Applications of Moment Inequalities: Ho (2009) + Ho, Ho, and Mortimer (2011)

Michael J. Dickstein
Stanford University

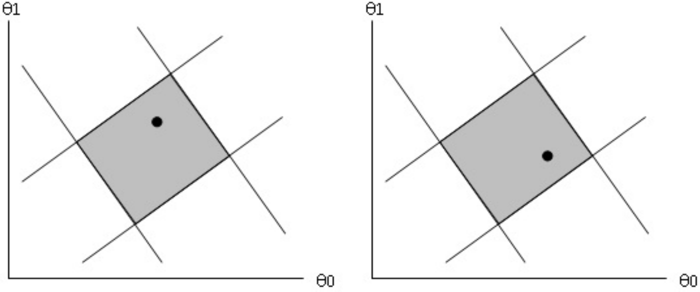
Economics 258

March 9, 2014

Identified Set

- Moment inequalities will generically lead to set identification. Given a set S of moment inequalities, the identified set is:

$$\Theta^S = \underset{\theta}{\operatorname{argmin}} \sum_{s=1}^{S} \Big(\min \big\{ 0, \mathbb{E}[m_s(Y, X, Z; \theta)] \big\} \Big)^2$$



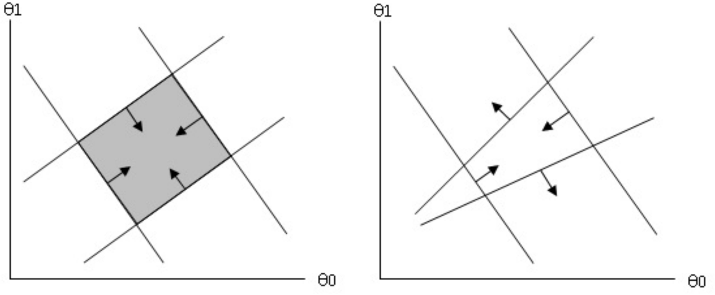
Steps for Estimation

- Step 1: Estimate the identified set given sample moments.
- Step 2: Perform inference on one or more of the following parameters:
  - Interval contained in the identified set: Pakes, Porter, Ho and Ishii (2011).
  - Identified set: Chernozhukov, Hong and Tamer (Econometrica, 2007).
  - True parameter vector: Andrews and Soares (Econometrica, 2010).

Estimation of the Identified Set

- Estimation is based on the sample analogue of the moment inequalities:

$$\overline{m}_{n,s}(\theta) = \frac{1}{n}\sum_{i=1}^{n} m_s(Y_i, X_i, Z_i; \theta)$$



(c) Case 1　　(d) Case 2

Estimation of the identified set

- Two possible criterion functions to define the estimated set:
  - Unweighted criterion function:

$$\hat{\Theta}_n^S = \underset{\theta}{\operatorname{argmin}} \sum_{s=1}^{S} \Big(\min\{0, \overline{m}_{n,s}(\theta)\} \Big)^2$$

  - Weighted criterion function:

$$\hat{\Theta}_n^S = \underset{\theta}{\operatorname{argmin}} \sum_{s=1}^{S} \Big(\min\{0, \Big[\frac{\overline{m}_{n,s}(\theta)}{\hat{\sigma}_{n,s}^2(\theta)}\Big]\} \Big)^2,$$

  with

$$\hat{\sigma}_{n,s}^2(\theta) = \frac{1}{n} \sum_{i=1}^{n} (m_s(Y_i, X_i, Z_i; \theta) - \overline{m}_{n,s}(\theta))^2$$

- The weighting lessens the influence of sample moments that have high variance (likely to be further away from their population analogues).

