0.92224, 1.1125]	0.33465
T=50					
θ_1	5.0218	5.0288	0.08552	[4.6863, 5.3162]	0.20485
θ_2	1.4794	1.4752	0.16211	[1.1823, 1.8771]	0.21398
θ_3	1.0322	1.0176	0.06815	[0.82246, 1.3440]	0.19895
θ_4	-1.0293	-1.0116	0.02993	[1.0837, 0.98737]	0.08357
β	1.0172	1.0000	0.00448	[0.95352, 1.0168]	0.23988
T=100					
θ_1	5.0323	5.0224	0.05724	[4.8738, 5.2040]	0.12690
θ_2	1.4938	1.4766	0.08945	[1.3246, 1.7082]	0.14590
θ_3	1.0253	1.0165	0.03221	[0.92628, 1.2529]	0.17947
θ_4	-1.0221	-1.0112	0.01746	[-1.0594, -0.99170]	0.058417
β	0.99637	1.0001	0.00100	[0.96680, 1.0083]	0.017842

Monte Carlo results of the estimator on simulated data sets of N=25, 50, and 100. The estimation was repeated 100 times for each data set size. The true parameters ($\theta_1,\theta_2,\theta_3,\theta_4,\beta$) are (5, 1.5, 1.0, -1.0, 1.0).

A1. For every i and $a_{-i} \in A_{-i}$, we let $f_i(\underline{a_i}, a_{-i}, x) = 0$ for some chosen $\underline{a_i} \in A_i$ and for all $a_{-i} \in A_{-i}$.

A2. For every i and for every a, $\varepsilon_i(a)$ are distributed i.i.d. standard normal.

A3. λ does not depend on the stochastic preference shocks ε.

A4. Given x, λ only depends only on the support of the elements in $\mathcal{E}(u)$.

- The observable moments of conditional probabilities $P\left(a|x\right)$ are determined by the mean utilities and the selection mechanism. Denote this mapping as $P\left(a|x\right) = H\left(f\left(x\right), \lambda\left(x\right)\right)$.

A5. The map H is continuously differentiable. Also suppose that for all x, the Jacobian matrix $DH_{f,\lambda}$ has rank equal to the dimension of the parmeter vector $\left(f(a,x), \lambda(x)\right)$.

A6. λ is a function of x and z, where z can be excluded from f_i for all i.

	L	R
T	$(\varepsilon_1(TL), \varepsilon_2(TL))$	$(\varepsilon_1(TR), f_2(TR, x) + \varepsilon_2(TR))$
B	$(f_1(BL, x) + \varepsilon_1(BL), \varepsilon_2(BL))$	$(f_1(BR, x) + \varepsilon_1(BR), f_2(BR, x) + \varepsilon_2(BR))$

Lemma 1. (2 by 2 Equilibrium). With probability one, the set of equilibrium is either unique or has three elements. If it has three elements (*i*) One equilibrium is in mixed strategies and (*ii*) In the two pure strategy equilibrium, no player plays the same strategy in both equilibria.

Lemma 2. Given A1-A4, λ can be characterized by a finite dimensional vector of parameters holding x fixed.

- Equilibrium selection probabilities: $\lambda_1(x), ..., \lambda_4(x)$.
- If the equilibrium set is $\{(T, L), (B, R)$, a mixed strategy equilibrium$\}$, select (T, L) with probability $\lambda_1(x)$, (B, R) with probability $\lambda_2(x)$ and the mixed strategy equilibrium with probability $1 - \lambda_1(x) - \lambda_2(x)$.
- If the equilibrium set is $\{(T, R), (B, L)$, a mixed strategy equilibrium$\}$, select (T, R) with probability $\lambda_3(x)$, (B, L) with probability $\lambda_4(x)$ and the mixed strategy equilibrium with probability $1 - \lambda_3(x) - \lambda_4(x)$.

