Model Selection Tests for Nonnested Moment Inequality Models

Xiaoxia Shi

Yale University

November 2009
Aim of this Paper

- Construct model selection tests for nonnested *moment inequality models*

- Choose the *correct* model if one of the candidate models is correct

- Choose the model *better fitting* the data if both candidates are misspecified
Moment Inequality Models

- Example: Airline Entry Game with Complete Information, Ciliberto and Tamer (2009) \((\Delta_u \leq 0, \Delta_c \leq 0)\)

<table>
<thead>
<tr>
<th></th>
<th>Continental</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Out</td>
</tr>
<tr>
<td>United</td>
<td>Out</td>
</tr>
<tr>
<td>In</td>
<td>(\theta_u - \varepsilon_u, 0)</td>
</tr>
</tbody>
</table>

- GMM approach: \(E[p_{(I,O)}(\theta) - 1\{Y_i = (I, O)\}] = 0\)

- Multiple equilibria when

\[
\begin{align*}
\theta_u - \varepsilon_u & \geq 0 \geq \theta_u + \Delta_u - \varepsilon_u \\
\theta_c - \varepsilon_c & \geq 0 \geq \theta_c + \Delta_c - \varepsilon_c
\end{align*}
\]

- Without an equilibrium selection rule, \(p_{(I,O)}(\theta)\) is not available
Moment Inequality Models

- $p_{(I,O)}(\theta)$ is bounded from above by

$$\bar{p}_{(I,O)}(\theta) = \Pr((I, O) \text{ is a unique equilibrium}) + \Pr(\text{multiple equilibria})$$

- Moment inequality:

$$E[\bar{p}_{(I,O)}(\theta) - 1\{Y_i = (I, O)\}] \geq 0$$
Moment Inequality Models

- $p(I,O)(\theta)$ is bounded from above by
  
  $$\bar{p}(I,O)(\theta) = \Pr((I,O) \text{ is a unique equilibrium}) + \Pr(\text{multiple equilibria})$$

- Moment inequality:
  
  $$E[\bar{p}(I,O)(\theta) - 1\{Y_i = (I,O)\}] \geq 0$$

- In general
  
  $$E[m(X_i, \theta)] \geq 0$$
Moment Inequality Models

- \( p_{(I,O)}(\theta) \) is bounded from above by
  \[
  \bar{p}_{(I,O)}(\theta) = \Pr((I,O) \text{ is a unique equilibrium}) + \Pr(\text{multiple equilibria})
  \]

- Moment inequality:
  \[
  E[\bar{p}_{(I,O)}(\theta) - 1\{Y_i = (I,O)\}] \geq 0
  \]

- In general
  \[
  E[m(X_i, \theta)] \geq 0
  \]

- \( m \): moment function; \( X_i \): data; \( \theta \): finite-dimensional parameter
Moment Inequality Models

- $p_{(I,O)}(\theta)$ is bounded from above by

$$\bar{p}_{(I,O)}(\theta) = \Pr((I,O) \text{ is a unique equilibrium})$$

$$+ \Pr(\text{multiple equilibria})$$

- Moment inequality:

$$E[\bar{p}_{(I,O)}(\theta) - 1\{Y_i = (I,O)\}] \geq 0$$

- In general

$$E[m(X_i, \theta)] \geq 0$$

- $m$: moment function; $X_i$: data; $\theta$: finite-dimensional parameter

- Compared to GMM models: $E[m(X_i, \theta)] = 0$. 
Applications of Moment Inequality Models

- **Multiple equilibria with no equilibrium selection rule** (e.g. Andrews, Berry and Jia (2004), Ciliberto and Tamer (2009))

- **Games with huge strategy space** (e.g. Pakes, Porter, Ho and Ishii (2004), Ellickson, Houghton and Timmins (2007)), Pakes (2009)

- **Incomplete data (bracket data, missing data)** (e.g. Manski and Tamer (2002))
Estimation of Moment Inequality Models

- Parameter $\theta$ typically not point-identified

- Estimate the identified set $\Theta_{P_0}^*$: $\{\theta : E_{P_0}[m(X_i, \theta)] \geq 0\}$

  - **GMM-type:** Chernozhukov, Hong and Tamer (2007), Andrews and Guggenberger (2009), Andrews and Soares (2010), Bugni (2008), Romano and Shaikh (2009)

  - **Empirical Likelihood:** Canay (2008)

  - **Exponential Tilting:** this paper
Examples:

- $\Delta_u \leq 0, \Delta_c \leq 0$ (business stealing) vs. $\Delta_u \geq 0, \Delta_c \geq 0$ (positive externality)

- e.g. complete information vs. incomplete information?

