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) ($\Delta_u \leq 0$, $\Delta_c \leq 0$)

		Continental	
		Out	In
United	Out	$0, 0$	$0, \theta_c - \varepsilon_c$
	In	$\theta_u - \varepsilon_u, 0$	$\theta_u + \Delta_u - \varepsilon_u, \theta_c + \Delta_c - \varepsilon_c$

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

- Multiple equilibria when

$$\theta_u - \varepsilon_u \geq 0 \geq \theta_u + \Delta_u - \varepsilon_u$$
$$\theta_c - \varepsilon_c \geq 0 \geq \theta_c + \Delta_c - \varepsilon_c$$

- 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; θ: 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; θ: 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

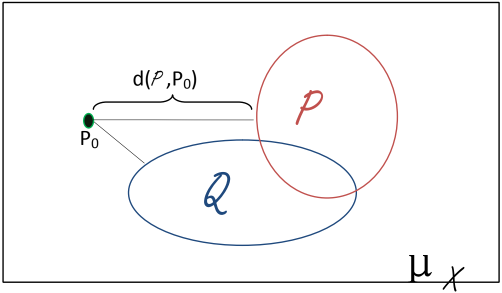
- Parameter θ typically not point-identified
- Estimate the identified set $\Theta^*_{P_0}$: $\{\theta : E_{P_0}[m(X_i, \theta)] \geq 0\}$
  - GMM-type: Chernozhukov, Hong and Tamer (2007), Andrews and Guggenberger (2009), Andrews and Soares (2010), Bugni (2008), Romano and Shaikh (2009)
  - Empirical Likelihood: Canay (2008)
  - *Exponential Tilting*: this paper

Model Selection Problem

- 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

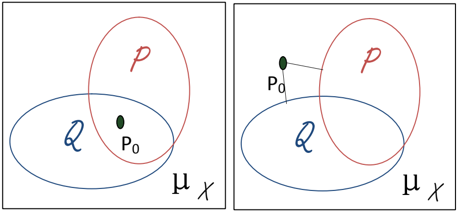
- $\mathcal{P} = \{P : E_P m(X_i, \theta) \geq 0, \ \theta \in \Theta\}$
- $\mathcal{Q} = \{Q : E_Q g(X_i, \beta) \geq 0, \ \beta \in B\}$



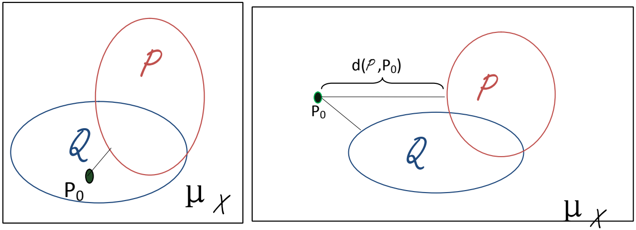
- $d(\mathcal{P}, P_0)$ "distance" from the true distribution P_0 to $\mathcal{P}$
- model $\mathcal{P}$ correctly specified iff $P_0 \in \mathcal{P}$ or $d(\mathcal{P}, P_0) = 0$

Model Selection Test

- $H_0 : d(\mathcal{P}, P_0) = d(\mathcal{Q}, P_0)$



- $H_{1,\mathcal{Q}} : d(\mathcal{P}, P_0) > d(\mathcal{Q}, P_0)$



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)

Literature

- 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:

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

Measure the Goodness-of-fit

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

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

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

Measure the Goodness-of-fit

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

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

  - 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:

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

  - 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

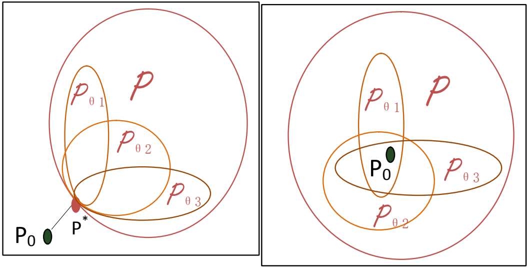
- Why Kullback-Leibler?
  - Minimizing KLIC $\Leftrightarrow$ Maximum Likelihood in parametric models
  - Standard AIC, BIC procedures use KLIC
  - Compared to Kolmogorov-Smirnov distance: more tractable
  - Compared to GMM-type criteria: automatic weighting
- $d(P, P_0)$ (exponential tilting) or $d(P_0, P)$ (empirical likelihood)?
  - Exponential tilting is robust to misspecification
  - Results extend to empirical likelihood, when moment function bounded

