Practical and Theoretical Advances for Inference in Partially Identified Models

by

Azeem M. Shaikh, University of Chicago

August 2015

amshaikh@uchicago.edu

Collaborator:
Ivan Canay, Northwestern University
Introduction

Partially Identified Models:

– Param. of interest is not uniquely determined by distr. of obs. data.
– Instead, limited to a set as a function of distr. of obs. data.
  (i.e., the identified set)
– Due largely to pioneering work by C. Manski, now ubiquitous.
  (many applications!)

Inference in Partially Identified Models:

– Focused mainly on the construction of confidence regions.
– Most well-developed for moment inequalities.
– Important practical issues remain subject of current research.
Outline of Talk

1. Definition of partially identified models
2. Confidence regions for partially identified models
   - Importance of uniform asymptotic validity
3. Moment inequalities
   - Common framework to describe five distinct approaches
4. Subvector inference for moment inequalities
5. More general framework
   - Unions of functional moment inequalities
Partially Identified Models

Obs. data $X \sim P \in \mathbf{P} = \{P_\gamma : \gamma \in \Gamma\}$.

($\gamma$ is possibly infinite-dim.)

Identified set for $\gamma$:

$$\Gamma_0(P) = \{\gamma \in \Gamma : P_\gamma = P\}.$$

Typically, only interested in $\theta = \theta(\gamma)$.

Identified set for $\theta$:

$$\Theta_0(P) = \{\theta(\gamma) \in \Theta : \gamma \in \Gamma_0(P)\},$$

where $\Theta = \theta(\Gamma)$. 
Partially Identified Models (cont.)

θ is identified relative to P if

\[ \Theta_0(P) \] is a singleton for all \( P \in \mathbf{P} \).

θ is unidentified relative to P if

\[ \Theta_0(P) = \Theta \] for all \( P \in \mathbf{P} \).

Otherwise, θ is partially identified relative to P.

\( \Theta_0(P) \) has been characterized in many examples ...

... can often be characterized using moment inequalities.
Confidence Regions

If \( \theta \) is identified relative to \( \mathbf{P} \) (so, \( \theta = \theta(P) \)), then we require that

\[
\liminf_{n \to \infty} \inf_{P \in \mathbf{P}} P\{\theta(P) \in C_n\} \geq 1 - \alpha .
\]

Now we require that

\[
\liminf_{n \to \infty} \inf_{P \in \mathbf{P}} \inf_{\theta \in \Theta_0(P)} P\{\theta \in C_n\} \geq 1 - \alpha .
\]

Refer to as conf. region for points in id. set unif. consistent in level.

**Remark:** May also be interested in conf. regions for identified set itself:

\[
\liminf_{n \to \infty} \inf_{P \in \mathbf{P}} P\{\Theta_0(P) \subseteq C_n\} \geq 1 - \alpha .
\]

See Chernozkukov et al. (2007) and Romano & Shaikh (2010).
Confidence Regions (cont.)

Unif. consistency in level vs. pointwise consistency in level, i.e.,

\[ \liminf_{n \to \infty} P\{\theta \in C_n\} \geq 1 - \alpha \text{ for all } P \in \mathcal{P} \text{ and } \theta \in \Theta_0(P). \]

May be for every \( n \) there is \( P \in \mathcal{P} \) and \( \theta \in \Theta_0(P) \) with cov. prob. \( \ll 1 - \alpha \).

In well-behaved prob., distinction is entirely technical issue.

(e.g., conf. regions for the univariate mean with i.i.d. data.)

In less well-behaved prob., distinction is more important.

(e.g., conf. regions in even simple partially id. models!)

Some “natural” conf. reg. may need to restrict \( \mathcal{P} \) in non-innocuous ways.

(e.g., may need to assume model is “far” from identified.)

**Moment Inequalities**

Henceforth, $W_i, i = 1, \ldots, n$ are i.i.d. with common marg. distr. $P \in \mathbf{P}$. Numerous ex. of partially identified models give rise to mom. ineq., i.e.,

\[ \Theta_0(P) = \{ \theta \in \Theta : E_P[m(W_i, \theta)] \leq 0 \} , \]

where $m$ takes values in $\mathbf{R}^k$.