- e.g. use which covariates and which function form?

$\Rightarrow$ Need a data-driven model selection procedure

- Allow partial identification
Model Selection Test

- $\mathcal{P} = \{ P : E_P m(X_i, \theta) \geq 0, \theta \in \Theta \}$
- $\mathcal{Q} = \{ Q : E_Q g(X_i, \beta) \geq 0, \beta \in B \}$

- $d(\mathcal{P}, P_0)$ "distance" from the true distribution $P_0$ to $\mathcal{P}$
- model $\mathcal{P}$ correctly specified iff $P_0 \in \mathcal{P}$ or $d(\mathcal{P}, P_0) = 0$
Model Selection Test

- \( H_0 : d(P, P_0) = d(Q, P_0) \)

- \( H_{1,Q} : d(P, P_0) > d(Q, P_0) \)
My Model Selection Tests

- Choose the correct model if one of the candidate models is correct.

- Choose the model better fitting the data if both candidates are misspecified.

- Allow the parameters to be partially identified.

- Are computationally tractable (use standard normal critical value).

\[ H_0: d(\mathcal{P}, \mathcal{P}_0) = d(\mathcal{Q}, \mathcal{P}_0) \]
Model Selection tests

- for parametric models: Vuong (1989)
- for moment equality models: Kitamura (2000)

Model selection criterion: AIC, BIC

- do not provide probabilistic conclusion

Cox-type tests (Cox (1961)):

- model evaluation rather than model selection
Road Map Ahead

- Goodness-of-fit measure
- Test statistics
- Model selection tests and their size and power properties
- Monte Carlo
Measure the Goodness-of-fit

- Kullback-Leibler Information Criterion (KLIC) between two distributions:

\[
d(P, P_0) = \begin{cases} 
\int f_{P_0} \log f_{P_0} dP_0 & \text{if } P \ll P_0 \\
\infty & \text{otherwise}
\end{cases}
\]
Measure the Goodness-of-fit

- Kullback-Leibler Information Criterion (KLIC) between two distributions:

\[ d(P, P_0) = \begin{cases} 
\int f_{P_0} \log f_{P_0} \, dP_0 & \text{if } P \ll P_0 \\
\infty & \text{otherwise}
\end{cases} \]

- \( f_{P_0} \): density of \( P \) with respect to \( P_0 \)

\[ H_0: d(P, P_0) = d(Q, P_0) \]
Measure the Goodness-of-fit

- Kullback-Leibler Information Criterion (KLIC) between two distributions:

\[
d(P, P_0) = \begin{cases} 
\int f_{P_0} \log f_{P_0} \, dP_0 & \text{if } P \ll P_0 \\
\infty & \text{otherwise}
\end{cases}
\]

- \(f_{P_0}\): density of \(P\) with respect to \(P_0\)

- \(d(P, P_0) \geq 0\) unless \(P = P_0\):

\[
\int f_{P_0} \log f_{P_0} \, dP_0 \geq (\int f_{P_0} \, dP_0) \log (\int f_{P_0} \, dP_0) = 0
\]
Measure the Goodness-of-fit

- Kullback-Leibler Information Criterion (KLIC) between two distributions:

\[
d(P, P_0) = \begin{cases} \int f_{P_0} \log f_{P_0} dP_0 & \text{if } P \ll P_0 \\ \infty & \text{otherwise} \end{cases}
\]

- \(f_{P_0}\): density of \(P\) with respect to \(P_0\)