Computation of the estimated set

- We characterize the set $\hat{\Theta}_n^S$ by finding its boundaries along any linear combination of the dimensions of vector θ.
- If the moment functions $\{\overline{m}_{n,s}(\theta) : s = 1, \ldots, S\}$ are linear in θ, use linear programming to find the extremum

$$\begin{aligned} \max_{\theta} \quad & f \cdot \theta \\ \text{s.t.} \\ & \overline{m}_{n,s}(\theta) \geq 0, \text{ for } s = 1, ..., S. \end{aligned} \tag{1}$$

- To find the maximum and minimum of our two-dimensional parameter θ, we use:

$$f = \{[1, 0], [-1, 0], [0, 1], [0, -1]\}.$$

- Apply simplex routine in Matlab via *linprog*

Computation of the estimated set

- If there is no value of θ that verifies all the constraints, $\hat{\Theta}_n^S$ will be a singleton.
- This singleton is the outcome of a nonlinear optimization problem:

$$\hat{\Theta}_n^S = \underset{\theta}{\operatorname{argmin}} \sum_{s=1}^{S} \Big(\min\{0, \Big[\frac{\overline{m}_{n,s}(\theta)}{\hat{\sigma}^2_{n,s}(\theta)}\Big]\}\Big)^2.$$

- Use a nonlinear optimization package, like *KNITRO* (in Matlab via *ktrlink* with license).

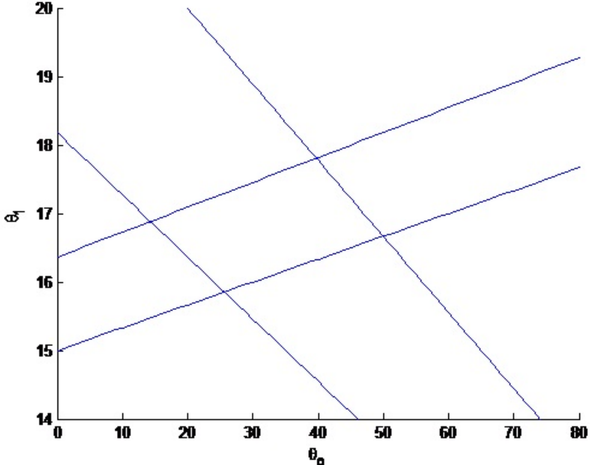
Computation of the estimated set: example

- Sample moments:

$$
\begin{aligned}
900 - \theta_0(-2) - \theta_1(60) &\geq 0 \\
-900 - \theta_0(2) - \theta_1(-55) &\geq 0 \\
200 - \theta_0(1) - \theta_1(9) &\leq 0 \\
-200 - \theta_0(-1) - \theta_1(-11) &\leq 0
\end{aligned}
$$

Computation of the estimated set: example

Vertex, θ0 min		Vertex, θ0 max		Vertex, θ1 min		Vertex, θ1 max	
θ0	θ1	θ0	θ1	θ0	θ1	θ0	θ1
13.0	16.8	51.0	16.6	23.9	15.8	41.3	17.8



Inference: General Intuition

- Consider we want to test the null hypothesis: $H_0 : \theta = \theta_0$.
- We use the following statistic:

$$T_n(\theta_0) = \sum_{s=1}^{S} \left(\min\{0, \left[\frac{\overline{m}_{n,s}(\theta_0)}{\hat{\sigma}^2_{n,s}(\theta_0)}\right]\}\right)^2.$$

- The finite-sample null distribution of $T_n(\theta_0)$ depends on the degree of *slackness* of the population moments–i.e. how much greater than 0 is:

$$\mathbb{E}[m_s(Y_i, X_i, Z_i; \theta)], \quad \text{for } s = 1, \ldots, S.$$

Inference: General Intuition

- Key: need to infer whether a population moment binds at a particular value θ_0.
- Compute slackness factor, $SF_{n,s}(\theta_0)$
  - accounts for whether moment is likely to be binding
  - moments likely to be nonbinding asymptotically –i.e. $\overline{m}_{n,s}(\theta) >>> 0$–will have larger slackness factors

Inference: General Intuition

- Three slackness factors proposed in the literature:
  - Assume that all the S moments are binding at θ_0: $SF_{I,s} = 0$.
    - yields the most conservative test
  - Moment Selection:

$$SF^{MS}_{n,s}(\theta_0) = \mathbb{1}\{\sqrt{n}(\frac{\overline{m}_{n,s}(\theta_0)}{\hat{\sigma}_{n,s}(\theta_0)}) \leq \sqrt{2\ln(\ln(n))}\}$$