Theorem

In a game with two players and two strategies, if we make assumptions A1-A5, the deterministic utility components

$$f_1(BL, x), f_1(BR, x), f_2(TR, x), f_2(BR, x)$$

are not identified from the distribution of $P(a|x)$ even if the selection mechanism

$$\lambda_1(x), \ldots, \lambda_4(x)$$

is known.

Proof: To begin with, consider the identification problem holding a given realization of x fixed. Since there are two players with two strategies, the econometrician observes four conditional moments,

$$P(TL|x),\ P(TR|x),\ P(BL|x),\ \text{and } P(BR|x),$$

Since the probability of the actions must sum to one, there are effectively three moments that the econometrician observes. This leaves us with 4 utility parameters,

$$f_1(BL,x), f_1(BR,x), f_2(TR,x), f_2(BR,x)$$

to identify. Clearly, for a given realization of x we are not identified. Q.E.D.

Theorem

In a game with more than two players and at least two strategies per player, if we make assumptions A1-A5, the deterministic utility parameters $f_i(a, x)$ are not identified from the distribution of $P(a|x)$, even if the selection mechanism $\lambda(\cdot)$ is known.

Proof: Consider a game with N players and $\#A_i$ strategies for player i. Holding x fixed, the total number of mean utility parameters $f_i(a, x)$ is equal to

$$N \cdot \prod_i \#A_i - \sum_i \prod_{j \neq i} \#A_j.$$

This is equal to the cardinality of the number of strategies, times the number of players, minus the normalizations allowed by assumption A1. The number of moments that the economist can observe, conditional on x, is only equal to

$$\prod_i \#A_i - 1.$$

If each player has at least two strategies and if there are at least 2 players in the game, then for each given x the difference between the number of utility parameters, f_i, to estimate and the number of available moment conditions is bounded from below by

$$\left((N-1) - \frac{N}{2}\right) \prod_i \#A_i + 1 \geq 0.$$

A6. λ is a function of x and z, where z can be excluded from f_i for all i.

A7. The covariates, x, can be partitioned into

$$x = (x_\lambda, x_u)$$

such that $\lambda(x, z)$ depends only on x_λ and not x_u:

$$\lambda(x, z) = \lambda(x_\lambda, z).$$

Theorem

In the two by two game, suppose that A1-A7 are satisfied. Also suppose that (x_λ, x_u, z) takes on a discrete number of values and $\#x_u > 3$ and $\#z > 3$. Then the mean utilities $f_i(a, x)$ and the selection parameters $\lambda(x_\lambda, z)$ are locally identified.

Proof

The number of moments generated by observable population conditional outcome probabilities, $P(a|x_\lambda, x_u, z)$ is

$$3 \times (\#x_\lambda) \times (\#x_u) \times (\#z)\,, \tag{1}$$

Note that, in this equation, we multiply by 3 because the probabilities of the various actions must sum to one. The total number of parameters needed to characterize both the utility functions and the equilibrium selection probabilities is

$$4 \times (\#x_\lambda) \times (\#x_u) + 4 \times (\#x_\lambda) \times (\#z)\,. \tag{2}$$

Alternatively, we can think for each each given x_λ, there are

$$3 \times (\#x_u) \times (\#z)\,, \tag{3}$$

conditional outcome probabilities and there are

$$4 \times (\#x_u) + 4 \times (\#z)\,, \tag{4}$$

parameters to estimate. It is clear that as long as $\#x_u > 3$ and $\#z > 3$,

$$3 \times (\#x_u) \times (\#z) > 4 \times (\#x_u) + 4 \times (\#z) \tag{5}$$

and the model is locally identified by the implicit function theorem. Q.E.D.

Theorem

In a general N player game, suppose that A1-A7 are satisfied. Also suppose that (x_λ, x_u, z) takes on a discrete number of values. If $\#x_u$ and $\#z$ are sufficiently large, the model is locally identified.