- \(d(P, P_0) \geq 0\) unless \(P = P_0\):

\[
\int f_{P_0} \log f_{P_0} dP_0 \geq (\int f_{P_0} dP_0) \log(\int f_{P_0} dP_0) = 0
\]

- Goodness-of-fit of model \(\mathcal{P}\):

\[
d(\mathcal{P}, P_0) = \min_{P \in \mathcal{P}} d(P, P_0)
\]
Measure the Goodness of Fit

- Why Kullback-Leibler?
  - Minimizing KLIC $\iff$ Maximum Likelihood in parametric models
  - Standard AIC, BIC procedures use KLIC
  - Compared to Kolmogorov-Smirnov distance: more tractable
  - Compared to GMM-type criteria: automatic weighting

- $d(P, P_0)$ (exponential tilting) or $d(P_0, P)$ (empirical likelihood)?
  - Exponential tilting is robust to misspecification
  - Results extend to empirical likelihood, when moment function bounded

$H_0: d(P, P_0) = d(Q, P_0)$
Minimum KLIC - Pseudo-true Distribution

- Rewrite

$$\mathcal{P} = \bigcup_{\theta \in \Theta} \mathcal{P}_\theta$$
$$= \bigcup_{\theta \in \Theta} \{ P : E_P m(X_i, \theta) \geq 0 \}$$

$$d(\mathcal{P}, P_0) = \min_{P \in \mathcal{P}} d(P, P_0) = \min_{\theta \in \Theta} \min_{P \in \mathcal{P}_\theta} d(\mathcal{P}_\theta, P_0)$$

- $$P^{*}_{P_0} = \arg \min_{P \in \mathcal{P}} d(P, P_0): \text{pseudo-true distribution}$$

  - When $$P_0 \in \mathcal{P}$$, $$P^{*}_{P_0} = P_0$$
**Pseudo-true set:** $\Theta^*_P = \arg \min_{\theta \in \Theta} \min_{P \in \mathcal{P}_\theta} d(\mathcal{P}_\theta, P_0)$

**vs. point identification:** $\theta^* = \arg \min_{\theta \in \Theta} \min_{P \in \mathcal{P}_\theta} d(\mathcal{P}_\theta, P_0)$
**Assumption**: the pseudo-true distribution \( P^* \) is unique

- this assumption has nothing to do with the identification of the parameter \( \theta \)
- depends on the shape of the model:

\[
P \quad P^* \quad P_0 \quad (=P^*)
\]

- a reasonable assumption made in most papers that allow misspecification
**Assumption**: the pseudo-true distribution $P^*$ is unique.

- This assumption has nothing to do with the identification of the parameter $\theta$.
- It depends on the shape of the model:

  ![Diagram showing the relationship between $P$, $P^*$, and $P_0$](image)

  - $P_0$
  - $P^*$
  - $P$

  $P_0$ ($= P^*$)

- A reasonable assumption made in most papers that allow misspecification.
Solving Minimum KLIC

\[ d(\mathcal{P}, P_0) = \min_{\theta \in \Theta} \min_{P \in \mathcal{P}_\theta} d(\mathcal{P}_\theta, P_0). \]

is solved using a Lagrangian:

\[
\min_{P \in \mathcal{P}_\theta} \int f_{P_0} \log f_{P_0} dP_0 + \gamma' \int m(\cdot, \theta) f_{P_0} dP_0 + \lambda (1 - \int f_{P_0} dP_0) \]

solution:

\[
f^*_P(x; \theta) = \frac{\exp(\gamma^*_P(\theta)' m(x, \theta))}{E_{P_0} \exp(\gamma^*_P(\theta)' m(X_i, \theta))}
\]

\[
\gamma^*_P(\theta) = \arg \min_{\gamma \geq 0} E_{P_0} \exp(\gamma' m(X_i, \theta))
\]

"exponential tilt"
Solving Minimum KLIC

Let $\mathcal{M}_{P_0}(\gamma, \theta) = E_{P_0} \exp(\gamma' m(X_i, \theta))$.