  - Shifted Mean: shift each moment proportionately to how far away from binding it is in the sample.

$$SF^{SM}_{n,s}(\theta_0) = (\frac{\overline{m}_{n,s}(\theta_0)}{\hat{\sigma}_{n,s}(\theta_0)})(\frac{1}{\sqrt{2\ln(\ln(n))}})\mathbb{1}\{\frac{\overline{m}_{n,s}(\theta_0)}{\hat{\sigma}_{n,s}(\theta_0)} > 0\}$$

Inference for an Interval: PPHI (2011)

- Objective: build confidence intervals for the vertices of the estimated set, and use the outer bounds to form a unique confidence interval.
- We need four elements for inference:
    - Vertices of the estimated set.
    - Approximation to the asymptotic distribution of all the (weighted) moments recentered at zero.
    - Jacobian of the moments.
    - Slackness factors.

Inference for an Interval

- Approximation to asymptotic distribution of all the recentered moments.
  - Draw $r = 1, ..., R$ times from a multivariate normal with zero mean, and covariance equal to the variance of the weighted moments
    - Take R standard normal draws.
    - Premultiply each draw by the Cholesky decomposition of the correlation matrix evaluated at the vertex of interest, $\widehat{\Omega}_{n,S}(\hat{\theta})$:

$$\widehat{\Omega}_{n,S}(\hat{\theta}) = diag(\widehat{\Sigma}_{n,S}(\hat{\theta}))^{-\frac{1}{2}}\widehat{\Sigma}_{n,S}(\hat{\theta})diag(\widehat{\Sigma}_{n,S}(\hat{\theta}))^{-\frac{1}{2}}.$$

    - Result:

$$q_r(\hat{\theta}) = chol(\widehat{\Omega}_{n,S}(\hat{\theta}))N(0_S, I_S).$$

Inference for an Interval

- Jacobian of the moments.
    - Compute the Jacobian of the sample unweighted moments, $\overline{m}_{n,s}(\theta)$, and evaluate the result at the vertex of interest:
    - When the moments are linear in θ, the derivative matrix multiplied by θ simply equals the mean of the weighted moments:

$$\widehat{\Gamma}_{n,s}(\theta) * \theta = \frac{1}{n}\left[\sum_{i=1}^{n}\frac{\Delta x_{i,s}}{\widehat{\sigma}_{n,s}(\theta)}, \sum_{i=1}^{n}\frac{\Delta y_{i,s}}{\widehat{\sigma}_{n,s}(\theta)}\right] * \begin{pmatrix}\theta_0\\ \theta_1\end{pmatrix} = \frac{m_{n,s}(\theta)}{\widehat{\sigma}_{n,s}(\theta)}$$

- evaluate the weights, $\widehat{\sigma}_{n,s}(\theta)$, at θ values equal to the relevant vertex.

Inference for an Interval

- Evaluate the slackness factor at the vertex of interest and normalize by $\sqrt{n}$.
    - We could use either $SF^{MS}_{n,s}$ or $SF^{SM}_{n,s}$.
      The option described in Pakes, Porter, Ho, and Ishii (2011) is Shifted Mean:

$$SF^{SM}_{n,s}(\hat{\theta})\sqrt{n} = \left(\frac{\overline{m}_{n,s}(\hat{\theta})}{\hat{\sigma}_{n,s}(\hat{\theta})}\right)\left(\frac{1}{\sqrt{2\ln(\ln(n))}}\right)\mathbb{1}\left\{\frac{\overline{m}_{n,s}(\hat{\theta})}{\hat{\sigma}_{n,s}(\hat{\theta})} > 0\right\}\sqrt{n}$$