Proof. Holding x_λ and z fixed, by Lemma 2, it must be possible to characterize λ with a finite dimensional parameter vector. Since this vector depends on the supports of the elements in $\mathcal{E}(u)$, it is possible to create a bound on the size of this vector that is independent of x_λ and z. Let $\#\mathcal{E}$ denote this number. Holding x_u fixed, the number of vNM utilities is equal to $N \cdot \prod_i \#A_i - \sum_i \prod_{j \neq i} \#A_j$.

Then the number of parameters is bounded by above by:

$$\#\mathcal{E}\,(\#z) + \left(N \cdot \prod_i \#A_i - \sum_i \prod_{j \neq i} \#A_j \right) (\#x_u)$$

The number of moments is proportional to

$$(\#A - 1) \times (\#x_u) \times (\#z)\,.$$

The number of moments grows at a rate involving the product of $(\#x_u) \times (\#z)$ while the number of parameters is a linear combination of these terms. For sufficiently large $\#x_u$ and $\#z$, the number of moments is greater than the number of parameters.

A8. For each agent i, there exists some covariate, x_i that enters the utility of agent i, but not the utility of other agents. That is, i's utility can be written as $f_i(a, x, x_i)$. Furthermore, x_i can be excluded from λ.

Theorem

Suppose that A1-A5 and A8 hold. If $\#x_i$ are sufficiently large, the model is nonparametrically (locally) identified.

Proof: The proof follows similarly to the previous section. Hold x fixed. Consider a large, but finite number of values of x_i equal to K for each agent. Consider the all the distinct vectors of the form $x = (x_1, ..., x_N)$ that can be formed. The number will be equal to K^N. Consider the moments generated by these K^N distinct covariates. The number of moments is equal to $K^N \cdot \sum_i (\#A_i - 1)$. The number of mean utility parameters is equal to $\sum_i K(\#A_i - 1) \prod_{j \neq i} \#A$ plus the number of parameters required to characterize λ (which is independent of the x_i). Thus, the number of moments depends linearly on K but the number of moments grows exponentially with K.

Application 1

- Small business decision to go online.
- Focus on a single industry.
- Golf course in the Carolinas.
- Owner-operated, compete in spatially separated markets.
- Small business as marginal internet adopters.
- Slower and more varied than larger firms.
- Netwok dynamics in internet adoption.
- Long evolution of the internet.
- Web sites that aggregate information.

- Strategic considerations of web site creation.
- Negative demand consequences if competitors adopt.
- Negative supply side network effect even in absence of competition.
- Model the decision of internet technology adoption as a strategic game.
- Low number of firms deciding on a discrete action.
- Information aggregators in the online golf market.
- generic listing: location, rating, yardage, slope, price.
- Some golf courses maintain own web page.
- Largely informational.

- North Carolina and South Carolina
- variables:
  - Course type (public, private, resort, military) and Location
  - Have web site or not
  - Number of holes, rating, slope,
  - weekday and weekend prices
  - Local population, median rent, median house value, median household income.
  - Internet golf guides for course characteristics
  - Census sources for housing and income information.
- Market definitions critical: chaining all courses within 10 miles of any other course in the market.
- Only use markets with five or fewer courses.

Table 3: Website Adoption Rates for Golf Courses in Carolinas.

Number of Firms in Market	*Adoption Rate*
1	15%
2	19%
3	16%
4	45%
5	27%

- Are adoption decisions by other firms strategic compliments or substitutes?
- What type of equilibrium is most likely to be played (efficient, mixed strategies)?

$$\lambda_i = \frac{\exp\left(\beta_1 MIXED_i + \beta_2\left(\pi_i - \pi_{eff}\right)\right)}{\sum \exp\left(\beta_1 MIXED_j + \beta_2\left(\pi_j - \pi_{eff}\right)\right)}$$

- Utility of firm *i* is a linear function of the weekend price, number of competing firms adopting internet, population, home price, income.
- Weekend price excluded from the other firms' utility and from the selection mechanism.
- Two step procedure.
- First stage: private information game.
- Starting values in regression coefficients in a probit regression with predicted number of competitors.

