Inference for Large-Scale Systems of Linear Inequalities

Zheng Fang
Andres Santos
Azeem Shaikh
Alex Torgovitsky

TAMU
UCLA
U. Chicago
U. Chicago

March 30, 2022
The Question

Let i.i.d. sample \( \{Z_i\}_{i=1}^n \) with \( Z \sim P \in \mathcal{P} \) and suppose there is a parameter \( \beta(P) \in \mathbb{R}^p \) that is unknown but estimable.
The Question

Let i.i.d. sample \( \{ Z_i \}_{i=1}^n \) with \( Z \sim P \in \mathcal{P} \) and suppose there is a parameter \( \beta(P) \in \mathbb{R}^p \) that is unknown but estimable

We aim to test whether distribution \( P \) satisfies the following null hypothesis

\[
H_0 : P \in \mathcal{P}_0 \quad \quad H_1 : P \in \mathcal{P} \setminus \mathcal{P}_0
\]

where

\[
\mathcal{P}_0 \equiv \{ P \in \mathcal{P} : \beta(P) = Ax \text{ for some } x \geq 0 \}
\]

Key Structure

- The \( p \times d \) matrix \( A \) is known.
- \( x \geq 0 \) with \( x \in \mathbb{R}^d \) denotes all coordinates of \( x \) are non-negative.
Type $h \in \{1, \ldots, H\}$ consumer, data plans $k \in \{1, \ldots, K\}$, time $t$ utility

$$u_h(c_t, y_t, v_t; k) = v_t\left(\frac{c_t^{1 - \zeta_h}}{1 - \zeta_h}\right) - c_t(\kappa_{1h} + \frac{\kappa_{2h}}{\log(s_k)}) + y_t$$

for i.i.d. shock $v_t$, data usage $c_t$, data speed $s_k$, numeraire good $y_t$. 
Example: Nevo et al. (2016)

Type $h \in \{1, \ldots, H\}$ consumer, data plans $k \in \{1, \ldots, K\}$, time $t$ utility

$$u_h(c_t, y_t, v_t; k) = v_t\left(\frac{c_t^{1-\zeta_h}}{1 - \zeta_h}\right) - c_t(\kappa_{1h} + \frac{\kappa_{2h}}{\log(s_k)}) + y_t$$

for i.i.d. shock $v_t$, data usage $c_t$, data speed $s_k$, numeraire good $y_t$.

For overage price $p_k$, fee $F_k$, data allowance $\bar{C}_k$, type $h$ utility from plan $k$ is

$$\max_{c_1, \ldots, c_T} \sum_{t=1}^{T} E_h[u_h(c_t, y_t, v_t; k)]$$

s.t. $F_k + p_k \max\{C_T - \bar{C}_k, 0\} + Y_T \leq I$, $C_T = \sum_{t=1}^{T} c_t$, $Y_T = \sum_{t=1}^{T} y_t$
For $Z$ observed plan choice and data usage, and $m$ known moment function

$$E_P[m(Z)] = \sum_{h=1}^{H} E_h[m(Z)]x_h$$

where $x = (x_1, \ldots, x_H)$ are unknown proportions of each type in population.
Example: Nevo et al. (2016)

For $Z$ observed plan choice and data usage, and $m$ known moment function

$$E_P[m(Z)] = \sum_{h=1}^{H} E_h[m(Z)]x_h$$

where $x = (x_1, \ldots, x_H)$ are unknown proportions of each type in population.

**Goal:** Inference on counterfactual demand, which for known $a_h$ equals

$$\sum_{h=1}^{H} a_h x_h$$
Example: Nevo et al. (2016)

For $Z$ observed plan choice and data usage, and $m$ known moment function

$$EP[m(Z)] = \sum_{h=1}^{H} Eh[m(Z)]x_h$$

where $x = (x_1, \ldots, x_H)$ are unknown proportions of each type in population.

**Goal:** Inference on counterfactual demand, which for known $a_h$ equals

$$\sum_{h=1}^{H} a_h x_h$$

**Current Approach**

- Build large grid of types, solve $E_h[m(Z)]$ for each type.
- Estimate proportions $x = (x_1, \ldots, x_h)$ by constrained GMM.
- Inference via bootstrap ... but bootstrap can fail.
Example: Nevo et al. (2016)

Instead, test if counterfactual demand equals hypothesized \( \lambda \) by testing if

\[
\beta(P) = Ax \text{ for some } x \geq 0
\]

with

\[
\beta(P) \equiv \begin{pmatrix}
E_P[m(Z)] \\
1 \\
\lambda
\end{pmatrix}, \quad A \equiv \begin{pmatrix}
E_1[m(Z)] & \cdots & E_H[m(Z)] \\
1 & \cdots & 1 \\
a_1 & \cdots & a_H
\end{pmatrix}
\]

Comments

- Confidence region through test inversion (in \( \lambda \)).
- We do not require proportion of types to be identified.
- In Nevo et al. (2016) \( p \approx 120000 \) and \( d \approx 16800 \).
Impact of War on Cancer

- $(S_1, S_2)$ competing risks (e.g. cardio vascular disease and cancer).
- $D$ an indicator for whether war on cancer policy in effect.
- Unspecified distribution for $(S_1, S_2)$, and for unknown $\alpha$ and $\beta$ assume

$$(T^*, I) = \begin{cases} 
(\min\{S_1, S_2\}, \arg\min\{S_1, S_2\}) & \text{if } D = 0 \\
(\min\{\alpha S_1, \beta S_2\}, \arg\min\{\alpha S_1, \beta S_2\}) & \text{if } D = 1
\end{cases}$$

Impact of War on Cancer

- \((S_1, S_2)\) competing risks (e.g. cardio vascular disease and cancer).
- \(D\) an indicator for whether war on cancer policy in effect.
- Unspecified distribution for \((S_1, S_2)\), and for unknown \(\alpha\) and \(\beta\) assume

\[
(T^*, I) = \begin{cases} 
(\min\{S_1, S_2\}, \arg\min\{S_1, S_2\}) & \text{if } D = 0 \\
(\min\{\alpha S_1, \beta S_2\}, \arg\min\{\alpha S_1, \beta S_2\}) & \text{if } D = 1 
\end{cases}
\]

Partial Identification

- We see \((T, D, I)\) where \(T\) is interval censored version of \(T^*\).
- Parameter \((\alpha, \beta)\) partially identified (even without interval censoring).

Goal: Construct confidence region for identified set for \((\alpha, \beta)\).