Inference for an Interval

- Compute the following linear programing problem for each draw $r = 1, ..., R$ and each vertex $\hat{\theta}$: (total of $2xdxR$ optimizations)

$$
\begin{aligned}
\theta_r = \max_{\theta} \quad & f \cdot \sqrt{n}(\hat{\theta} - \theta) \\
\text{s.t.} \quad & \\
& \widehat{\Gamma}_{n,S}(\hat{\theta})\sqrt{n}(\hat{\theta} - \theta) + q_r(\hat{\theta}) + SF_{n,S}^{SM}(\hat{\theta})\sqrt{n} \geq 0
\end{aligned} \tag{2}
$$

- As before, to find the maximum and minimum of our two-dimensional parameter θ, we use:

$$f = \{[1, 0], [-1, 0], [0, 1], [0, -1]\}.$$

In equation (2), use the estimated vertex $\hat{\theta}$ that corresponds to each vector f.

- We obtain R draws of the asymptotic distribution of each of

Inference for an Interval

- For each pair of vertices corresponding to a given dimension d of θ.
  - For the min vertex, take the $\alpha/2$ quantile of the set of simulated vertices, θ_r, $r = 1, \ldots, R$. Denote this number:

    $$\underline{\theta}_{d,\alpha/2}.$$

  - For the max vertex, take the $(1 - \alpha/2)$ quantile of the set of simulated vertices, θ_r, $r = 1, \ldots, R$

    $$\overline{\theta}_{d,1-\alpha/2}.$$

- The confidence interval for θ in the dimension d with significance level α is:

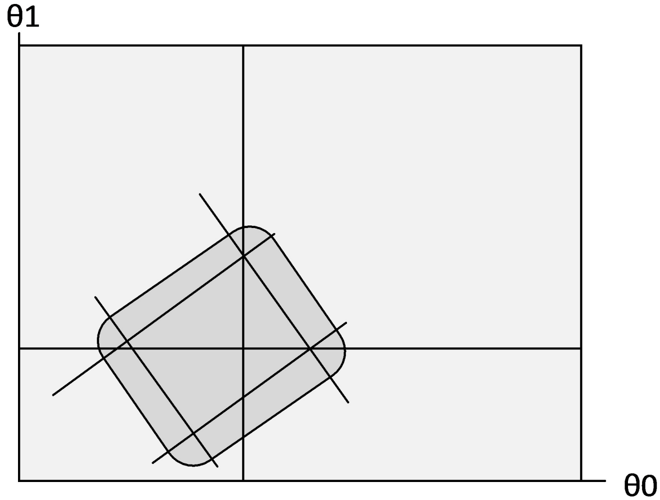
  $$(\underline{\theta}_{d,\alpha/2}, \overline{\theta}_{d,\alpha/2}).$$

Set/Point Inference: General Intuition

- Based on the inversion of an Anderson-Rubin T statistic.
- General steps in the algorithm:
  1. Define θ grids, $\widehat{\Theta}_n^{Grid}$ and $\widehat{\Theta}_n^{\epsilon}$, where $\widehat{\Theta}_n^{\epsilon} \subset \widehat{\Theta}_n^{Grid}$.
  2. Calculate $T_r(\theta)$, at a set of points in either $\widehat{\Theta}_n^{Grid}$ or $\widehat{\Theta}_n^{\epsilon}$ depending on whether the focus of inference is the identified set or the true value of the parameter.
  3. Determine a critical value as a quantile of $T_r(\theta)$ for $r = 1, ..., R$
  4. Calculate $T^{obs}(\theta)$ at each $\theta \in \widehat{\Theta}_n^{Grid}$ with the observed data for all moments.
  5. Define the confidence set as those θ points where $T^{obs}(\theta)$ falls below the critical value.