Table 4: Parameter Estimates.

Parameter	*Mean*	*Median*	*Confidence Interval*
Constant	-1.1830	-1.1511	[-1.5017,-1.1218]
Competition Penalty	-6.1076	-6.1226	[-6.0331,-6.4116]
Weekend Price	0.3365	0.3337	[0.3321, 0.3447]
Population	0.0006	-0.0012	[-0.0207,0.0161]
Median House Value	-0.0264	-0.0243	[-0.0766, 0.0065]
Median House Income	-0.0878	-0.0525	[-0.2520,0.0128]
Mixed Strategy Selection	0.1984	0.1841	[0.1841, 0.2459]
Most Efficient Strategy Selection	0.5807	0.6239	[0.6407,0.3242]

- Negative constant shows disincentive to adopt a web site absent all other effects.
- Adoption is driven by the price differences.
- Significant penalties with adopting a web site when your competition has also done so.
- But not significant determinants of the decision of adopting web site.
- Price is most informative.
- Mixed strategy equilibria are more likely to be played than pure strategy equilibria! But marginal effect is small.
- Potentially large number of mixed strategy equilibria.
- Sampler draws games with only mixed strategy equilibria.
- Efficient equilibrium most likely to be played.
- No dynamics yet.

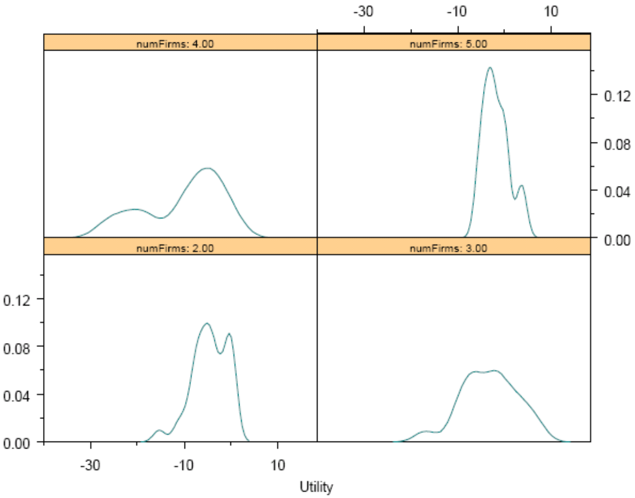
Model Simulations

- How close the expected utility of an observed outcome is to the joint utility maximizing equilibrum.
- Solve for all equilibria of observed games.
- Compute the expected surplus as the expected payoff of the observed action and the most efficient equilibrium.
- Distributions of the observed utility surplus.
- Negatively skewed to the left of zero.
- Percentage difference: two modes at 0 and 100.
- Larger market tends to be more efficient.

Table 5: Simulation Results.

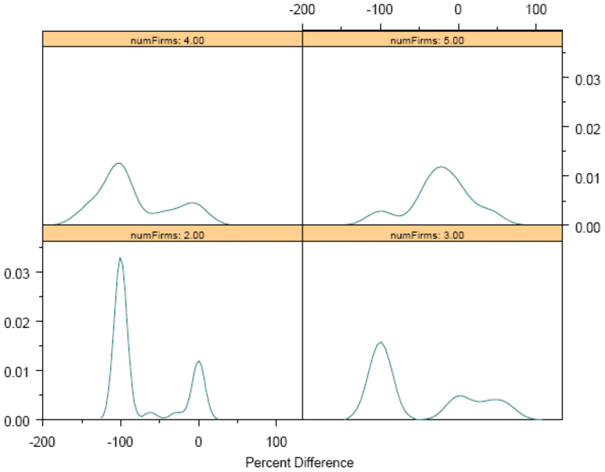
Number of Firms in Market	*2*	*3*	*4*	*5*
Minimum	-15.33	-16.89	-26.02	-5.38
1st Quartile	-6.76	-6.79	-15.73	-3.7
Mean	-4.56	-3.22	-10.23	-1.74
Median	-4.58	-2.94	-7.84	-2.26
3rd Quartile	-1.06	0.46	-4.74	0
Maximum	0	7.2	0	3.68
Number Observations	32	20	10	8
Standard Deviation	3.83	5.98	8.67	2.89