**Key:** \((\alpha, \beta)\) in the identified set iff there is some distribution \(p\) on \(S(\alpha, \beta)\) with

\[
\sum_{(s_1, s_2) \in S_{k,i,d}(\alpha, \beta)} p(s_1, s_2) = P(T = t_k, I = i \mid D = d)
\]

where \(S(\alpha, \beta), S_{k,i,d}(\alpha, \beta) \subseteq S(\alpha, \beta)\) are finite sets depending on \((\alpha, \beta)\).
**Example: Honore and Lleras-Muney (2006)**

**Key:** $(\alpha, \beta)$ in the identified set iff there is some distribution $p$ on $S(\alpha, \beta)$ with

$$
\sum_{(s_1, s_2) \in S_{k,i,d}(\alpha, \beta)} p(s_1, s_2) = P(T = t_k, I = i | D = d)
$$

where $S(\alpha, \beta), S_{k,i,d}(\alpha, \beta) \subseteq S(\alpha, \beta)$ are finite sets depending on $(\alpha, \beta)$.

**For Confidence Region**

- Map $\beta(P)$ into conditional probabilities (and adding up restriction).
- Map $x$ into unknown distribution $p$ satisfying restriction.
- For each candidate $(\alpha, \beta)$ sets $S_{k,i,d}(\alpha, \beta)$ map into matrix $A$.
- Test null hypothesis that $(\alpha, \beta)$ is in identified set by testing whether

$$
\beta(P) = Ax \text{ for some } x \geq 0
$$
Additional Applications

Treatment Effects

Feasibility of Linear Program

Revealed Preferences

Key Challenge: “Large” $p$ and $d \Rightarrow$ Computational scalability important
Moment Inequalities

- $P \in P_0$ if and only if $\beta(P)$ is in set defined by inequalities (in $\mathbb{R}^p$).
- Challenge: For large $p$, $d$, computing inequalities is prohibitive.

Other Related Work

- Kitamura and Stoye (2018) test imposes restrictions on $A$ (satisfied in the revealed preferences problem that motivates them).
- Cox & Shi (2021) derive tuning parameter free method for inference.
Related Literature

Moment Inequalities

- $P \in \mathbf{P}_0$ if and only if $\beta(P)$ is in set defined by inequalities (in $\mathbb{R}^p$).
- Challenge: For large $p, d$, computing inequalities is prohibitive.

Shape Restrictions

- $P \in \mathbf{P}_0$ if and only if $\beta(P)$ is in convex set.
- We employ specific structure in computation and assumptions.

Kitamura and Stoye (2018) test imposes restrictions on $A$ (satisfied in the revealed preferences problem that motivates them).


Cox & Shi (2021) derive tuning parameter free method for inference.
Related Literature

Moment Inequalities

• $P \in P_0$ if and only if $\beta(P)$ is in set defined by inequalities (in $\mathbb{R}^p$).
• Challenge: For large $p, d$, computing inequalities is prohibitive.

Shape Restrictions

• $P \in P_0$ if and only if $\beta(P)$ is in convex set.
• We employ specific structure in computation and assumptions.

Other Related Work

• Kitamura and Stoye (2018) test imposes restrictions on $A$ (satisfied in the revealed preferences problem that motivates them).
• Andrews, Pakes & Roth (2019) find least favorable for subvector inference in a class of (conditional) moment inequalities models.
• Cox & Shi (2021) derive tuning parameter free method for inference.
1 The Geometry

2 The Test

3 Simulations
Some Notation

**Question:** For any $\beta \in \mathbb{R}^p$, when is $\beta = Ax$ for some $x \geq 0$?
Some Notation

**Question:** For any $\beta \in \mathbb{R}^p$, when is $\beta = Ax$ for some $x \geq 0$?

**Three Subspaces**

\[
R \equiv \{ b \in \mathbb{R}^p : b = Ax \text{ for some } x \in \mathbb{R}^d \}
\]

\[
N \equiv \{ x \in \mathbb{R}^d : Ax = 0 \}
\]

\[
N^\perp \equiv \{ y \in \mathbb{R}^d : \langle y, x \rangle = 0 \text{ for all } x \in N \}
\]
Question: For any $\beta \in \mathbb{R}^p$, when is $\beta = Ax$ for some $x \geq 0$?

Three Subspaces

\[ R \equiv \{ b \in \mathbb{R}^p : b = Ax \text{ for some } x \in \mathbb{R}^d \} \]

\[ N \equiv \{ x \in \mathbb{R}^d : Ax = 0 \} \]

\[ N^\perp \equiv \{ y \in \mathbb{R}^d : \langle y, x \rangle = 0 \text{ for all } x \in N \} \]

Some Intuition

- If $\beta = Ax$ text for some $x \geq 0$, then in particular we must have that ...

\[ \beta \in R \]
Some Notation

**Question:** For any $\beta \in \mathbb{R}^p$, when is $\beta = Ax$ for some $x \geq 0$?

Three Subspaces

$$R \equiv \{ b \in \mathbb{R}^p : b = Ax \text{ for some } x \in \mathbb{R}^d \}$$

$$N \equiv \{ x \in \mathbb{R}^d : Ax = 0 \}$$

$$N^\perp \equiv \{ y \in \mathbb{R}^d : \langle y, x \rangle = 0 \text{ for all } x \in N \}$$

Some Intuition

- If $\beta = Ax$ text for some $x \geq 0$, then in particular we must have that ...

  $$\beta \in R$$

- If $\beta = Ax_1$ for some $x_1$ and $x_2 \in N$ then $\beta = A(x_1 + x_2)$ so ...

  $\Rightarrow$ Intuitively, if $x_1 \not< 0$, then maybe can fix it by moving along $N$
**Lemma:** If $\beta \in \mathbb{R}$, then there is unique $x^* \in N^\perp$ satisfying the equality

$$\beta = Ax^*$$
Simple Lemma

**Lemma:** If $\beta \in R$, then there is unique $x^* \in N^\perp$ satisfying the equality

$$\beta = Ax^*$$

**Key Implications**

- $\beta = Ax$ with $x \geq 0$ requires $\beta \in R$. 