Then,

$$d(\mathcal{P}, P_0) = \min_{\theta \in \Theta} \left[ -\log \mathcal{M}_{P_0}(\gamma, \theta) \right]$$

$$= -\log \left[ \max_{\theta \in \Theta} \min_{\gamma \geq 0} \mathcal{M}_{P_0}(\gamma, \theta) \right]$$

$max_{\theta \in \Theta} \min_{\gamma \geq 0} \mathcal{M}_{P_0}(\gamma, \theta)$:

- is the finite-dimensional dual problem of $\min_{P \in \mathcal{P}} d(P, P_0)$.

Pseudo-true set $\Theta^*_{P_0} = \arg \max_{\theta \in \Theta} \min_{\gamma \geq 0} \mathcal{M}_{P_0}(\gamma, \theta)$
Introduce the Second Model

models:
\[ \mathcal{P} = \bigcup_{\theta \in \Theta} \mathcal{P}_\theta \]
\[ \mathcal{Q} = \bigcup_{\beta \in \mathcal{B}} \mathcal{Q}_\beta \]

submodels:
\[ \mathcal{P}_\theta = \{ P : E_P m(X_i, \theta) \geq 0 \} \]
\[ \mathcal{Q}_\beta = \{ Q : E_Q g(X_i, \beta) \geq 0 \} \]

representation:
\[ f_P : E_P m(X_i, \theta) \]
\[ f_Q : E_Q g(X_i, \beta) \]

moment function:
\[ m : \mathbb{R}^{d_x + d_\theta} \to \mathbb{R}^{d_m} \]
\[ g : \mathbb{R}^{d_x + d_\beta} \to \mathbb{R}^{d_g} \]

parameter:
\[ \theta \in \Theta \]
\[ \beta \in \mathcal{B} \]

Lagrange multiplier:
\[ \gamma \in \mathbb{R}^{d_m}_+ \]
\[ \mu \in \mathbb{R}^{d_g}_+ \]

dual criterion:
\[ \mathcal{M}_{P_0}(\gamma, \theta) \]
\[ \mathcal{N}_{P_0}(\mu, \beta) \]
Test Statistics

- $H_0 : QLR_{P_0} = 0$, where the quasi-likelihood ratio:

$$QLR_{P_0} = \max_{\theta \in \Theta} \min_{\gamma \geq 0} \mathcal{M}_{P_0}(\gamma, \theta) - \max_{\beta \in B} \min_{\mu \geq 0} \mathcal{N}_{P_0}(\mu, \beta)$$

$$= \mathcal{M}_{P_0}(\gamma_{P_0}^*(\theta^*), \theta^*) - \mathcal{N}_{P_0}(\mu_{P_0}^*(\beta^*), \beta^*)$$

- $\gamma_{P_0}^*(\theta)$: minimizer of $\mathcal{M}_{P_0}(\gamma, \theta)$

- $\theta^* \in \Theta_{P_0}^*$, the pseudo-true set of model $P$

- $\mu_{P_0}^*(\beta)$: minimizer of $\mathcal{N}_{P_0}(\mu, \beta)$

- $\beta^* \in B_{P_0}^*$, the pseudo-true set of model $Q$
Test Statistics

- Quasi-likelihood ratio statistic:

\[
\hat{QLR}_n = \hat{M}_n(\hat{\gamma}_n(\hat{\theta}_n), \hat{\theta}_n) - \hat{N}_n(\hat{\mu}_n(\hat{\beta}_n), \hat{\beta}_n)
\]

- \(\hat{M}_n(\gamma, \theta) = n^{-1} \sum_{i=1}^{n} \exp(\gamma' m(W_i, \theta))\)

- \(\hat{\gamma}_n(\theta)\): minimizer of \(\hat{M}_n(\gamma, \theta)\)

- \(\hat{\theta}_n \in \hat{\Theta}_n\), the set of maximizers of \(\hat{M}_n(\hat{\gamma}_n(\theta), \theta)\)