Figure 1: Distribution of Surplus by Size of Market



This figure shows the distribution of expected surplus conditional on the size of the market. The expected surplus is calculated as the difference between the expected utility of the observed outcome and the expected utility of the most efficient equilibrium.

Figure 2: Distribution of Percentage of Expected Maximum Surplus by Size of Market



This figure shows the distribution of the percentage of expected maximum surplus conditional on the size of the market. The distribution is calculated as the percent difference between the expected utility of the observed outcome and the expected utility of the most efficient equilibrium.

- Strategic entry by bidder in highway procurement auctions.
- California Department of Transportation, 1999-2000.
- Multiple entry equilibria in bidding markets.
- Equilibrium selection important for simulating entry process.
- Two stage game.
- Stage 1: simultaneous entry decision by bidders.
- Stage 2: first price sealed bid auction.
- $i = 1, \ldots, N$ potential bidders.
- $t = 1, \ldots, T$ highway paving contracts.
- first price auction, independent private values.
- $N(t) \in \{1, \ldots, N\}$.

- Private information of cost estimate $F_i\left(c_{it}|z_{it}\right)$
- z_{it}: distance of contractor i to project t, project fixed effects.
- Bidder i chooses b_{it} to maximize

$$
\begin{aligned}
&(b_{i,t} - c_{i,t}) \prod_{j \in N(t), j \neq i} \left(1 - F_j\left(\phi_{j,t}(b_{i,t})|z_{j,t}\right)\right) \\
=&(b_{i,t} - c_{i,t}) \prod_{j \in N(t), j \neq i} \left(1 - G_j(b_{i,t}|z_t)\right).
\end{aligned}
$$

- First condition of profit maximization

$$
c_{i,t} = b_{i,t} - \left[\sum_{j \in N(t), j \neq i} \frac{g_j(b_{i,t}|z_t)}{(1 - G_j(b_{i,t}|z_t))}\right]^{-1}
$$

- Guerre, Perrigne and Vuong (GPV).

- Bid cost substantial part of the profit margins.
- 271 bidding firms. 4 major firms.
- Entry of other bidders $N(t) - \{1, 2, 3, 4\}$ predetermined.
- Participating bidders $N(t|a)$ include large entering firms and exogenous fringe firms.
- Zero profit if a firm does not participate.
- Profit if a firm participates

$$u_i(a_i = 1, a_{-i}; z_t) = \int (b_{i,t} - c_{i,t}) \prod_{j \neq i} (1 - G_j(b_{i,t}|z_t, N(t|a))) dF(c_{i,t}|z_{i,t}) - f_i$$

- $b_{it} = b_{it}(c_{it}; z_t, N(t|a))$ equilibrium bidding strategy.

- Step 1, form estimates of $\widehat{G}_j(b_{i,t}|z_t, N(t|a))$ and $\widehat{g}_j(b_{i,t}|z_t, N(t|a))$.
- Step 2, recover private cost information:

$$\widehat{c}_{i,t} = b_{i,t} - \left[\sum_{j \neq i} \frac{\widehat{g}_j(b_{i,t}|z_t, N(t|a))}{(1 - \widehat{G}_j(b_{i,t}|z_t, N(t|a))} \right]^{-1}$$