**Goal:** Conf. reg. for points in the id. set that are unif. consistent in level.

**Remark:** Assume throughout mild uniform integrability condition ...

... ensures CLT and LLN hold unif. over $P \in \mathbf{P}$ and $\theta \in \Theta_0(P)$. 
**Moment Inequalities (cont.)**

**How:** Construct tests $\phi_n(\theta)$ of

$$H_\theta : E_P[m(W_i, \theta)] \leq 0$$

that provide unif. asym. control of Type I error, i.e.,

$$\limsup_{n \to \infty} \sup_{P \in \mathcal{P}} \sup_{\theta \in \Theta_0(P)} E_P[\phi_n(\theta)] \leq \alpha .$$

Given such $\phi_n(\theta)$,

$$C_n = \{ \theta \in \Theta : \phi_n(\theta) = 0 \}$$

satisfies desired coverage property.

Below describe **five different tests**, all of form

$$\phi_n(\theta) = I\{T_n(\theta) > \hat{c}_n(\theta, 1 - \alpha)\} .$$
Moment Inequalities (cont.)

Some Notation:

\[ \mu(\theta, P) = E_P[m(W_i, \theta)]. \]
\[ \bar{m}_n(\theta) = \text{sample mean of } m(W_i, \theta). \]
\[ \hat{\Omega}_n(\theta) = \text{sample correlation of } m(W_i, \theta). \]
\[ \sigma_j^2(\theta, P) = \text{Var}_P[m_j(W_i, \theta)]. \]
\[ \hat{\sigma}_{n,j}^2(\theta) = \text{sample variance of } m_j(W_i, \theta). \]
\[ \hat{D}_n(\theta) = \text{diag}(\hat{\sigma}_{n,1}(\theta), \ldots, \hat{\sigma}_{n,k}(\theta)). \]
Moment Inequalities (cont.)

Test Statistic:

In all cases,

\[ T_n(\theta) = T(\hat{D}_n^{-1}(\theta)\sqrt{n}\hat{m}_n(\theta), \hat{\Omega}_n(\theta)) \]

for an appropriate choice of \( T(x, V) \), e.g.,

- modified method of moments: \( \sum_{1 \leq j \leq k} \max\{x_j, 0\}^2 \)
- maximum: \( \max_{1 \leq j \leq k} \max\{x_j, 0\} \)
- quasi-likelihood ratio: \( \inf_{t \leq 0} (x - t)'V^{-1}(x - t) \)

Main requirement is that \( T \) weakly increasing in first argument.
Moment Inequalities (cont.)

Critical Value:

Useful to define

\[ J_n(x, s(\theta), \theta, P) = P \left\{ T(\hat{D}_n^{-1}(\theta)Z_n(\theta) + \hat{D}_n^{-1}(\theta)s(\theta), \hat{\Omega}_n(\theta)) \leq x \right\}, \]

where

\[ Z_n(\theta) = \sqrt{n}(\bar{m}_n(\theta) - \mu(\theta, P)), \]

which is easy to estimate.

On the other hand,

\[ J_n(x, \sqrt{n}\mu(\theta, P), \theta, P) = P\{T_n(\theta) \leq x\} \]

is difficult to estimate. See, e.g., Andrews (2000).

Indeed, not even possible to estimate \( \sqrt{n}\mu(\theta, P) \) consistently!

Five diff. tests distinguished by how they circumvent this problem.
Moment Inequalities (cont.)

Test #1: Least Favorable Tests:

Main Idea: \( \sqrt{n} \mu(\theta, P) \leq 0 \) for any \( P \in \mathcal{P} \) and \( \theta \in \Theta_0(P) \)

\[ \Rightarrow J_n^{-1}(1 - \alpha, \sqrt{n} \mu(\theta, P), \theta, P) \leq J_n^{-1}(1 - \alpha, 0, \theta, P) \, . \]

Choosing

\[ \hat{c}_n(1 - \alpha, \theta) = \text{estimate of } J_n^{-1}(1 - \alpha, 0, \theta, P) \]

therefore leads to valid tests.


Moment Inequalities (cont.)

Test #1: Least Favorable Tests (cont.):

**Remark:** Deemed “conservative,” but criticism not entirely fair:
- In Gaussian setting, these tests are \((\alpha\text{- and } d\text{-})\) admissible.
- Some are even maximin **optimal** among restricted class of tests.
- See Lehmann (1952) and Romano & Shaikh (unpublished).

Nevertheless, unattractive:
- Tend to have best power against alternatives with *all* moments > 0.
- As \(\theta\) varies, many alternatives with only *some* moments > 0.
- May therefore not lead to smallest confidence regions.