Forming the Grids: $\widehat{\Theta}_I^{Grid}$ and $\widehat{\Theta}_I^{\epsilon}$

$\widehat{\Theta}_n^{\epsilon} \subset \widehat{\Theta}_n^{Grid}$



Inference for the Identified Set

Chernozhukov, Hong and Tamer (Econometrica, 2007)

- Steps of the procedure:
  - (1) At $\theta \in \widehat{\Theta}_n^{\varepsilon}$, compute R draws $\{q^r(\theta); r = 1, \ldots, R\}$ such that:

    $$q_r(\theta) = chol(\widehat{\Omega}_{n,S}(\theta))N(0_S, I_S),$$

    with

    $$\widehat{\Omega}_{n,S}(\hat{\theta}) = diag(\widehat{\Sigma}_{n,S}(\hat{\theta}))^{-\frac{1}{2}}\widehat{\Sigma}_{n,S}(\hat{\theta})diag(\widehat{\Sigma}_{n,S}(\hat{\theta}))^{-\frac{1}{2}}.$$

    Note that we are taking draws from the asymptotic distribution of the normalized recentered moments, evaluated at each point θ.

Inference for the Identified Set

- Steps of the procedure (cont.)
  - (2) Compute one of the following T-statistic for each value of θ and draw r:

$$T_r^N(\theta) = \sum_{s=1}^{S} \left(\min\{0, q_{r,s}(\theta)\}\right)^2$$

$$T_r^{MS}(\theta) = \sum_{s=1}^{S} \left\{\left(\min\{0, q_{r,s}(\theta)\}\right)^2 \times SF_{n,s}^{MS}(\theta)\right\}$$

$$T_r^{SM}(\theta) = \sum_{s=1}^{S} \left(\min\{0, q_{r,s}(\theta) + SF_{n,s}^{SM}(\theta)\}\right)^2$$

  - (3) For each draw r, take the maximum across θ:

$$T_r^{\max} = \max_{\theta \in \widehat{\Theta}_n^{\varepsilon}} T_r^k(\theta), \quad k = \{N, MS, SM\}.$$

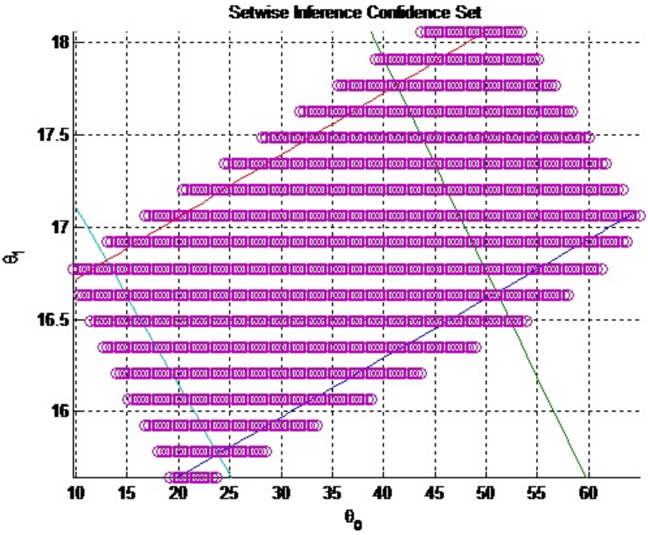
Inference for the Identified Set

- Steps of the procedure (cont.)
  - (4) Compute the critical value c_α as the $1-\alpha$ quantile of the distribution of $\{T_r^{\max}; r = 1, \ldots, R\}$.
  - (5) Return to the larger grid of theta points, $\widehat{\Theta}_n^{Grid}$, and calculate $T^{obs}(\theta)$ at each candidate value $\theta \in \widehat{\Theta}_n^{Grid}$:

$$T^{obs}(\theta) = \sum_{s=1}^{S} (\min\{0, \frac{\overline{m}_{n,s}(\theta)}{\hat{\sigma}_{n,s}(\theta)}\})^2$$

  - (6) Compare $T^{obs}(\theta)$ against c_α and accept θ into the confidence set whenever $T^{obs}(\theta) < c_\alpha$.