- Step 3, simulate $s = 1, \ldots, S$ random draws $b_{it}^{(s)}$ from estimated bid distribution

$$c_{i,t}^{(s)} = b_{i,t}^{(s)} - \left[\sum_{j \neq i} \frac{\widehat{g}_j(b_{i,t}^{(s)}|z_t, N(t|a))}{(1 - \widehat{G}_j(b_{i,t}^{(s)}|z_t, N(t|a))} \right]^{-1}$$

- Simulate expected profit

$$\widehat{u}_i(b_{i,t}, c_{i,t}; z_t, N(t|a)) = \frac{1}{S} \sum_s (b_{i,t}^{(s)} - c_{i,t}^{(s)}) \prod_{j \neq i, j \in N(t)} (1 - \widehat{G}_j(b_{i,t}^{(s)}|z_t, N(t|a))) - f_i \tag{6}$$

Equilibrium Selection

- Pure strategy equilibium vs mixed strategy equilibrium
- Efficiency of equilibrium: joint payoff maximizing.
- does not account for revenue or other fringe participants.
- Dummy for pareto dominance.
- Nash product of players' utilities.

Table: Bidder Identities and Summary Statistics

Company	Share	No. Wins	No. Bids Entered	Participation Rate	Total Bids for Contracts Awarded
Granite Construction Company	27.2%	76	244	58.9%	343,987,526
E. L. Yeager Construction Co Inc	10.4%	13	31	7.5%	132,790,460
Kiewit Pacific Co	6.6%	5	30	7.2%	112,057,627
M. C. M. Construction Inc	6.5%	2	6	1.4%	89,344,972
J. F. Shea Co Inc	3.3%	9	40	9.7%	43,030,861
Teichert Construction	3.3%	16	43	10.4%	40,177,076
W. Jaxon Baker Inc	2.9%	13	65	15.7%	37,702,631
All American Asphalt	2.2%	14	33	8.0%	30,764,962
Tullis And Heller Inc	2.1%	10	16	3.9%	27,809,535
Sully Miller Contracting Co	1.9%	17	49	11.8%	27,889,186

Covariates

- $b_{i,t}$—The bid of contractor i on project t.
- EST_t—The engineer's estimate for project t.
- $DIST_{i,t}$—The distance (in miles) of firm i to project t.
- $CAP_{i,t}$—The capacity utilization of firm i at the time of bidding for project t.
- $FRINGE_{i,t}$—A dummy variable equal to one if firm i is a fringe firm.

Table: Bidding Summary Statistics

	Obs	Mean	Std. Dev.	Min	Max
Winning Bid	414	3,203,130	7,384,337	70,723	86,396,096
Markup: (Winning Bid-Estimate)/Estimate	414	-0.0617	0.1763	-0.6166	0.7851
Normalized Bid: Winning Bid/Estimate	414	0.9383	0.1763	0.3834	1.7851
Second Lowest Bid	414	3,394,646	7,793,310	84,572	92,395,000
Money on the Table: Second Bid-First Bid	414	191,516	477,578	68	5,998,904
Normalized Money on the Table: (Second Bid-First Bid)/Estimate	414	0.0679	0.0596	0.0002	0.3476
Number of Bidders	414	4.68	2.30	2	19
Distance of the Winning Bidder	414	47.47	60.19	0.27	413.18
Travel Time of the Winning Bidder	414	56.95	64.28	1.00	411.00
Utilization Rate of the Winning Bidder	414	0.1206	0.1951	0.0000	0.9457
Distance of the Second Lowest Bidder	414	73.55	100.38	0.19	679.14
Travel Time of the Second Lowest Bidder	414	82.51	97.51	1.00	614.00
Utilization Rate of the Second Lowest Bidder	414	0.1401	0.2337	0.000	0.9959

Bid Regression

Table: Bid Function Regressions

Variable	$b_{i,t}$	$b_{i,t}/EST_t$	$b_{i,t}/EST_t$	$b_{i,t}/EST_t$	$b_{i,t}/EST_t$
EST_t	1.025 (56.26)				
$DIST_{i,t}$		.000246 (5.66)	.000249 (5.73)	.000223 (5.01)	
$UTIL_{i,t}$				0.02539 (0.93)	
$FRINGE_{i,t}$					04288 (4.65)
Constant	-25686 (0.56)	1.19 (94.9)	1.007486 (XXXX)	1.001 (79.98)	
Fixed Effects	No	Project	Project	Project	Project/Firm
R^2	0.989	0.5245	0.5290	0.5292	0.5321

Number of observations = 1938.