Following tests incorporate info. about \(\sqrt{n}\mu(\theta, P)\) in some way.
\[\implies \text{better power against such alternatives.}\]
Moment Inequalities (cont.)

Test #2: Subsampling:

See Politis & Romano (1994).

**Main Idea:** Fix \( b = b_n < n \) with \( b \to \infty \) and \( b/n \to 0 \).

Compute \( T_n(\theta) \) on each of \( \binom{n}{b} \) subsamples of data.

Denote by \( L_n(x, \theta) \) the empirical distr. of these quantities.

Use \( L_n(x, \theta) \) as estimate of distr. of \( T_n(\theta) \), i.e.,

\[
J_n(x, \sqrt{n} \mu(\theta, P), \theta, P).
\]

Choosing

\[
\hat{c}_n(1 - \alpha, \theta) = L_n^{-1}(1 - \alpha, \theta)
\]

leads to valid tests.

Moment Inequalities (cont.)

Test #2: Subsampling (cont.):

**Why:** $L_n(x, \theta)$ is a “good” estimate of distr. of $T_b(\theta)$, i.e.,

$$J_b(x, \sqrt{b}\mu(\theta, P), \theta, P).$$

See general results in Romano & Shaikh (2012).

Moreover,

$$\sqrt{n}\mu(\theta, P) \leq \sqrt{b}\mu(\theta, P)$$

for any $P \in \mathcal{P}$ and $\theta \in \Theta_0(P)$

$$\implies J_n^{-1}(1 - \alpha, \sqrt{n}\mu(\theta, P), \theta, P) \leq J_n^{-1}(1 - \alpha, \sqrt{b}\mu(\theta, P), \theta, P).$$

Desired conclusion follows.

**Remark:** Incorporates information about $\sqrt{n}\mu(\theta, P)$ ...

... but remains unattractive because *choice of $b$ problematic.*
Test #3: Generalized Moment Selection:

See Andrews & Soares (2010).

**Main Idea:** Perhaps possible to estimate $\sqrt{n} \mu(\theta, P)$ “well enough”?

Consider, e.g., $\hat{s}^{gms}_n(\theta) = (\hat{s}^{gms}_{n,1}(\theta), \ldots, \hat{s}^{gms}_{n,k}(\theta))'$ with

$$\hat{s}^{gms}_{n,j}(\theta) = \begin{cases} 0 & \text{if } \frac{\sqrt{n} \tilde{m}_{n,j}(\theta)}{\tilde{\sigma}_{n,j}(\theta)} > -\kappa_n, \\ -\infty & \text{otherwise} \end{cases}$$

where $0 < \kappa_n \to \infty$ and $\kappa_n / \sqrt{n} \to 0$.

Choosing

$$\hat{c}_n(1 - \alpha, \theta) = \text{estimate of } J_n^{-1}(1 - \alpha, \hat{s}^{gms}_n(\theta), \theta, P)$$

leads to valid tests.
**Moment Inequalities (cont.)**

**Test #3: Generalized Moment Selection (cont.):**

**Why:** For any sequence $P_n \in \mathbf{P}$ and $\theta_n \in \Theta_0(P_n)$

$$\hat{s}_{n,j}^{\text{gms}}(\theta_n) = \begin{cases} 0 & \text{if } \sqrt{n}\mu_j(\theta_n, P_n) \to c \leq 0 \\ -\infty & \text{if } \sqrt{n}\mu_j(\theta_n, P_n) \to -\infty \end{cases} \text{ w.p.a.1.}$$

In this sense, $\hat{s}_n^{\text{gms}}(\theta)$ provides an asymptotic upper bound on $\sqrt{n}\mu(\theta, P)$.

**Remark:** Also incorporates information about $\sqrt{n}\mu(\theta, P)$ ...

... and, for typical $\kappa_n$ and $b$, more powerful than subsampling.

Main drawback is choice of $\kappa_n$:

- In finite-samples, smaller choice always more powerful.
- First- and higher-order properties do not depend on $\kappa_n$.

See Bugni (2014).

- Precludes data-dependent rules for choosing $\kappa_n$. 

18
Test #4: Refined Moment Selection:

See Andrews & Barwick (2012).

**Main Idea:** In order to develop data-dep. rules for choosing $\kappa_n$, ...

... change asymp. framework so $\kappa_n$ does not depend on $n$.