Inference for the Identified Set: Example



Inference for the True Parameter

Andrews and Soares (2010)

- Steps of the procedure:
  - (1) At every $\theta \in \widehat{\Theta}_n^{Grid}$, calculate $\{q_r(\theta);\ r = 1, \ldots, R\}$:

    $$q_r(\theta) = chol(\widehat{\Omega}_{n,S}(\theta))N(0_S, I_S)$$

  - (2) For each of these θ and r, calculate: $T_r(\theta)$, $T_r^{MS}(\theta)$, or $T_r^{SM}(\theta)$.
  - (3) For each θ, calculate the $(1-\alpha)$ quantile. This the critical value, $c(\alpha, \theta)$.
  - (4) Calculate $T^{obs}(\theta)$ at each candidate value $\theta \in \widehat{\Theta}_n^{Grid}$.
  - (5) Compare $T^{obs}(\theta)$ against $c(\alpha, \theta)$ and accept θ into the confidence set whenever $T^{obs}(\theta) < c(\alpha, \theta)$.

Comparison of Inference Procedures

Table 1: Estimated Confidence Intervals				
	θ0 min	θ0 max	θ1 min	θ1 max
True Θ Interval	14,376	15,298	1,305	1,321
Average Confidence Intervals From				
PPHI Procedure	13,573	16,170	1,283	1,343
Ptwise Procedure	8,500	20,307	832	1,850
Setwise Procedure	8,485	20,321	830	1,852
Moment Selection, Ptwise Procedure	8,514	20,280	832	1,850
Moment Selection, Setwise Procedure	8,485	20,321	830	1,852
Shifted Mean, Ptwise Procedure	10,775	19,272	927	1,658
Shifted Mean, Setwise Procedure	8,822	20,014	858	1,820

Ho 2009

- Theory testing
- Measurement
- Methodology

Ho 2009

- Theory testing
  - Can a bargaining model explain the hospital-insurance plan contracting process, rationalizing the observed network of hospital-plan relationships?
- Measurement
  - What characteristics of hospitals and plans explain the level of surplus hospitals can extract from the relationship?
  - What is the effect of capacity constraints on producer welfare? Might the level of capacity be a relevant choice variable for a profit-maximizing firm?

Ho 2009

- Methodology
  - What assumptions are needed on behavior to develop a moment inequality estimator for static contracting problems?
  - What can information on ex-post network formation reveal about private negotiated prices?

Ho 2009

Main Idea

- Model demand for hospitals and health plans, accounting for the hospital network of each plan in the consumer's plan choice
- Model the supply side negotiation between hospitals and plans in forming equilibrium networks, which determines the division of profits
- To increase their share of the surplus from contracting, hospitals have incentives to:
  - Invest in quality to attract more patients, lower costs
  - Merge with other providers, to improve bargaining position
  - Under-invest in capacity

Ho 2009

Main Idea

- Findings
  - “Star“ hospitals capture $6700 more per patient than other providers, on costs of $11,000
  - Hospitals with capacity constraints have markups of $6900 per patient more than those without constraints
  - System hospitals have $180,000/month greater profits than other providers

Ho 2009

Data

- Insurer plan data cover all managed care insurers in 43 major markets across the US for Q3, Q4 of 2002 (cross-section)
  - Premiums earned, number of enrollees, tax status of each carrier
  - Data on clinical performance and patient satisfaction with health plans
- Hospital data from Medstat from private insurers; includes encounter-level data on hospital admissions during 2 year period.
  - Patient's diagnosis and characteristics, identity of hospital, type of plan
  - Hospital characteristics from AHA
- Data on network of hospitals for every HMO/POS plan in every market considered in March/April 2003

Ho 2009

Model: Stages

- 0. Plans choose quality and products; Hospitals choose capacity, location, product mix, system mergers.
- 1. Hospitals make simultaneous take-it-or-leave-it price offers to all plans in the market
- 2. Plans choose whether to accept these offers, forming their hospital network
- 3. Plans set premiums to maximize profits after a change in networks
- 4. Consumers and employers jointly choose plans
- 5. Sick consumers visit hospitals; plans pay hospitals per service provided.