- \(\hat{N}_n(\mu, \beta), \hat{\mu}_n(\beta), \hat{B}_n\) are defined analogously
Theorem

under $H_0$ and regularity conditions:

$$n^{1/2} \sqrt{QLR_n} \rightarrow_d N(0, \omega^2_{P_0}), \text{ where}$$

$$\omega^2_{P_0} = E_{P_0} \left[ \exp\left( \gamma^{*}_{P_0}(\theta^*)'m(X_i, \theta^*) \right) - \exp\left( \mu^{*}_{P_0}(\beta^*)'g(X_i, \beta^*) \right) \right]^2$$

Heuristics:

$$n^{1/2} \sqrt{QLR_n}$$

$$= n^{-1/2} \sum_{i=1}^{n} \left[ \exp(\hat{\gamma}_n(\hat{\theta}_n)'m(X_i, \hat{\theta}_n)) - \exp(\hat{\mu}_n(\hat{\beta}_n)'g(X_i, \hat{\beta}_n)) \right]$$

$$= n^{-1/2} \sum_{i=1}^{n} \left[ \exp(\gamma^{*}_{P_0}(\theta^*)'m(X_i, \theta^*)) - \exp(\mu^{*}_{P_0}(\beta^*)'g(X_i, \beta^*)) \right] + o_p(1)$$
Asymptotic Variance of Q\(\hat{LR}_n\)

\[
\omega^2_{P_0} = E_{P_0}[\exp(\gamma^*_{P_0}(\theta^*)'m(X_i, \theta^*)) - \exp(\mu^*_{P_0}(\beta^*)'g(X_i, \beta^*))]^2
\]

- Under \(H_0\) and the unique pseudo-true distribution assumption:

\[
\omega^2_{P_0} = \exp(-d(\mathcal{P}, P_0)^2) \cdot E_{P_0}[f^*_{\mathcal{P}}(X_i) - f^*_{Q}(X_i)]^2
\]

- \(f^*_{\mathcal{P}}, f^*_{Q}\): density functions of \(P^*_{P_0}\) and \(Q^*_{P_0}\) with respect to \(P_0\)

- \(P^*_{P_0}\) and \(Q^*_{P_0}\): pseudo-true distributions of models \(\mathcal{P}\) and \(\mathcal{Q}\), respectively
Asymptotic Variance of $Q_{\hat{LR}n}$

- Non-overlapping models: $\omega_{P_0}^2 \geq \epsilon > 0$ if $d(P, P_0) \leq C < \infty$

- Overlapping models: $\omega_{P_0}^2 \geq \epsilon > 0$ or $\omega_{P_0}^2 \approx 0$ or $\omega_{P_0}^2 = 0$
Model Selection Test for Non-overlapping Models

- Under $H_0$, 
  \[ n^{1/2} \frac{QLR_n}{\hat{\omega}_n} \to_d N(0, 1) \]

- Our test rejects $H_0$ if 
  \[ n^{1/2} \left| \frac{QLR_n}{\hat{\omega}_n} \right| > z_{\alpha/2} : \]

- \( \hat{\omega}_n^2 = \sup_{\theta \in \hat{\Theta}_n, \beta \in \hat{B}_n} \hat{\omega}_n^2(\theta, \beta) : \)

\[ \hat{\omega}_n^2(\theta, \beta) = n^{-1} \sum_{i=1}^{n} \left[ \exp(\hat{\gamma}_n(\theta) \cdot m(W_i, \theta)) - \exp(\hat{\mu}_n(\beta) \cdot g(W_i, \beta)) \right]^2 \]
Asymptotic Size

**Theorem**

Under suitable assumptions (**non-overlapping**, i.i.d., compact $\Theta$ and $B$, differentiable moment functions), for $\alpha \in (0, 1)$,

$$
\text{AsySZ}^{\text{no}}(\alpha) \equiv \lim_{n \to \infty} \sup_{P_0 \in \mathcal{H}_0^{\text{no}}} \Pr_{P_0}(n^{1/2}|QLR_n|/\hat{\omega}_n > z_{\alpha/2}) = \alpha
$$

$\mathcal{H}_0^{\text{no}}$ : the set of null distribution, $P_0$, that satisfies:

(i) $d(P, P_0) = d(Q, P_0) < C$

(ii) under $P_0$, the pseudo-true distributions of both models are unique

(iii) moment conditions, etc.
Model Selection Test for Overlapping Models

- Under $H_0$, it is possible $\omega^2_{P_0} \geq \varepsilon > 0$ or $\omega^2_{P_0} \approx 0$ or $\omega^2_{P_0} = 0$

- Want the test to have good size and power in finite samples

- Want to learn the distributions of $\hat{QLR}_n$ and $\hat{\omega}^2_n$ in all three cases

- Approximate by asymptotics: derive asymptotic distributions under drifting sequences of $P_{0,n}$, such that $n\omega^2_{P_{0,n}} \to c$

  - localization parameter $c \in [0, \infty]$
Model Selection Test for Overlapping Models

- Difficulty caused by partial identification:
  - the asymptotic distributions of $n^{1/2} \frac{QLR_n}{\hat{\omega}_n}$ and $\hat{\omega}_n^2$ depend on those of $\hat{\theta}_n$ and $\hat{\beta}_n$, when the localization parameter $c < \infty$
  - $\hat{\theta}_n$ and $\hat{\beta}_n$ do not converge in distribution under partial identification

- Solution:
  - Show every sequence $\{\hat{\theta}_n \in \hat{\Theta}_n\}$ converges to $\Theta^*_P_{0,n}$ at $n^{-1/2}$-rate
  - Then, $n\hat{\omega}_n^2 = O_p(1)$ when $c < \infty$
  - $n^{1/2} \frac{QLR_n}{\hat{\omega}_n} \to_p N(0, 1)$ when $c = \infty$
Model Selection Test for Overlapping Models

- Two-step test
- Rejects $H_0$ if

$$n\hat{\omega}_n^2 > b_n \& \ n^{1/2} \left| \tilde{QLR}_n \right| / \hat{\omega}_n > z_{\alpha/2}$$

- $b_n$ is user-chosen, $b_n^{-1} + n^{-1} b_n \rightarrow 0$. e.g. $b_n = 2 \log n$
Asymptotic Size

**Theorem**

*Under suitable conditions (i.i.d., compact $\Theta$ and $B$, differentiable moment functions), for any $\alpha \in (0, 1)$,*

\[
\text{AsySZ}^0 l(\alpha) \equiv \lim_{n \to \infty} \sup_{P_0 \in \mathcal{H}_0^l} \Pr_{P_0} \left( n\hat{\omega}_n^2 > b_n \& n^{1/2} |\widehat{QLR}_n| / \hat{\omega}_n > z_{\alpha/2} \right) \leq \alpha.
\]

*If there exists $P_0 \in \mathcal{H}_0^l$ such that $P_{P_0}^* \neq Q_{P_0}^*$,*

\[
\text{AsySZ}^0 l(\alpha) = \alpha.
\]

- Rules out nested case
- With nested models, the test still can be trusted if it rejects
$H_0^0$: the set of null distribution, $P_0$, that satisfies:

(i) $d(P, P_0) = d(Q, P_0)$

(ii) under $P_0$, the pseudo-true distributions of both models are unique

(iii) moment conditions, etc.
Fixed alternatives: $H_1, \mathcal{Q}: d(\mathcal{P}, P_0) > d(\mathcal{Q}, P_0)$

- The tests choose model $\mathcal{Q}$ with probability approaching one
Local Power- non-overlapping models

- $n^{-1/2}$-local alternatives: $n^{1/2}(d(\mathcal{P}, P_n) - d(Q, P_n)) \to h_1 \in R \setminus \{0\}$

\[ \lim_{n \to \infty} \Pr_{P_n}(n^{1/2} \hat{QLR}_n / \hat{\omega}_n < -z_{\alpha/2}) = \Phi(-z_{\alpha/2} + C_1 \cdot h_1) \]
Local Power- non-overlapping models