Consider, e.g., $\hat{s}_{n}^{\text{rms}}(\theta) = (\hat{s}_{n,1}^{\text{rms}}(\theta), \ldots, \hat{s}_{n,k}^{\text{rms}}(\theta))'$ with

$$
\hat{s}_{n,j}^{\text{rms}}(\theta) = \begin{cases} 
0 & \text{if } \frac{\sqrt{n}\hat{m}_{n,j}(\theta)}{\hat{\sigma}_{n,j}(\theta)} > -\kappa \\
-\infty & \text{otherwise}
\end{cases}.
$$

Note $\hat{s}_{n}^{\text{rms}}(\theta)$ no longer an asymp. upper bound on $\sqrt{n}\mu(\theta, P)$, so ...

... critical value replacing $\hat{s}_{n}^{\text{gms}}(\theta)$ with $\hat{s}_{n}^{\text{rms}}(\theta)$ is too small.

For appropriate size-corr. factor $\hat{\eta}_n(\theta) > 0$, choosing

$$
\hat{c}_n(1 - \alpha, \theta) = \text{estimate of } J_n^{-1}(1 - \alpha, \hat{s}_{n}^{\text{rms}}(\theta), \theta, P) + \hat{\eta}_n(\theta)
$$

leads to valid tests (whose first-order properties depend on $\kappa$.)
Moment Inequalities (cont.)

Test #4: Refined Moment Selection (cont.):

**Remark:** Incorporates information about $\sqrt{n}\mu(\theta, P) \ldots$

... in asymp. framework where first-order prop. depend on $\kappa$.

Main drawback is computation of $\hat{\eta}_n(\theta)$:

- Requires approx. max. rejection probability over $k$-dim. space.
- Andrews & Barwick (2012) examine $2^{k-1} - 1$ extreme points.
- Provide numerical evidence in favor of this simplification.
- Some results in McCloskey (2015).
- Even so, remains computationally infeasible for $k > 10$.

Precludes many applications, e.g.,

- Bajari, Benkard & Levin (2007) ($k \approx 500$ or more!)
- Ciliberto & Tamer (2009) ($k = 2^{m+1}$ where $m = \#$ of firms).
Test #5: Two-Step Tests:

See Romano, Shaikh & Wolf (2014).

Main Idea:

Step 1: Construct conf. region for $\sqrt{n}\mu(\theta, P)$, i.e., $M_n(1 - \beta, \theta)$ s.t.

$$\liminf_{n \to \infty} \inf_{P \in \mathcal{P}} \inf_{\theta \in \Theta_0(P)} \{ \sqrt{n}\mu(\theta, P) \in M_n(1 - \beta, \theta) \} \geq 1 - \beta,$$

where $0 < \beta < \alpha$.

An upper-right rect. conf. reg. is computationally attractive, i.e.,

$$M_n(1 - \beta, \theta) = \left\{ \mu \in \mathbb{R}^k : \mu_j \leq \bar{m}_{n,j}(\theta) + \frac{\hat{\sigma}_{n,j}(\theta)\hat{q}_n(1 - \beta, \theta)}{\sqrt{n}} \right\},$$

where $\hat{q}_n(1 - \beta, \theta)$ may be easily constructed using, e.g., bootstrap.
Moment Inequalities (cont.)

Test #5: Two-Step Tests:

Main Idea (cont.):

Step 2: Use $M_n(1 - \beta, \theta)$ to restrict possible values for $\sqrt{n}\mu(\theta, P)$.

Consider “largest” $s \leq 0$ with $s \in M_n(1 - \beta, \theta)$, i.e.,

$$\hat{s}_{n,\text{ts}}(\theta) = (\hat{s}_{n,1}(\theta), \ldots, \hat{s}_{n,k}(\theta))^t$$

with

$$\hat{s}_{n,j}(\theta) = \min\{\sqrt{n}\hat{m}_{n,j}(\theta) + \hat{\sigma}_{n,j}(\theta)\hat{q}_n(1 - \beta, \theta), 0\}.$$

Choosing

$$\hat{c}_n(1 - \alpha, \theta) = \text{estimate of } J_n^{-1}(1 - \alpha + \beta, \hat{s}_{n,\text{ts}}(\theta), \theta, P),$$

leads to valid tests (whose first-order properties depend on $\beta$).

Closed-form expression for $\hat{s}_{n,\text{ts}}(\theta)$ a key feature!
Moment Inequalities (cont.)

Test #5: Two-Step Tests (cont.):

Why: Argument hinges on simple Bonferroni-type inequality.