Ho 2009

Model: Hospital Demand

$$u_{i,h,l} = \eta_h + x_h\alpha + x_h\nu_{i,l}\beta + \varepsilon_{i,h,l}$$

- individual i, hospital h, diagnosis l
- x_h observed hospital characteristics
- $\nu_{i,l}$ observed characteristics of consumers
- Estimate via ML, using Medstat data
    - Medstat doesn't have hospital networks for managed care enrollees; use only data on indemnity and PPO enrollees whose choice set is unrestricted
    - Assume: indemnity/PPO enrollees have same preferences over hospitals as managed care enrollees (vertical preferences)

Ho 2009

Model: Health Plan Demand

$$\widetilde{u}_{ijm} = \xi_{jm} + z_{jm}\lambda + \gamma_1 EU_{ijm} + \gamma_2 \frac{prem_{j,m}}{y_i} + \omega_{ijm}$$

- individual i, plan j, market m
- (z_{jm}, ξ_{jm}) observed and unobserved plan characteristics
- outside option = choosing to be uninsured; indemnity/PPO is separate choice in each mkt
- IV for premium
  - plan char, avg hourly hospital wage, avg weekly nurse wage
  - exclusion restriction: health plan costs correlated with premiums but not with unobs plan quality
- Find: consumers value EU from network in plan choice

Ho 2009

Model: Producer surplus generated by network

$$S_{j,m}(H_j, H_{-j}) = \sum_i (n_i s_{ijm}(H_j, H_{-j})[prem_{j,m} - p_i \sum_{h \in H_j} s_{i,h}(H_j)\text{cost}_h])$$

- The shares $s_{ijm}(H_j, H_{-j})$ are plan j's predicted shares of type i people when networks (H_j, H_{-j}) offered (flow of consumers to plans after network changes)
- $s_{i,h}(H_j)$ hospitals h's predicted share of type i people (flow of consumers to hospitals after network change)

Ho 2009

Model: Producer surplus generated by network

$$S_{j,m}(H_j, H_{-j}) = \sum_i (n_i s_{i,j,m}(H_j, H_{-j})[prem_{j,m} - p_i \sum_{h \in H_j} s_{i,h}(H_j)\text{cost}_h])$$

- Premiums adjust in response to changes in hospital network
  - (1) Estimate supply model assuming fixed premiums
  - (2) Allow all plans to simultaneously adjust their premiums to max profits
  - Comment: with panel data, could push further
- No non-hospital variable costs
- Adjusts for capacity constraints at 85% level

Ho 2009

Model: Negotiation

- All hospitals make TIOLI offers of {contract,null offer}
- All plans simultaneously respond
- Offers are private info; plans have passive beliefs (if plan gets an alternative offer from h, doesn't change plan's beliefs about offers h makes to its competitors)

$$\pi^P_{j,m} = S_{j,m}(H_j, H_{-j}) - c^{Hosp}_{j,m}(H_j, H_{-j}, X, \theta) - c^{nonhosp}_{j,m}(H_j, H_{-j}, X, \theta$$

$$\begin{aligned} \pi^{P,o}_{j,m}(.) &= \pi^P_{j,m} + \mu_{j,H_j} \\ E[\pi^P_{j,m}(H_j, H_{-j}, X, \theta)|Ij, m] &= \pi^P_{j,m}(H_j, H_{-j}, X, \theta) - \varphi_{j,H_j} \end{aligned}$$

Ho 2009

Model: Negotiation

- Key assumption: plan j's expected profits from $H_j >$ expected profits from alternative network formed by reversing contract with h

$$E[\pi_{j,m}^{P,o}(H_j, H_{-j}, X, \theta) - \pi_{j,m}^{P,o}(H_j^h, H_{-j}, X, \theta)|Z_{j,m}] \geq 0$$

- form unconditional moments using positive-valued function of $Z_{j,m}$
  - must be known to firms when they make their choice
  - use char in fixed cost and markup terms other than cost/admission
  - use indicators for some plan and market characteristics

Ho 2009

Model: Negotiation

- Choose counterfactuals of reversing a single contract.
- Plans may respond by changing its response to other hospital's offers (passive beliefs rules out the following: plan responds to changes in h's offer by assuming other plans have different offers and therefore change their own networks)
- Two possible routes