Fang, Santos, Shaikh, Torgovitsky. March 30, 2022. UCLA
**Simple Lemma**

**Lemma:** If $\beta \in \mathbb{R}$, then there is unique $x^* \in N^\perp$ satisfying the equality

$$\beta = Ax^*$$

**Key Implications**

- $\beta = Ax$ with $x \geq 0$ requires $\beta \in \mathbb{R}$.
- Moreover, the above lemma implies set of solutions to $\beta = Ax$ equals

$$\{x \in \mathbb{R}^d : Ax = \beta\} = x^* + N$$
**Simple Lemma**

**Lemma:** If $\beta \in \mathbb{R}$, then there is unique $x^* \in N^\perp$ satisfying the equality

$$\beta = Ax^*$$

**Key Implications**

- $\beta = Ax$ with $x \geq 0$ requires $\beta \in \mathbb{R}$.
- Moreover, the above lemma implies set of solutions to $\beta = Ax$ equals

$$\{ x \in \mathbb{R}^d : Ax = \beta \} = x^* + N$$

- Whether $\beta = Ax$ for some $x \geq 0$ characterized by $\beta \in \mathbb{R}^p$, $x^* \in \mathbb{R}^d$ via

$$\begin{align*}
(\text{i}) & \quad \beta \in \mathbb{R} \\
(\text{ii}) & \quad \{x^* + N\} \cap \mathbb{R}^d_+ \neq \emptyset
\end{align*}$$

**Key Challenge:** Obtaining tractable characterization for (ii).
Geometric Intuition

Suppose $\beta = Ax$ with $x^* \in N^\perp$... is there positive solution?

Question: What if instead $\beta = Ax^* \in N^\perp$?

Note: In this example positive solution always exists (provided $\beta \in \mathbb{R}$)

$\mathbb{R}_+^2 \cap N \cap N^\perp$
Example: Suppose $\beta = Ax^*_1$ with $x^*_1 \in N^\perp$ ... is there positive solution?
Example: Suppose $\beta = Ax^*_1$ with $x^*_1 \in N^\perp$ ... is there positive solution?

Note: In this example positive solution always exists (provided $\beta \in \mathbb{R}$).
Example: Suppose $\beta = Ax_1^*$ with $x_1^* \in N^\perp$ ... is there positive solution?

Note: In this example positive solution always exists (provided $\beta \in \mathbb{R}$).
Example: Suppose $\beta = Ax^*_1$ with $x^*_1 \in N^\perp$ ... is there positive solution?

Note: In this example positive solution always exists (provided $\beta \in \mathbb{R}$)

$\mathbb{R}_2^+$

$\mathbb{R}_2^-$

$N^\perp$

$N + x^*_1$

$N$
**Geometric Intuition**

**Question:** What if instead $\beta = Ax_2^*$ with $x_2^* \in N^\perp$?
Question: What if instead $\beta = Ax^*_2$ with $x^*_2 \in N^\perp$?
**Geometric Intuition**

**Question:** What if instead $\beta = Ax^*_2$ with $x^*_2 \in N^\perp$?

Note: In this example positive solution always exists (provided $\beta \in \mathbb{R}$).
Geometric Intuition

**Note:** In this example positive solution always exists (provided $\beta \in R$)
Geometric Intuition

Example: Suppose $\beta = A x^\star_1$ with $x^\star_1 \in N \perp \ldots$ is there a positive solution?

Question: What if instead $\beta = A x^\star_2$ with $x^\star_2 \in N \perp$?

Note: Positive solution exists if and only if $x^\star \in \mathbb{R}^2_+$ (provided $\beta \in \mathbb{R}$).
Geometric Intuition

**Example:** Suppose $\beta = Ax^*_1$ with $x^*_1 \in N^\perp$ ... is there a positive solution?
Example: Suppose $\beta = Ax^*_1$ with $x^*_1 \in N^\perp$ ... is there a positive solution?
**Geometric Intuition**

**Question:** What if instead $\beta = Ax^*_2$ with $x^*_2 \in N^\perp$?
Question: What if instead $\beta = Ax^*_2$ with $x^*_2 \in N^\perp$?
Question: What if instead $\beta = Ax^*_2$ with $x^*_2 \in N^\perp$?
Geometric Intuition

**Note:** Positive solution exists if and only if \( x^* \in \mathbb{R}_+^2 \) (provided \( \beta \in \mathbb{R} \))
Geometric Characterization

(i) $\beta \in R$  \quad (ii) $\{x^* + N\} \cap \mathbb{R}^d_+ \neq \emptyset$

**Goal:** Obtain alternative characterization that suggests natural test statistic.
Geometric Characterization

(i) \( \beta \in \mathbb{R} \) \quad (ii) \( \{x^* + N\} \cap \mathbb{R}^d_+ \neq \emptyset \)

**Goal:** Obtain alternative characterization that suggests natural test statistic.

**Theorem:** There is an \( x_0 \in \mathbb{R}^d_+ \) satisfying \( Ax_0 = \beta \) if and only if

(i) \( \beta \in \mathbb{R} \) \quad (ii) \( \langle s, x^* \rangle \leq 0 \) for all \( s \in N^\perp \cap \mathbb{R}^d_- \)

**Comments**

- Condition (i) yields “equalities” and (ii) yields “inequalities.”
- (ii) equivalent to angles between \( x^* \) and \( N^\perp \cap \mathbb{R}^d_- \) are obtuse.
- Reflects dependence on \( x^* \) and “orientation” of \( N^\perp \) in \( \mathbb{R}^d \).
Geometric Intuition

\[ \langle s, x^* \rangle \leq 0 \text{ for all } s \in N^\perp \cap \mathbb{R}^d \]
Geometric Intuition

\[ \langle s, x^* \rangle \leq 0 \text{ for all } s \in N^\perp \cap \mathbb{R}^d \]

Equivalent angle between \( x^* \) and \( N^\perp \cap \mathbb{R}^d = \{(0, 0, \lambda) : \lambda \leq 0\} \) obtuse.
Geometric Intuition

\[ \langle s, x^* \rangle \leq 0 \text{ for all } s \in N^\perp \cap \mathbb{R}^d \]

Equivalent angle between \( x^* \) and \( N^\perp \cap \mathbb{R}^d = \{(0, 0, \lambda) : \lambda \leq 0\} \) obtuse.
\( \langle s, x^* \rangle \leq 0 \) for all \( s \in N^\perp \cap \mathbb{R}_d^+ \)

Equivalent angle between \( x^* \) and \( N^\perp \cap \mathbb{R}_d^+ = \{(0, 0, \lambda) : \lambda \leq 0\} \) obtuse.
Geometric Intuition

\[ \langle s, x^* \rangle \leq 0 \text{ for all } s \in N^\perp \cap \mathbb{R}^d \]

Equivalent angle between \( x^* \) and \( N^\perp \cap \mathbb{R}^d = \{(0, 0, \lambda) : \lambda \leq 0\} \) obtuse.
Geometric Intuition

\[ \langle s, x^* \rangle \leq 0 \text{ for all } s \in N^\perp \cap \mathbb{R}^d \]

Equivalent angle between \( x^* \) and \( N^\perp \cap \mathbb{R}^d = \{(0, 0, \lambda) : \lambda \leq 0\} \) obtuse.
Geometric Intuition

\[ \langle s, x^* \rangle \leq 0 \text{ for all } s \in N^\perp \cap \mathbb{R}_d^d \]