Remark: Also incorporates information about $\sqrt{n}\mu(\theta, P)$ ...

... in asymp. framework where first-order prop. depend on $\beta$.

But, importantly:

– Remains feasible even for large values of $k$.
– Despite “crudeness” of ineq., remains competitive in terms of power.

Many earlier antecedents:

– In economics, e.g., Stock & Staiger (1997) and McCloskey (2012).
– Computational simplicity key novelty here.
Subvector Inference for Moment Inequalities

Despite advances, methods not commonly employed.

Methods difficult (infeasible?) when $\dim(\theta)$ even moderately large ...

... but interest often only in few coord. of $\theta$ (or a fcn. of $\theta$)!

Let $\lambda(\cdot) : \Theta \to \Lambda$ be function of $\theta$ of interest.

Identified set for $\lambda(\theta)$ is

$$\Lambda_0(P) = \lambda(\Theta_0(P)) = \{ \lambda(\theta) : \theta \in \Theta_0(P) \} ,$$

where

$$\Theta_0(P) = \{ \theta \in \Theta : E_P[m(W_i, \theta)] \leq 0 \} .$$

**Goal:** Conf. reg. for points in id. set that are unif. consistent in level.

**Remark:** Methods require same assumptions plus possibly others.
Subvector Inference for Moment Inequalities (cont.)

**How:** Construct tests $\phi_n(\lambda)$ of

$$H_\lambda : \exists \theta \in \Theta \text{ with } E_P[m(W_i, \theta)] \leq 0 \text{ and } \lambda(\theta) = \lambda$$

that provide unif. asym. control of Type I error, i.e.,

$$\limsup_{n \to \infty} \sup_{P \in \mathcal{P}} \sup_{\lambda \in \Lambda_0(P)} E_P[\phi_n(\lambda)] \leq \alpha .$$

Given such $\phi_n(\lambda)$,

$$C_n = \{ \lambda \in \Lambda : \phi_n(\lambda) = 0 \}$$

satisfies desired coverage property.

Below describe three different tests.
Subvector Inference for Moment Inequalities (cont.)

Test #1: Projection:

Main Idea: Utilize previous tests $\phi_n(\theta)$:

$$\phi_n^{\text{proj}}(\lambda) = \inf_{\theta \in \Theta_\lambda} \phi_n(\theta),$$

where

$$\Theta_\lambda = \{\theta \in \Theta : \lambda(\theta) = \lambda\}.$$  

Properties of $\phi_n(\theta)$ imply this is a valid test.

Remark: As noted by Romano & Shaikh (2008) ...

... generally conservative, i.e., may severely over cover $\lambda(\theta)$.

Computationally difficult when $\dim(\theta)$ large.

Related work by Kaido, Molinari & Stoye (in progress) ...

... adjust critical value in $\phi_n(\theta)$ to avoid over-coverage.
Subvector Inference for Moment Inequalities (cont.)

Test #2: Subsampling:


Main Idea: Reject $H_\lambda$ for large values of profiled test statistic:

$$T_n^{\text{prof}}(\lambda) = \inf_{\theta \in \Theta_\lambda} T_n(\theta),$$

where $T_n(\theta)$ is one of test statistics from before.

Use subsampling to estimate distribution of $T_n^{\text{prof}}(\lambda)$.

High-level conditions for validity given by Romano & Shaikh (2008).

Remark: Less conservative than proj., but choice of $b$ problematic.
Subvector Inference for Moment Inequalities (cont.)

Test #3: Minimum Resampling:

See Bugni, Canay & Shi (2014).

Also rejects for large values of $T_n^{\text{prof}}(\lambda)$.

In order to describe critical value, useful to define

$$J_n(x, \Theta_\lambda, s(\cdot), \lambda, P) = P\left\{ \inf_{\theta \in \Theta_\lambda} T(\hat{D}_n^{-1}(\theta)Z_n(\theta) + \hat{D}_n^{-1}(\theta)s(\theta), \hat{\Omega}_n(\theta)) \leq x \right\}. $$

Note

$$J_n(x, \Theta_\lambda, \sqrt{n}\mu(\cdot, P), \lambda, P) = P\{T_n^{\text{prof}}(\lambda) \leq x\}. $$
Subvector Inference for Moment Inequalities (cont.)

Test #3: Minimum Resampling (cont.):

**Old Idea:** Replace $s(\cdot)$ with 0 or $\hat{s}^\text{gms}(\cdot)$.

Does not lead to valid tests.