Minimum KLIC - Pseudo-true Distribution

- Rewrite

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

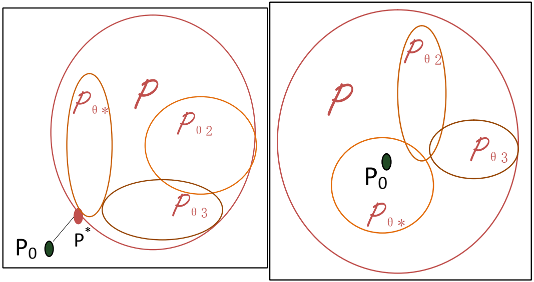
- $d(\mathcal{P}, P_0) = \min_{P\in\mathcal{P}} d(P, P_0) = \min_{\theta\in\Theta}\min_{P\in\mathcal{P}_\theta} d(\mathcal{P}_\theta, P_0)$



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

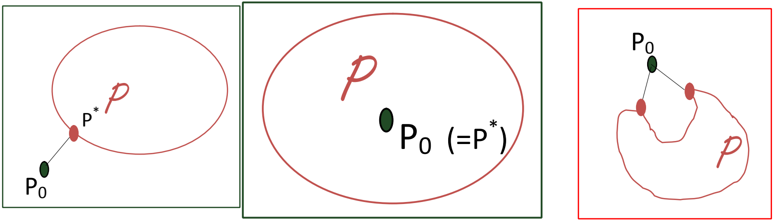
Minimum KLIC - Pseudo-true Set

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



Minimum KLIC - Uniqueness of Pseudo-true Distribution

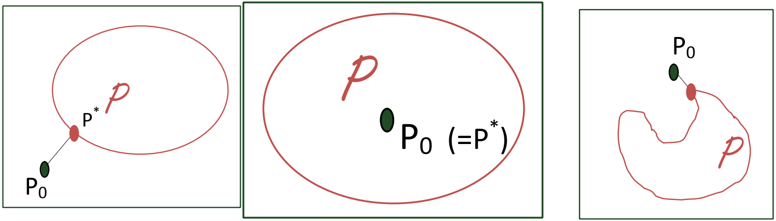
- **Assumption**: the pseudo-true distribution P^* is unique
  - this assumption has nothing to do with the identification of the parameter θ
  - depends on the shape of the model:



  - a reasonable assumption made in most papers that allow misspecification

Minimum KLIC - Uniqueness of Pseudo-true Distribution

- **Assumption**: the pseudo-true distribution P^* is unique
  - this assumption has nothing to do with the identification of the parameter θ
  - depends on the shape of the model:



  - a reasonable assumption made in most papers that allow misspecification

Solving Minimum KLIC

- $d(\mathcal{P}, P_0) = \min_{\theta \in \Theta} \boxed{\min_{P \in \mathcal{P}_\theta} d(\mathcal{P}_\theta, P_0)}$.
- $\Box$ 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:

$$f^*_{\mathcal{P}}(x; \theta) = \frac{\exp(\gamma^*_{P_0}(\theta)' m(x, \theta))}{E_{P_0} \exp(\gamma^*_{P_0}(\theta)' m(X_i, \theta))}$$

  - $\gamma^*_{P_0}(\theta) = \arg\min_{\gamma \geq 0} E_{P_0} \exp(\gamma' m(X_i, \theta))$
  - "exponential tilt"

Solving Minimum KLIC

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

- Then,

$$\begin{aligned} d(\mathcal{P}, P_0) &= \min_{\theta \in \Theta} [-\log \mathcal{M}_{P_0}(\gamma, \theta)] \\ &= -\log[\max_{\theta \in \Theta} \min_{\gamma \geq 0} \mathcal{M}_{P_0}(\gamma, \theta)] \end{aligned}$$

- $\max_{\theta \in \Theta} \min_{\gamma \geq 0} \mathcal{M}_{P_0}(\gamma, \theta)$:
  - 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:	$\mathcal{P} = \cup_{\theta \in \Theta} \mathcal{P}_\theta$	$\mathcal{Q} = \cup_{\beta \in B} \mathcal{Q}_\beta$
submodels: representation:	$\mathcal{P}_\theta = \{P : E_P m(X_i, \theta) \geq 0\}$	$\mathcal{Q}_\beta = \{Q : E_Q g(X_i, \beta) \geq 0\}$
moment function:	$m : R^{d_x + d_\theta} \longrightarrow R^{d_m}$	$g : R^{d_x + d_\beta} \longrightarrow R^{d_g}$
parameter:	$\theta \in \Theta$	$\beta \in B$
Lagrange multiplier:	$\gamma \in R_+^{d_m}$	$\mu \in R_+^{d_g}$
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:

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

  - $\gamma^*_{P_0}(\theta)$: minimizer of $\mathcal{M}_{P_0}(\gamma,\theta)$
  - $\theta^* \in \Theta^*_{P_0}$, the pseudo-true set of model $\mathcal{P}$
  - $\mu^*_{P_0}(\beta)$: minimizer of $\mathcal{N}_{P_0}(\mu,\beta)$
  - $\beta^* \in B^*_{P_0}$, the pseudo-true set of model $\mathcal{Q}$