Equivalent angle between \( x^* \) and \( N^\perp \cap \mathbb{R}_d^d = \{(0, 0, \lambda) : \lambda \leq 0\} \) obtuse.
Geometric Intuition

\[ \langle s, x^* \rangle \leq 0 \text{ for all } s \in N^\perp \cap \mathbb{R}^d \]

Equivalent angle between \( x^* \) and \( N^\perp \cap \mathbb{R}^d = \{(0, 0, \lambda) : \lambda \leq 0\} \) obtuse.
Geometric Intuition

\[ \langle s, x^* \rangle \leq 0 \text{ for all } s \in N^\perp \cap \mathbb{R}_d^d \]

Equivalent angle between \( x^* \) and \( N^\perp \cap \mathbb{R}_d^d = \{ (0, 0, \lambda) : \lambda \leq 0 \} \) obtuse.
Geometric Intuition

\[ \langle s, x^* \rangle \leq 0 \text{ for all } s \in N^\perp \cap \mathbb{R}_d^d \]

Equivalent angle between \( x^* \) and \( N^\perp \cap \mathbb{R}_d^d = \{(0, 0, \lambda) : \lambda \leq 0\} \) obtuse.
1 The Geometry

2 The Test

3 Simulations
Test Statistic

Key: For $x^*(P) \in N^\perp$ solving $\beta(P) = Ax^*(P)$, $P \in P_0$ if and only if

(i) $\beta(P) \in R$

(ii) $\langle s, x^*(P) \rangle \leq 0$ for all $s \in N^\perp \cap R^d$
**Test Statistic**

**Key:** For $x^*(P) \in N^\perp$ solving $\beta(P) = A x^*(P)$, $P \in P_0$ if and only if

1. $\beta(P) \in R$
2. $\langle s, x^*(P) \rangle \leq 0$ for all $s \in N^\perp \cap \mathbb{R}_d$

**For talk only:** Assume $R = \mathbb{R}^p$ so condition (i) is automatically satisfied.
Test Statistic

Key: For $x^*(P) \in N^\perp$ solving $\beta(P) = Ax^*(P)$, $P \in P_0$ if and only if

(i) $\beta(P) \in R$ (ii) $\langle s, x^*(P) \rangle \leq 0$ for all $s \in N^\perp \cap R^d$

For talk only: Assume $R = R^p$ so condition (i) is automatically satisfied.

The Pseudoinverse

- Under $R = R^p$, for any $b \in R^p$ there is unique $x(b) \in N^\perp$ solving

  $$b = Ax(b)$$
Test Statistic

**Key:** For $x^*(P) \in N^\perp$ solving $\beta(P) = Ax^*(P)$, $P \in P_0$ if and only if

(i) $\beta(P) \in R$

(ii) $\langle s, x^*(P) \rangle \leq 0$ for all $s \in N^\perp \cap \mathbb{R}^d$

*For talk only:* Assume $R = \mathbb{R}^p$ so condition (i) is automatically satisfied.

**The Pseudoinverse**

- Under $R = \mathbb{R}^p$, for any $b \in \mathbb{R}^p$ there is unique $x(b) \in N^\perp$ solving

  $b = Ax(b)$

- Under $R = \mathbb{R}^p$, the (MP) pseudoinverse $A^\dagger$ of $A$ is $d \times p$ matrix solving

  $x(b) = A^\dagger b$
\[ \langle s, x^*(P) \rangle \leq 0 \text{ for all } s \in N^\perp \cap \mathbb{R}^d \]
\[ \langle s, x^*(P) \rangle \leq 0 \text{ for all } s \in N^\perp \cap \mathbb{R}^d \]

... or equivalently, since \( A^\dagger \beta(P) = x^*(P) \), we may re-write condition as

\[ \langle s, A^\dagger \beta(P) \rangle \leq 0 \text{ for all } s \in N^\perp \cap \mathbb{R}^d \]
Test Statistic

\[ \langle s, x^*(P) \rangle \leq 0 \text{ for all } s \in N^\perp \cap \mathbb{R}^d \]

... or equivalently, since \( A^\dagger \beta(P) = x^*(P) \), we may re-write condition as

\[ \langle s, A^\dagger \beta(P) \rangle \leq 0 \text{ for all } s \in N^\perp \cap \mathbb{R}^d \]

... or equivalently, since \( \text{range}\{A^\dagger\} = N^\perp \), we may re-write condition as

\[ \langle A^\dagger s, A^\dagger \beta(P) \rangle \leq 0 \text{ for all } s \in \mathbb{R}^p \text{ s.t. } A^\dagger s \leq 0 \text{ (in } \mathbb{R}^d) \]
\[ \langle A^\dagger s, A^\dagger \beta(P) \rangle \leq 0 \text{ for all } s \in \mathbb{R}^p \text{ s.t. } A^\dagger s \leq 0 \text{ (in } \mathbb{R}^d) \]
\[ \langle A^\dagger s, A^\dagger \beta(P) \rangle \leq 0 \text{ for all } s \in \mathbb{R}^p \text{ s.t. } A^\dagger s \leq 0 \text{ (in } \mathbb{R}^d) \]

Test Statistic

\[ T_n = \sup_{s \in \hat{V}_n} \langle A^\dagger s, A^\dagger \hat{\beta}_n \rangle \]
Test Statistic

\[ \langle A^\dagger s, A^\dagger \beta(P) \rangle \leq 0 \text{ for all } s \in \mathbb{R}^p \text{ s.t. } A^\dagger s \leq 0 \text{ (in } \mathbb{R}^d) \]

Test Statistic

\[
T_n = \sup_{s \in \hat{\mathcal{V}}_n} \langle A^\dagger s, A^\dagger \hat{\beta}_n \rangle
\]

\[
\hat{\mathcal{V}}_n = \{ s \in \mathbb{R}^p : A^\dagger s \leq 0 \text{ and } \|\hat{\Omega}_n (AA')^\dagger s\|_1 \leq 1 \}
\]

Comments

- Weighting matrix \( \hat{\Omega}_n \) can be used to obtain scale invariance.
- Norm constraint ensures \( T_n \neq +\infty \) with positive probability.
- Test statistic can be computed by linear programming.
- The norm \( \| \cdot \|_1 \) yields better coupling rates than, e.g., \( \| \cdot \|_2 \).
Test Statistic

Assumption T

- $\hat{\beta}_n$ is function of i.i.d. sample $\{Z_i\}_{i=1}^n$ with $Z_i \sim P \in \mathcal{P}$.
- $\hat{\Omega}_n$ is consistent for $\Omega$ uniformly in $P \in \mathcal{P}$ (under $\| \cdot \|_{o,\infty}$).
- For some sequence $a_n \downarrow 0$ and influence function $\psi$ we have

$$\|\Omega^\dagger \{ \sqrt{n} \{ \hat{\beta}_n - \beta(P) \} - \frac{1}{\sqrt{n}} \sum_{i=1}^n \psi(Z_i) \} \|_{\infty} = O_P(a_n)$$