- \( n^{-1/2} \)-local alternatives:
  - \( n^{1/2} \omega_{P_n}^{-1}(d(P, P_n) - d(Q, P_n)) \rightarrow h_1^o \in \mathbb{R}\setminus\{0\} \)
  - \( b_n^{1/2} n^{-1/2} \omega_{P_n}^{-1} \rightarrow 0 \)

\[
\lim_{n \to \infty} \Pr_{P_n}(n \hat{\omega}_n > b_n \& n^{1/2} \hat{QLR}_n / \hat{\omega}_n < -z_{\alpha/2}) = \Phi(-z_{\alpha/2} + C_2 \cdot h_1^o)
\]
Simulation Design

- The entry game example

<table>
<thead>
<tr>
<th></th>
<th>Out</th>
<th>In</th>
</tr>
</thead>
<tbody>
<tr>
<td>United</td>
<td>Out</td>
<td>In</td>
</tr>
<tr>
<td></td>
<td>0, 0</td>
<td>0, $\theta_c - \varepsilon_c$</td>
</tr>
<tr>
<td></td>
<td>$\theta_u - \varepsilon_u, 0$</td>
<td>$\theta_u + \Delta_u - \varepsilon_u, \theta_c + \Delta_c - \varepsilon_c$</td>
</tr>
</tbody>
</table>

- $\theta_u = \theta_c = \theta$, $\Delta_u = \Delta_c = \Delta$ (Two-parameter Design)

- $(\varepsilon_u, \varepsilon_c) \sim N(0, I_2)$

- Selection problem: assume complete information

  - $\Delta \geq 0$ (business stealing effect dominates)
  - vs. $\Delta \leq 0$ (positive externality dominates)
Simulation Design

- $\Delta \leq 0$: $(I,O)$ and $(O,I)$ can appear in multiple equilibria:

  \[ P_Y : E_{P_Y}[\bar{p}_j(\theta, \Delta) - 1\{Y = j\}] = 0, \ j = (I, O) \text{ or } (O, I) \\]
  \[ E_{P_Y}[\bar{p}_j(\theta, \Delta) - 1\{Y = j\}] \geq 0, \ j = (I, I) \text{ or } (O, O) \]

- $\Delta \geq 0$: $(I,I)$ and $(O,O)$ can appear in multiple equilibria:

  \[ Q_Y : E_{Q_Y}[\bar{p}_j(\beta, \delta) - 1\{Y = j\}] \geq 0, \ j = (I, O) \text{ or } (O, I) \\]
  \[ E_{Q_Y}[\bar{p}_j(\beta, \delta) - 1\{Y = j\}] = 0, \ j = (I, I) \text{ or } (O, O) \]
Simulation Design

- $P_Y$ and $Q_Y$ are distributions on $Y$
- $Y$ only takes four values: $(I, O)$, $(O, I)$, $(I, I)$ or $(O, O)$
- Models $\mathcal{P}$ and $\mathcal{Q}$ in picture:

\[ H_0: d(\mathcal{P}, P_0) = d(\mathcal{Q}, P_0) \]
Simulation Result

Overlapping test, $b_n = 2 \log n$

<table>
<thead>
<tr>
<th>True DGP: θ = 0.5, Δ = −0.25</th>
<th>θ = 0.5, Δ = 0.25</th>
</tr>
</thead>
<tbody>
<tr>
<td>in multiple equilibria:</td>
<td>(I, O) always selected</td>
</tr>
<tr>
<td>$n = 500$</td>
<td>(0.5058, 0)</td>
</tr>
<tr>
<td>$n = 800$</td>
<td>(0.7596, 0)</td>
</tr>
<tr>
<td>$n = 1000$</td>
<td>(0.8850, 0)</td>
</tr>
<tr>
<td>$n = 1500$</td>
<td>(0.9780, 0)</td>
</tr>
</tbody>
</table>

- Number of repetitions: 5000
Summary

- Proposed model selection tests for moment inequality models
- Showed the tests have correct asymptotic size
- Showed the tests have power against fixed alternatives and $n^{-1/2}$-local alternatives
- Demonstrated by simulation that the tests choose the right model with high probability even in small samples
- ...the probability approaches one as sample size gets large