Test Statistics

- Quasi-likelihood ratio statistic:

$$\widehat{QLR}_n = \hat{\mathcal{M}}_n(\hat{\gamma}_n(\hat{\theta}_n), \hat{\theta}_n) - \hat{\mathcal{N}}_n(\hat{\mu}_n(\hat{\beta}_n), \hat{\beta}_n)$$

  - $\hat{\mathcal{M}}_n(\gamma, \theta) = n^{-1} \sum_{i=1}^{n} \exp(\gamma' m(W_i, \theta))$
  - $\hat{\gamma}_n(\theta)$: minimizer of $\hat{\mathcal{M}}_n(\gamma, \theta)$
  - $\hat{\theta}_n \in \hat{\Theta}_n$, the set of maximizers of $\hat{\mathcal{M}}_n(\hat{\gamma}_n(\theta), \theta)$
  - $\hat{\mathcal{N}}_n(\mu, \beta)$, $\hat{\mu}_n(\beta)$, $\hat{B}_n$ are defined analogously

Test Statistics

Theorem

under H_0 and regularity conditions:

$$n^{1/2}\widehat{QLR}_n \to_d N(0, \omega^2_{P_0}), \text{ where}$$

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

Heuristics:

$$\begin{aligned} & n^{1/2}\widehat{QLR}_n \\ = \; & n^{-1/2}\sum_{i=1}^{n}[\exp(\hat{\gamma}_n(\hat{\theta}_n)'m(X_i, \hat{\theta}_n)) - \exp(\hat{\mu}_n(\hat{\beta}_n)'g(X_i, \hat{\beta}_n))] \\ = \; & n^{-1/2}\sum_{i=1}^{n}[\exp(\gamma^*_{P_0}(\theta^*)'m(X_i, \theta^*)) - \exp(\mu^*_{P_0}(\beta^*)'g(X_i, \beta^*))] + o_p(1) \end{aligned}$$

Asymptotic Variance of $Q\hat{L}R_n$

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

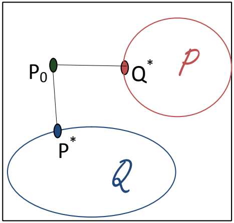
- Under H_0 and the unique pseudo-true distribution assumption:

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

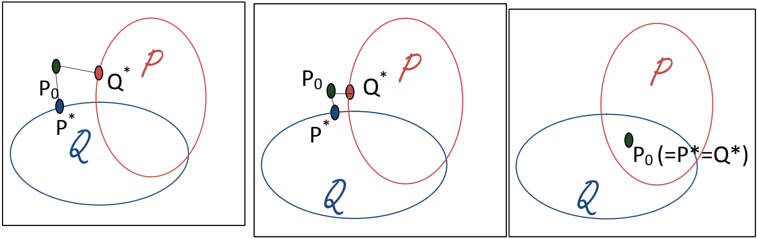
  - $f^*_{\mathcal{P}}$, $f^*_{\mathcal{Q}}$: density functions of $P^*_{P_0}$ and $Q^*_{P_0}$ with respect to P_0
  - $P^*_{P_0}$ and $Q^*_{P_0}$: pseudo-true distributions of models $\mathcal{P}$ and $\mathcal{Q}$, respectively

Asymptotic Variance of $\hat{QLR}n$

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



- Overlapping models: $\omega^2_{P_0} \geq \varepsilon > 0$ or $\omega^2_{P_0} \approx 0$ or $\omega^2_{P_0} = 0$

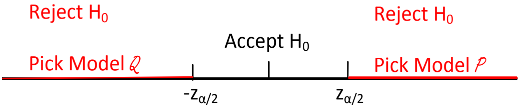


Model Selection Test for Non-overlapping Models

- Under H_0,

$$n^{1/2}\widehat{QLR}_n/\hat{\omega}_n \to_d N(0,1)$$

- Our test rejects H_0 if $n^{1/2}|\widehat{QLR}_n|/\hat{\omega}_n > z_{\alpha/2}$:



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

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

Asymptotic Size

Theorem

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

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

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

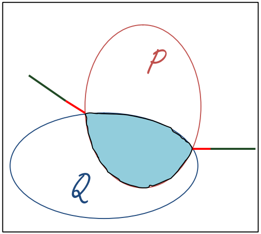
(i) $d(P, P_0) = d(Q, P_0) < C$

(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 $\widehat{QLR}_n$ 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 $n^{1/2}\widehat{QLR}_n/\hat{\omega}_n$ 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 $c < \infty$
  - $n^{1/2}\widehat{QLR}_n/\hat{\omega}_n \to_p N(0,1)$ when $c = \infty$