Comments

- Weighting matrix $\Omega$ need not be invertible.
- Estimator $\hat{\beta}_n$ is asymptotically linear.
- Norm $\| \cdot \|_{\infty}$ leads to favorable rate conditions in $p$. 
Asymptotic Distribution

**Theorem:** Under Assumption T and regularity conditions, we have

\[ T_n \equiv \sup_{s \in \hat{V}_n} \sqrt{n} \langle A^\dagger s, A^\dagger \hat{\beta}_n \rangle \]

\[ = \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger G_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle + O_P(r_n) \]

for some centered gaussian \( G_n \in \mathbb{R}^p \) (uniformly in \( P \in P \))

**Comments**

- Set \( \mathcal{V} \subset \mathbb{R}^p \) just population analogue to \( \hat{V}_n \).
- Under moment conditions, \( r_n \downarrow 0 \) provided \( p^2/n + a_n \downarrow 0 \) (up to logs).
- \( \| \cdot \|_1 \) constraint defining \( \hat{V}_n \) (and \( \mathcal{V} \)) facilitate coupling under \( \| \cdot \|_\infty \).
Critical Value

\[ T_n = \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger G_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle + O_P(r_n) \]

- Can be simulated
- Nuisance parameter

Like Moment Inequalities

- From geometry section, \( \langle A^\dagger s, A^\dagger \beta(P) \rangle \leq 0 \) for all \( s \in \mathcal{V}, P \in \mathbf{P}_0 \).
- Multiple techniques available from moment inequalities literature.
Critical Value

\[ T_n = \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger G_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle + O_P(r_n) \]

- Can be simulated
- Nuisance parameter

Like Moment Inequalities

- From geometry section, \( \langle A^\dagger s, A^\dagger \beta(P) \rangle \leq 0 \) for all \( s \in \mathcal{V}, P \in P_0 \).
- Multiple techniques available from moment inequalities literature.

... But Different

- Replace \( \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle \) with zero (may not be least favorable).
- Moment selection (e.g., Andrews & Soares 2010), two step procedures (e.g., Romano, Shaikh & Wolf 2014) can suffer in power.

Key: Nuisance parameter has additional structure beyond it being negative!
Critical Value

First Step

\[
\hat{\beta}^r_n \in \arg \min_{b \in \mathbb{R}^p} \sup_{s \in \hat{\mathcal{V}}_n} |\langle A^\dagger s, A^\dagger \hat{\beta}_n - A^\dagger b \rangle| \quad \text{s.t. } Ax = b \text{ for some } x \geq 0
\]
Critical Value

First Step

\[ \hat{\beta}_n^r \in \arg \min_{b \in \mathbb{R}^p} \sup_{s \in \hat{V}_n} |\langle A^\dagger s, A^\dagger \hat{\beta}_n - A^\dagger b \rangle| \quad \text{s.t. } Ax = b \text{ for some } x \geq 0 \]

Bootstrap Statistic

\[ T_n^* \equiv \sup_{s \in \hat{V}_n} \langle A^\dagger s, A^\dagger \hat{G}_n^* \rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \hat{\beta}_n^r \rangle \]

where \( 1 \geq \lambda_n \downarrow 0 \) and \( \hat{G}_n^* = \sqrt{n} \{ \hat{\beta}_n^* - \hat{\beta}_n \} \) with \( \hat{\beta}_n^* \) “bootstrapped” \( \hat{\beta}_n \)
Critical Value

First Step

\[ \hat{\beta}_n^r \in \arg \min_{b \in \mathbb{R}^p} \sup_{s \in \hat{V}_n} |\langle A^\dagger s, A^\dagger \hat{\beta}_n - A^\dagger b \rangle| \quad \text{s.t. } Ax = b \text{ for some } x \geq 0 \]

Bootstrap Statistic

\[ T_n^* \equiv \sup_{s \in \hat{V}_n} \langle A^\dagger s, A^\dagger \hat{G}_n^* \rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \hat{\beta}_n^r \rangle \]

where \( 1 \geq \lambda_n \downarrow 0 \) and \( \hat{G}_n^* = \sqrt{n} \{ \hat{\beta}_n^* - \hat{\beta}_n \} \) with \( \hat{\beta}_n^* \) “bootstrapped” \( \hat{\beta}_n \)

Critical Value

\[ \hat{c}_n(1 - \alpha) \equiv \inf \{ u : P(T_n^* \leq u | \{ Z_i \}_{i=1}^n) \geq 1 - \alpha \} \]
Some Intuition

**Question:** Why does this bootstrap approximation control size?

\[ T_n^* \equiv \sup_{s \in \hat{V}_n} \langle A^\dagger s, A^\dagger \hat{G}_n^* \rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \beta_n^r \rangle \]
Some Intuition

**Question:** Why does this bootstrap approximation control size?

\[
T_n^* \equiv \sup_{s \in \hat{V}_n} \langle A^\dagger s, A^\dagger \hat{G}_n^* \rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \hat{\beta}_n^r \rangle \\
\approx \sup_{s \in \hat{V}_n} \langle A^\dagger s, A^\dagger \hat{G}_n^* \rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle \\
\approx T_n \quad \text{(if } \lambda_n \to 0) 
\]
Some Intuition

**Question:** Why does this bootstrap approximation control size?

\[ T_n^* \equiv \sup_{s \in \hat{V}_n} \langle A^\dagger s, A^\dagger \hat{G}_n^* \rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \hat{\beta}^r_n \rangle \]

\[ \approx \sup_{s \in \hat{V}_n} \langle A^\dagger s, A^\dagger \hat{G}_n^* \rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle \quad \text{(if } \lambda_n \to 0) \]

\[ \geq \sup_{s \in \hat{V}_n} \langle A^\dagger s, A^\dagger \hat{G}_n^* \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle \quad \text{(by } \langle A^\dagger s, A^\dagger \beta(P) \rangle \leq 0) \]
**Some Intuition**

**Question:** Why does this bootstrap approximation control size?

\[
T_n^* = \sup_{s \in \hat{V}_n} \langle A^\dagger s, A^\dagger \hat{G}_n^* \rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \hat{\beta}^r_n \rangle \\
\approx \sup_{s \in \hat{V}_n} \langle A^\dagger s, A^\dagger \hat{G}_n^* \rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle \\
\geq \sup_{s \in \hat{V}_n} \langle A^\dagger s, A^\dagger \hat{G}_n^* \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle \\
\approx \sup_{s \in \hat{V}_n} \langle A^\dagger s, A^\dagger \hat{G}_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle
\]