Indeed, for $P \in \mathcal{P}$ and $\lambda \in \Lambda_0(P)$,

$$\sqrt{n} \mu(\theta, P) \text{ need not be } \leq 0 \text{ for } \theta \in \Theta_\lambda.$$

$\implies$ neither 0 nor $\hat{s}^\text{gms}(\cdot)$ provide (asympt.) upper bounds on $\sqrt{n} \mu(\cdot, P)$.

In simple ex., may lead to tests with size 30% (vs. nominal size 5%).
Subvector Inference for Moment Inequalities (cont.)

Test #3: Minimum Resampling (cont.):

Main Idea: (a) Replace $\Theta_\lambda$ with a subset, e.g.,

$$\hat{\Theta}_n \approx \text{minimizers of } T_n(\theta) \text{ over } \theta \in \Theta_\lambda ,$$

over which $\hat{s}^{\text{gms}}_n(\cdot)$ provides asymp. upper bound on $\sqrt{n}\mu(\cdot, P)$.

(b) Replace $s(\theta)$ with $\hat{s}^{\text{bcs}}_n(\theta) = (\hat{s}^{\text{bcs}}_{n,1}(\theta), \ldots, \hat{s}^{\text{bcs}}_{n,k}(\theta))'$ with

$$\hat{s}^{\text{bcs}}_{n,j}(\theta) = \frac{\sqrt{n\tilde{m}_{n,j}(\theta)}}{\kappa_n \hat{\sigma}_{n,j}(\theta)} ,$$

which does provide asymp. upper bound on $\sqrt{n}\mu(\cdot, P)$.

Critical values from (a) and (b) both lead to valid tests.

Combination of two ideas leads to even better test!
Subvector Inference for Moment Inequalities (cont.)

Test #3: Minimum Resampling (cont.):

Remark: By combining both (a) and (b):

- Power advantages over both projection and subsampling
- Not true for (a) or (b) alone.

Main drawback is choice of $\kappa_n$.

Possible to generalize Romano, Shaikh & Wolf (2014) ...

... but even further generalizations possible!
General Framework

Unions of Functional Moment Inequalities:

Canay, Santos & Shaikh (in progress).

Extend Romano, Shaikh & Wolf (2014) to following problem:
For $\bar{\Theta} \subseteq \Theta$, consider null hypothesis

$$H_{\bar{\Theta}} : \exists \theta \in \bar{\Theta} \text{ with } E_P[f(W_i)] \leq 0 \text{ for all } f \in F_\theta,$$

where $f$ is a function taking values in $\mathbb{R}$.

With appropriate choice of $\bar{\Theta}$ and $F_\theta$, includes previous problems:

- moment inequalities:
  $$\bar{\Theta} = \{\theta\} \text{ and } F_\theta = \{m_j(W_i, \theta) : 1 \leq j \leq k\}.$$  

- subvector inference for moment inequalities:
  $$\bar{\Theta} = \Theta_\lambda \text{ and } F_\theta = \{m_j(W_i, \theta) : 1 \leq j \leq k\}.$$
General Framework (cont.)

Unions of Functional Moment Inequalities (cont.):

But framework includes many other problems:

- conditional moment inequalities:
  Following Andrews & Shi (2013),
  \[ \tilde{\Theta} = \{\theta\} \text{ and } F_\theta = \{m_j(W_i, \theta)  | W_i \in V : V \in \mathcal{V}, 1 \leq j \leq k\}, \]
  where \( \mathcal{V} \) is a suitable class of sets.

- subvector inference for conditional moment inequalities:
  \[ \tilde{\Theta} = \Theta_\lambda \text{ and } F_\theta = \{m_j(W_i, \theta)  | W_i \in V : V \in \mathcal{V}, 1 \leq j \leq k\} \]

- specification testing for (conditional) moment inequalities:
  \[ \tilde{\Theta} = \Theta \text{ and appropriate } F_\theta \text{ from above.} \]

As well as others, e.g., tests of stochastic dominance.
**Important Omissions**

1. Many Moment Inequalities, e.g.,
   - Chernozhukov, Chetverikov & Kato (2013) and Menzel (2014)

2. Conditional Moment Inequalities, e.g.,

3. Inference using Random Set Theory, e.g.,

4. Bayesian Approaches, e.g.,
   - Moon & Schorfheide (2012) and Kline & Tamer (2014)