Model Selection Test for Overlapping Models

- Two-step test
- Rejects H_0 if

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

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

Asymptotic Size

Theorem

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

$$\begin{aligned} & AsySZ^{ol}(\alpha) \\ \equiv & \lim_{n\to\infty} \sup_{P_0 \in \mathcal{H}_0^{ol}} \Pr_{P_0}\left(n\hat{\omega}_n^2 > b_n \ \& \ n^{1/2}|\widehat{QLR}_n|/\hat{\omega}_n > z_{\alpha/2}\right) \leq \alpha. \end{aligned}$$

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

$$AsySZ^{ol}(\alpha) = \alpha.$$

- $\square$ rules out nested case
- With nested models, the test still can be trusted if it rejects

Asymptotic Size

$\mathcal{H}_0^{ol}$: the set of null distribution, P_0, that satisfies:

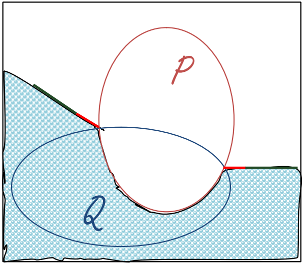
(i) $d(P, P_0) = d(Q, P_0)$

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

(iii) moment conditions, etc.

Power of the Tests

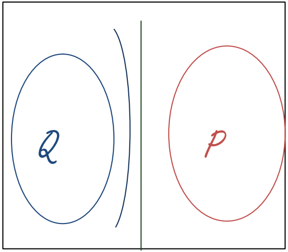
- Fixed alternatives: $H_{1,\mathcal{Q}}$: $d(\mathcal{P}, P_0) > d(\mathcal{Q}, P_0)$



- The tests choose model $\mathcal{Q}$ with probability approaching one

Local Power- non-overlapping models

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

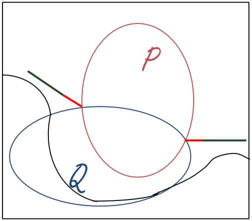


Theorem

$\lim_{n\to\infty} \Pr_{P_n}(n^{1/2}\widehat{QLR}_n/\hat{\omega}_n < -z_{\alpha/2}) = \Phi(-z_{\alpha/2} + C_1 \cdot h_1)$

Local Power- non-overlapping models

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



Theorem

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

Simulation Design

- The entry game example

		Continental	
		Out	In
United	Out	0, 0	0, $\theta_c - \varepsilon_c$
	In	$\theta_u - \varepsilon_u$, 0	$\theta_u + \Delta_u - \varepsilon_u$, $\theta_c + \Delta_c - \varepsilon_c$

- $\theta_u = \theta_c = \theta$, $\Delta_u = \Delta_c = \Delta$ (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

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

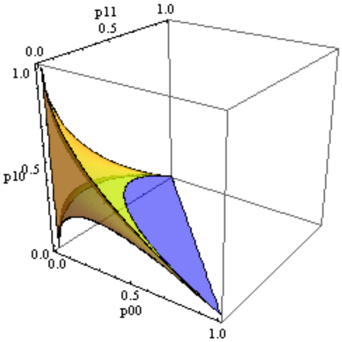
$$\mathcal{P} = \left\{ \begin{array}{c} P_Y : E_{P_Y}[\overline{p}_j(\theta, \Delta) - 1\{Y = j\}] = 0, \, j = (I, O) \text{ or } (O, I) \\ E_{P_Y}[\overline{p}_j(\theta, \Delta) - 1\{Y = j\}] \ge 0, \, j = (I, I) \text{ or } (O, O) \end{array} \right\}$$

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

$$\mathcal{Q} = \left\{ \begin{array}{c} Q_Y : E_{Q_Y}[\overline{p}_j(\beta, \delta) - 1\{Y = j\}] \ge 0, \, j = (I, O) \text{ or } (O, I) \\ E_{Q_Y}[\overline{p}_j(\beta, \delta) - 1\{Y = j\}] = 0, \, j = (I, I) \text{ or } (O, O) \end{array} \right\}$$

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:



- $\mathcal{P}$ and $\mathcal{Q}$ are overlapping

Simulation Result

Overlapping test, $b_n = 2 \log n$

Rejection Probability (Prob. of choosing $\mathcal{P}$, Prob. of choosing $\mathcal{Q}$)

True DGP: in multiple equilibria:	$\theta = 0.5$, $\Delta = -0.25$ (I, O) always selected	$\theta = 0.5$, $\Delta = 0.25$ (I, I) always selected
$n = 500$	(0.5058, 0)	(0.0002, 0.3246)
$n = 800$	(0.7596, 0)	(0, 0.6154)
$n = 1000$	(0.8850, 0)	(0, 0.7584)
$n = 1500$	(0.9780, 0)	(0, 0.9620)

- 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