(by \( \langle A^\dagger s, A^\dagger \beta(P) \rangle \leq 0 \))

(d bootstrap cons.)
Some Intuition

**Question:** Why does this bootstrap approximation control size?

\[
T_n^* \equiv \sup_{s \in \hat{V}_n} \langle A^\dagger s, A^\dagger \hat{G}_n^* \rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \beta_n^r \rangle \\
\approx \sup_{s \in \hat{V}_n} \langle A^\dagger s, A^\dagger \hat{G}_n^* \rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle \\
\geq \sup_{s \in \hat{V}_n} \langle A^\dagger s, A^\dagger \hat{G}_n^* \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle \\
\overset{d}{\approx} \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger G_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle \\
\approx T_n
\]

(if \(\lambda_n \to 0\))

(by \(\langle A^\dagger s, A^\dagger \beta(P) \rangle \leq 0\))

(bootstrap cons.)

(by theorem)

**Key:** Bootstrap provides uniform upper bound ... but is it conservative?
Suppose: $P$ is fixed and $n \to \infty$ (i.e. pointwise, not uniform analysis)

\[
T_n \approx \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger G_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle
\]
Some Intuition

**Suppose:** $P$ is fixed and $n \to \infty$ (i.e. pointwise, not uniform analysis)

\[
T_n \approx \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger G_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle
\]

\[
= \max\{0, \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger G_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle\} \quad \text{(since } 0 \in \mathcal{V})
\]
**Some Intuition**

**Suppose:** $P$ is fixed and $n \to \infty$ (i.e. pointwise, not uniform analysis)

\[
T_n \approx \sup_{s \in \mathcal{V}} \langle A^\dag s, A^\dag G_n \rangle + \sqrt{n} \langle A^\dag s, A^\dag \beta(P) \rangle
\]

\[
= \max\{0, \sup_{s \in \mathcal{V}} \langle A^\dag s, A^\dag G_n \rangle + \sqrt{n} \langle A^\dag s, A^\dag \beta(P) \rangle\} \quad \text{(since } 0 \in \mathcal{V})
\]

if $\ll 0$ then $s$ drops out
**Some Intuition**

**Suppose:** $P$ is fixed and $n \to \infty$ (i.e. pointwise, not uniform analysis)

\[
T_n \approx \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \mathcal{G}_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle
\]

\[
= \max\{0, \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \mathcal{G}_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle\} \quad (\text{since } 0 \in \mathcal{V})
\]

\[
\approx \max\{0, \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \mathcal{G}_n \rangle \text{ s.t. } \langle A^\dagger s, A^\dagger \beta(P) \rangle = 0\} \quad (P \text{ is fixed})
\]
**Some Intuition**

**Suppose:** $P$ is fixed and $n \to \infty$ (i.e. pointwise, not uniform analysis)

$$T_n \approx \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger G_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle$$

$$= \max\{0, \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger G_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle\} \quad \text{(since } 0 \in \mathcal{V})$$

$$\approx \max\{0, \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger G_n \rangle \text{ s.t. } \langle A^\dagger s, A^\dagger \beta(P) \rangle = 0\} \quad \text{(} P \text{ is fixed)}$$

**What About Bootstrap?**
Some Intuition

**Suppose:** $P$ is fixed and $n \to \infty$ (i.e. pointwise, not uniform analysis)

$$T_n \approx \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle$$

$$= \max \{0, \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle\}$$  \hspace{1cm} \text{(since $0 \in \mathcal{V}$)}

$$\approx \max \{0, \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n \rangle \text{ s.t. } \langle A^\dagger s, A^\dagger \beta(P) \rangle = 0\}$$  \hspace{1cm} \text{($P$ is fixed)}

**What About Bootstrap?**

$$T_n^* \approx \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \hat{\mathbb{G}}_n \rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle$$  \hspace{1cm} \text{(shown before)}
Some Intuition

Suppose: $P$ is fixed and $n \to \infty$ (i.e. pointwise, not uniform analysis)

\[
T_n \approx \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger G_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle \\
= \max \{0, \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger G_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle \} \\
\approx \max \{0, \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger G_n \rangle \ \text{s.t.} \ \langle A^\dagger s, A^\dagger \beta(P) \rangle = 0\} \\
\text{(since } 0 \in \mathcal{V}) \\
(P \text{ is fixed})
\]

What About Bootstrap?

\[
T_n^* \approx \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \hat{G}_n^* \rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle \\
= \max \{0, \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \hat{G}_n^* \rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle \} \\
\text{(shown before)} \\
\text{(since } 0 \in \mathcal{V})
\]
Some Intuition

Suppose: $P$ is fixed and $n \to \infty$ (i.e. pointwise, not uniform analysis)

$$T_n \approx \sup_{s \in V} \langle A^\dagger s, A^\dagger G_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle$$

$$= \max \{ 0, \sup_{s \in V} \langle A^\dagger s, A^\dagger G_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle \}$$  \hspace{1cm} (since $0 \in V$)

$$\approx \max \{ 0, \sup_{s \in V} \langle A^\dagger s, A^\dagger G_n \rangle \text{ s.t. } \langle A^\dagger s, A^\dagger \beta(P) \rangle = 0 \}$$  \hspace{1cm} ($P$ is fixed)

What About Bootstrap?

$$T_n^* \approx \sup_{s \in V} \langle A^\dagger s, A^\dagger \hat{G}_n^* \rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle$$  \hspace{1cm} (shown before)

$$= \max \{ 0, \sup_{s \in V} \langle A^\dagger s, A^\dagger \hat{G}_n^* \rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle \}$$  \hspace{1cm} (since $0 \in V$)

\[ \text{if } \ll 0 \text{ then } s \text{ drops out} \]
Some Intuition

**Suppose:** $P$ is fixed and $n \to \infty$ (i.e. pointwise, not uniform analysis)

\[ T_n \approx \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger G_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle \]

\[ = \max \{0, \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger G_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle\} \quad \text{(since 0 \in \mathcal{V})} \]

\[ \approx \max \{0, \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger G_n \rangle \ \text{s.t.} \ \langle A^\dagger s, A^\dagger \beta(P) \rangle = 0\} \quad \text{(P is fixed)} \]

**What About Bootstrap?**

\[ T^*_n \approx \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \hat{G}^*_n \rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle \quad \text{(shown before)} \]

\[ = \max \{0, \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \hat{G}^*_n \rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle\} \quad \text{(since 0 \in \mathcal{V})} \]

\[ \approx \max \{0, \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \hat{G}^*_n \rangle \ \text{s.t.} \ \langle A^\dagger s, A^\dagger \beta(P) \rangle = 0\} \quad \text{(if } \lambda_n \sqrt{n} \to \infty) \]
Critical Value

Assumption B

- There are random variables \( \{W_{i,n}\}_{i=1}^n \) independent of \( \{Z_i\}_{i=1}^n \) with

\[
\| \Omega^\dagger \{ \hat{G}^* - \frac{1}{\sqrt{n}} \sum_{i=1}^n (W_{i,n} - \bar{W}_n) \psi(Z_i) \} \|_\infty = O_P(a_n)
\]

- The distribution of \( \{W_{i,n}\}_{i=1}^n \) is exchangeable.