Logit entry model

Table: Logit Model of Entry

	I	II	III
Constant	-.9067 (7.91)	-1.6811 (7.53)	
$DIST_{i,t}$	-.00218 (5.42)	-.00322 (5.66)	-.00854 (4.85)
Granite	2.889	4.4537 (13.28)	(7.31)
E. L. Yeager		-	-
Kiewit Pacific		-.1527 (0.57)	1.1969 (2.1)
M. C. M.		-1.786 (3.94)	-.70779 (1.12)
Fixed Effects		Project	
Observations	1656	1656	1068
Number of Groups		261	
Log-Likelihood	-784.20	-511.86	-101.0728

The dependent variable is whether one of the 4 largest firms in the industry decides to submit a bid in a particular procurement. Standard errors are shown in parentheses.

Estimates of profits

- First stage linear regression

$$\frac{b_{i,t}}{EST_t} = x'_{i,t}\theta + u^{(t)} + \epsilon_{i,t}$$

- $u^{(t)}$: auction specific fixed effect.
- $\hat{H}$, Kaplan-Meier estimate of CDF of the fitted residuals $\hat{\epsilon}_{it}$.
- Impute bid distribution of nonparticipating firms.
- Need to compute distance of nonparticipating firms.

Table: Margin Estimates

Variable	Num. Obs.	Mean	Std. Dev.	Median	25th Percentile	75th Percentile
Profit Margin	1938	0.0644	0.1379	0.0271	0.0151	0.0520

The markup is defined as 1 minus the ratio of the estimated cost (private information) to the bid.

Equilibrium selection and bid costs

Table: Games Estimation Results

Variable	Mean	Median	Std.Dev.	95% Confidence Interval	
		Equilibrium Selection Parameters (λ)			
Pure Strategy	-1.3524	-1.5345	0.7979	-2.4903	0.1954
Efficient	6.4365	6.4226	0.5321	5.6151	7.5149
Dominated	-5.3841	-5.3316	0.7002	-6.7164	-4.0986
Nash Product	4.4143	4.2025	1.1017	2.9651	6.4836
		Profit Scale			
Profit Scale	0.0965	0.0954	0.0015	0.0954	0.0984
		Bid Preparation Costs (f_i)			
Granite Construction	0.2341	0.2393	0.0977	0.0679	0.4271
E. L. Yeager	1.4583	1.4757	0.0941	1.2563	1.6227
Kiewit Pacific	1.6751	1.6720	0.0511	1.5775	1.7789
M. C. M. Construction	2.4490	2.4360	0.1144	2.2547	2.6966

Estimation was run using LTE method of Chernozhukov and Hong citeyearhong-chernozhukov:03. A Markov chain was generated with 500 draws for each parameter. 409 importance games were used in the importance sampler for the 409 observations.

Discussion

- Bid costs consistent with participation rates and expected profits.
- Mixed strategy equilibrium more likely than pure strategy ones.
- Strong selection effect of efficiency.
- Potential collusive implications.
- Negative coefficient on dominated equilibrium.
- Positive coefficient on highest Nash product.

Conclusion

- Develop algorithms to estimate utilities and equilibrium selection parameters for static, discrete games.
- Computationally efficient.
- Works well in finite sample Monte Carlo simulations.
- Nonparametric identification: exclusion restrictions.
- Variables that shift equilibrium selection only.
- Variables that shift a specific agent utility only.
- Application to entry of four largest firms into California procurement auctions.