Ho 2009

Model: Negotiation

- Two possible routes
  - Assume hospital can make an alternative offer to j that will prompt j to drop h from its network and not change its contract with other hospitals
  - Allow plan to adjust its decisions wrt all other hospitals
    - find min of hospital h's profits from all possible choices plan j can make in response to the deviation (given its other contracts and networks of other plans)
    - form inequality with difference between realized profit and minimized counterfactual profit

Ho 2009

Results

- Estimate of θ for every specification is a singleton; could not satisfy all inequality constraints. Why?
  - random disturbances
  - no. of moments used
  - no structural error (some component of profit function not observed by econometrician but used by the agent, that varies at p,h level)
- Comment
  - Inference for moment inequalities that find a set more complicated
  - Counterfactuals in the case of set identification?

Ho 2009

Results: Substantive findings

- Hospitals in systems take a larger fraction of surplus, penalize plans that do not contract with all members
- Star hospitals capture high mark-ups
- Hospitals with higher costs/pt receive lower markups/pt

Goals of the Paper

- Theory Testing
  - What are the profit consequences for the manufacturer from offering retailers full-line forcing contracts?
  - Does it reduce consumer choice or lead to higher prices?
- Measurement
  - Quantify how consumer demand, retailer revenues and costs, and distributor revenues change when adding/removing FLF from contract mix.
- Methodology
  - "Role model" of bundling analysis: combine detailed demand side estimation of substitution with supply side model of firm's costs from adding inventory.

Setting

Innovation in recording rental transactions led to contract innovation:

- Linear pricing – $65-70 upfront fee per tape
- Revenue sharing - $8 upfront + 55% of rental revenue per tape
    - Have min and max quantity restrictions
- FLF – rental store purchases all titles of distributor
    - Terms like RS, but lower up-front fee ($3.60) and lower rev share (retailers keep 59%)
- Sell-through priced (STP) titles

Setting

Selection on contract type?

- What type of movies should retailer choose to accept under each contract type? (Mortimer (2008))
  - LP for high volume videos/new releases?
  - RS for niche films? RS usually have higher minimum quantity restrictions than the avg # of tapes bought under LP contracts.

Setting

How does inventory choice affect retailer profits?

- Can increase retailer profits by attracting new consumers to store. (included in costs of holding inventory)
- High inventory may lead to high initial demand (consumers see more tapes on shelf); can reduce later month sales (included in costs of holding inventory)
- Sales of substitute products fall with higher inventory on focal product (see demand model)

Model

Retailer portfolio choice problem

- Moment inequalities to bound the value of holding inventory (do not model complicated retailer equilibrium strategies)
- Intuition:
  - on average, stores' profits from the observed portfolio of titles and choices of inventory must exceed profits from alternative portfolios/inventory
  - Dropping a title gives you upper bound to costs of holding inventory
  - Adding tapes (say, 10%) provides lower bound on value of holding inventory.

Model

Retailer portfolio choice problem (continued)

Procedure

- Calculate share of title at store m at time t using demand estimates.
- Determine total returns to the store under its inventory constraints (determined by the contracts it entered).
- Subtract off payments to distributor and the costs of holding a tape to find profits.

Model

Inequalities

$$E[\pi_m^{obs}(.)|I_m] \geq E[\pi_m^{altj'}(.)|I_m]$$

$$\pi_m^{obs} = \sum_s \sum_{j \in J_s} (r_{jm}^{obs}(.) - C(.)\widetilde{c}_{jm}) + \eta_m + \rho(\widetilde{c}_{ms}, k_{ms}) + \varepsilon_{ms}$$

- Assume η_m, $\rho(\widetilde{c}_{ms}, k_{ms})$ difference out
- Have instruments, $Z_{ms'} \subset I_m$
- $E[\varepsilon_{ms}|Z_{ms'}] = 0$
- Use as IV's {constant; indicators for size of store}
- Rules out error term that differs by contract type and that are structural–that is, choice of contract depends on error.
    - STRONG - claim is that specification of inventory holding value captures all elements in I_m