Comments

- Asymptotically linear assumption analogous to requirement on \( \hat{\beta}_n \).
- Exchangeability covers multiplier, score, and nonparametric bootstrap.
- Derive coupling results for exchangeable bootstrap under \( \| \cdot \|_\infty \).
Asymptotic Size

**Theorem:** Under Assumptions T, B, regularity conditions, and $\alpha \in (0, 0.5)$

$$\limsup_{n \to \infty} \sup_{P \in P_0} P(T_n > \hat{c}_n(1 - \alpha)) \leq \alpha$$

**Comments**

- Bootstrap coupling requires $p^2/n \downarrow 0$ (up to logs).
- **Anti-concentration:** Under fixed $p$ and studentization automatic.
- **Anti-concentration:** Dependence on $p$ through $(AA')^\dagger \mathcal{V}$.
- Conservative universal (in $A$) bounds on dependence on $p$ available.
- Under same conditions, two stage critical value also valid.
1 The Geometry

2 The Test

3 Simulations
Simulation Design

\[ Y = 1\{C_0 + C_1W \geq U\} \]

\[ U \sim \text{logistic}, \text{ unobservable } V \equiv (C_0, C_1)' \text{ and observable } W \text{ discrete.} \]

Comments

- \( W, V, \) and \( U \) all mutually independent.
- Random coefficients logit (Fox, Kim, Ryan, Bajari, 2011).
- \( C_0 \in [0.5, 1], C_1 \in [-3, 0] \) with \( \sqrt{d} \) points of support each.
- Support of \( W \) is evenly spaced grid on \([0, 2]\) (cardinality equals \( p - 2 \)).
- 250 bootstrap draws, 5000 or 1000 replications.
Simulation Design

Restrictions

• For \( V \) support of \( V \), \( \pi(v) = P(V = v) \), and \( v = (c_1, c_2)' \) we have

\[
P(Y = 1|W = w) = \sum_{v \in V} \pi(v) \frac{1}{1 + \exp\{-c_0 - c_1 w\}}
\]

• Unknown probabilities \( \{\pi(v) : v \in V\} \) satisfy \( \sum_{v \in V} \pi(v) = 1 \).
Simulation Design

Restrictions

• For $\mathcal{V}$ support of $V$, $\pi(v) = P(V = v)$, and $v = (c_1, c_2)'$ we have

$$P(Y = 1|W = w) = \sum_{v \in \mathcal{V}} \pi(v) \frac{1}{1 + \exp\{-c_0 - c_1 w\}}$$

• Unknown probabilities $\{\pi(v) : v \in \mathcal{V}\}$ satisfy $\sum_{v \in \mathcal{V}} \pi(v) = 1$.

Parameter of Interest

• Consumer type $v = (c_0, c_1)$ with price $\bar{w}$ has purchase prob. elasticity

$$\epsilon(v, \bar{w}) \equiv c_0 \bar{w}(1 - \frac{1}{1 + \exp\{-c_0 - c_1 \bar{w}\}})$$

• Inference on $F(t|\bar{w}) \equiv P(\epsilon(V, \bar{w}) \leq t) = \sum_{v \in \mathcal{V}} \pi(v) 1\{\epsilon(v, \bar{w}) \leq t\}$
Design Partially Identified

Figure: Dark: $W$ with 4 support points, Lighter: $W$ with 16 support points
Simulation Design

The General Problem

\[ \beta(P) = Ax \text{ for some } x \geq 0 \]

In this Design

- \( x \in \mathbb{R}^d \) is the unknown probabilities \( \{\pi(v) : v \in \mathcal{V}\} \).
- \( \beta(P) \in \mathbb{R}^p \), first \( p - 2 \) coordinates correspond to \( P(Y = 1|W = w) \).
- The \( p - 1 \) coordinate of \( \beta(P) \) equals 1 (\( \sum_{v \in \mathcal{V}} \pi(v) = 1 \)).
- The \( p \) coordinate of \( \beta(P) \) equals hypothesized value for \( F(-1|1) \).

Bandwidth Selection

- Law of iterated logarithm: \( \lambda_n^r = (\log(e \vee p) \log(e \vee \log(e \vee n)))^{-1/2} \).
- Bootstrap: Set \( 1/\lambda_n^b \) to be \( 1 - (\log(e \vee \log(e \vee n)))^{-1/2} \) quantile of

\[ \sup_{s \in \hat{\mathcal{V}}_n} \langle A^\dagger s, A^\dagger \hat{G}_n \rangle \]
(Almost) Identified Case

<table>
<thead>
<tr>
<th>Test</th>
<th>BS Wald</th>
<th>BS Wald (RC)</th>
<th>FSST</th>
<th>FSST (RoT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d = 4, p = 6$</td>
<td><img src="image1.png" alt="Graph" /></td>
<td><img src="image2.png" alt="Graph" /></td>
<td><img src="image3.png" alt="Graph" /></td>
<td><img src="image4.png" alt="Graph" /></td>
</tr>
<tr>
<td>$d = 4, p = 18$</td>
<td><img src="image5.png" alt="Graph" /></td>
<td><img src="image6.png" alt="Graph" /></td>
<td><img src="image7.png" alt="Graph" /></td>
<td><img src="image8.png" alt="Graph" /></td>
</tr>
<tr>
<td>$d = 16, p = 18$</td>
<td><img src="image9.png" alt="Graph" /></td>
<td><img src="image10.png" alt="Graph" /></td>
<td><img src="image11.png" alt="Graph" /></td>
<td><img src="image12.png" alt="Graph" /></td>
</tr>
</tbody>
</table>

Fang, Santos, Shaikh, Torgovitsky. March 30, 2022. UCLA
## Null Rejection: Bootstrap Bandwidth

**Table:** Null Hypothesis that $F_\epsilon(-1 | 1)$ equals lower bound of identified set.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$p$</th>
<th>100</th>
<th>400</th>
<th>1600</th>
<th>4900</th>
<th>$100^2$</th>
<th>$225^2$</th>
<th>$317^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>6</td>
<td>.036</td>
<td>.034</td>
<td>.034</td>
<td>.037</td>
<td>.038</td>
<td>.036</td>
<td>.036</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>.040</td>
<td>.035</td>
<td>.036</td>
<td>.041</td>
<td>.039</td>
<td>.038</td>
<td>.036</td>
</tr>
<tr>
<td>2000</td>
<td>6</td>
<td>.042</td>
<td>.042</td>
<td>.049</td>
<td>.046</td>
<td>.047</td>
<td>.052</td>
<td>.061</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>.031</td>
<td>.028</td>
<td>.032</td>
<td>.032</td>
<td>.030</td>
<td>.030</td>
<td>.028</td>
</tr>
<tr>
<td></td>
<td>38</td>
<td>.053</td>
<td>.046</td>
<td>.051</td>
<td>.052</td>
<td>.052</td>
<td>.067</td>
<td>.053</td>
</tr>
<tr>
<td>4000</td>
<td>6</td>
<td>.045</td>
<td>.048</td>
<td>.049</td>
<td>.054</td>
<td>.058</td>
<td>.051</td>
<td>.065</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>.028</td>
<td>.031</td>
<td>.029</td>
<td>.028</td>
<td>.030</td>
<td>.038</td>
<td>.035</td>
</tr>
<tr>
<td></td>
<td>38</td>
<td>.031</td>
<td>.034</td>
<td>.039</td>
<td>.036</td>
<td>.040</td>
<td>.035</td>
<td>.037</td>
</tr>
<tr>
<td></td>
<td>51</td>
<td>.042</td>
<td>.051</td>
<td>.051</td>
<td>.040</td>
<td>.047</td>
<td>.047</td>
<td>.030</td>
</tr>
<tr>
<td>8000</td>
<td>6</td>
<td>.049</td>
<td>.055</td>
<td>.056</td>
<td>.048</td>
<td>.054</td>
<td>.055</td>
<td>.073</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>.034</td>
<td>.035</td>
<td>.036</td>
<td>.030</td>
<td>.032</td>
<td>.040</td>
<td>.041</td>
</tr>
<tr>
<td></td>
<td>38</td>
<td>.033</td>
<td>.035</td>
<td>.035</td>
<td>.037</td>
<td>.037</td>
<td>.025</td>
<td>.047</td>
</tr>
<tr>
<td></td>
<td>51</td>
<td>.034</td>
<td>.043</td>
<td>.035</td>
<td>.040</td>
<td>.037</td>
<td>.035</td>
<td>.038</td>
</tr>
<tr>
<td></td>
<td>83</td>
<td>.043</td>
<td>.042</td>
<td>.050</td>
<td>.048</td>
<td>.042</td>
<td>.054</td>
<td>.046</td>
</tr>
</tbody>
</table>
Table: Null Hypothesis that $F_\epsilon(-1|1)$ equals lower bound of identified set.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$p$</th>
<th>100</th>
<th>400</th>
<th>1600</th>
<th>4900</th>
<th>$100^2$</th>
<th>$225^2$</th>
<th>$317^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>6</td>
<td>.020</td>
<td>.019</td>
<td>.021</td>
<td>.021</td>
<td>.022</td>
<td>.019</td>
<td>.021</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>.037</td>
<td>.029</td>
<td>.029</td>
<td>.033</td>
<td>.033</td>
<td>.031</td>
<td>.030</td>
</tr>
<tr>
<td>2000</td>
<td>6</td>
<td>.030</td>
<td>.025</td>
<td>.033</td>
<td>.032</td>
<td>.033</td>
<td>.027</td>
<td>.039</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>.023</td>
<td>.021</td>
<td>.028</td>
<td>.027</td>
<td>.025</td>
<td>.027</td>
<td>.020</td>
</tr>
<tr>
<td></td>
<td>38</td>
<td>.048</td>
<td>.039</td>
<td>.043</td>
<td>.045</td>
<td>.047</td>
<td>.062</td>
<td>.046</td>
</tr>
<tr>
<td>4000</td>
<td>6</td>
<td>.034</td>
<td>.034</td>
<td>.038</td>
<td>.042</td>
<td>.046</td>
<td>.035</td>
<td>.058</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>.023</td>
<td>.026</td>
<td>.024</td>
<td>.022</td>
<td>.025</td>
<td>.032</td>
<td>.028</td>
</tr>
<tr>
<td></td>
<td>38</td>
<td>.026</td>
<td>.029</td>
<td>.033</td>
<td>.032</td>
<td>.035</td>
<td>.032</td>
<td>.033</td>
</tr>
<tr>
<td></td>
<td>51</td>
<td>.038</td>
<td>.044</td>
<td>.045</td>
<td>.034</td>
<td>.042</td>
<td>.041</td>
<td>.027</td>
</tr>
<tr>
<td>8000</td>
<td>6</td>
<td>.040</td>
<td>.046</td>
<td>.048</td>
<td>.040</td>
<td>.046</td>
<td>.050</td>
<td>.061</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>.028</td>
<td>.028</td>
<td>.032</td>
<td>.025</td>
<td>.027</td>
<td>.032</td>
<td>.034</td>
</tr>
<tr>
<td></td>
<td>38</td>
<td>.027</td>
<td>.029</td>
<td>.030</td>
<td>.032</td>
<td>.032</td>
<td>.021</td>
<td>.043</td>
</tr>
<tr>
<td></td>
<td>51</td>
<td>.029</td>
<td>.036</td>
<td>.028</td>
<td>.034</td>
<td>.033</td>
<td>.030</td>
<td>.031</td>
</tr>
<tr>
<td></td>
<td>83</td>
<td>.038</td>
<td>.035</td>
<td>.046</td>
<td>.041</td>
<td>.034</td>
<td>.048</td>
<td>.042</td>
</tr>
</tbody>
</table>
**Power Curves**

**Figure:** Power for 10% nominal level test

<table>
<thead>
<tr>
<th>$d = 16, p = 6$</th>
<th>$d = 1600, p = 6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = 2000$</td>
<td>$n = 4000$</td>
</tr>
</tbody>
</table>

Null hypothesis ($\gamma$)

Rejection probability

$\lambda$

- $\lambda_n^b$
- $\lambda_n^r$
- $\lambda_n$
- $0$

Fang, Santos, Shaikh, Torgovitsky. March 30, 2022. UCLA
Conclusion

Summary

- Mapped problems of interest into tests of $\beta(P) = Ax$ for some $x \geq 0$.
- Obtained new geometric characteristic of the null hypothesis.
- Derived test that can be evaluated by solving linear programs.
- Alternative tests also follow from geometric characterization.
- Immediate extension to (some) alternative sampling frameworks.