# Mis-Classified, Binary, Endogenous Regressors: Identification and Inference


Francis J. DiTraglia

Camilo García-Jimeno

University of Pennsylvania


January 18th, 2018

Mis-Classified, Binary, Endogenous Regressors: Identification and Inference

Francis J. DiTraglia
Camilo García-Jimeno
University of Pennsylvania

January 18th, 2018

Thanks for inviting me etc. Very nice meeting with so many of you today. For those I haven't spoken with, let me introduce myself. Applied econometrician. Interested in issues related to model selection, instrumental variables, inference in non-standard settings, and measurement error. Also do applied work, particularly in structural empirical micro. This paper is part of a research agenda on measurement error. Joint work with my co-author Camilo.

# What is the effect of $T^*$?

$$y = c(\mathbf{x}) + \beta(\mathbf{x})T^* + \varepsilon$$

- $y$ – Outcome of interest
- $T^*$ – Unobserved, endogenous binary regressor
- $T$ – Observed, mis-measured binary surrogate for $T^*$
- $\mathbf{x}$ – Exogenous covariates
- $z$ – Discrete (typically binary) instrumental variable

(Additively Separable $\varepsilon$ and binary $T^*$ $\Rightarrow$ linear model given $\mathbf{x}$)

What is the effect of $T^*$?

$$y = c(\mathbf{x}) + \beta(\mathbf{x})T^* + \varepsilon$$

- $y$ – Outcome of interest
- $T^*$ – Unobserved, endogenous binary regressor
- $T$ – Observed, mis-measured binary surrogate for $T^*$
- $\mathbf{x}$ – Exogenous covariates
- $z$ – Discrete (typically binary) instrumental variable

(Additively Separable $\varepsilon$ and binary $T^*$ $\Rightarrow$ linear model given $\mathbf{x}$)

No loss of generality from conditionally linear model provided $\varepsilon$ is additively separable:

$$y = h(T^*, \mathbf{x}) + \varepsilon$$

$$y = c(\mathbf{x}) + \beta(\mathbf{x})T^* + \varepsilon$$

$$\beta(\mathbf{x}) = h(1, \mathbf{x}) - h(0, \mathbf{x})$$

$$c(\mathbf{x}) = h(0, \mathbf{x})$$

# Using a discrete IV to learn about $\beta(\mathbf{x})$

$$y = c(\mathbf{x}) + \beta(\mathbf{x})T^* + \varepsilon$$

## Constributions of This Paper

1. Show that only existing point identification result is incorrect.
2. Derive sharp identified set for $\beta(\mathbf{x})$ under standard assumptions.
3. Prove point identification of $\beta(\mathbf{x})$ under slightly stronger assumptions.
4. Point out problem of weak identification in mis-classification models, develop identification-robust inference for $\beta(\mathbf{x})$.

Using a discrete IV to learn about $\beta(\mathbf{x})$

$$y = c(\mathbf{x}) + \beta(\mathbf{x})T^* + \varepsilon$$

Constributions of This Paper

1. Show that only existing point identification result is incorrect.
2. Derive sharp identified set for $\beta(\mathbf{x})$ under standard assumptions.
3. Prove point identification of $\beta(\mathbf{x})$ under slightly stronger assumptions.
4. Point out problem of weak identification in mis-classification models, develop identification-robust inference for $\beta(\mathbf{x})$.

Questions this paper sets out to answer. Simple model. If $T^*$ were continuous and measurement error were classical, this would be trivially easy. But, as we'll see things are much more complicated when $T^*$ is binary. First some motivating examples.

# Example: Smoking and Birthweight (SNAP Trial)

Coleman et al. (N Engl J Med, 2012)

RCT with 1050 pregnant smokers in England: 521 given nicotine patches, the rest given placebo patches.

- $y$ – Birthweight
- $T^*$ – True smoking behavior
- $T$ – Self-reported smoking behavior
- $\mathbf{x}$ – Mother characteristics
- $z$ – Indicator of nicotine patch

# Example: Schooling and Test Scores

Burde & Linden (2013, AEJ Applied)

RCT in Afghanistan: schools built in randomly selected villages. In treatment villages only some girls attend school; in control villages some girls attend school elsewhere.

- $y$ – Girl's score on math and language test
- $T^*$ – Girl's true school attendance
- $T$ – Parent's report of child's school attendance
- $\mathbf{x}$ – Child and household characteristics
- $z$ – School built in village

# Related Literature

## Continuous Regressor

Lewbel (1997, 2012), Schennach (2004, 2007), Chen et al. (2005), Hu & Schennach (2008), Song (2015), Hu et al. (2015)...

## Binary/Discrete, "Exogenous"

Aigner (1973), Bollinger (1996), Kane et al. (1999), Black et al. (2000), Frazis & Loewenstein (2003), Mahajan (2006), Lewbel (2007), Hu (2008), Molinari (2008)

## Binary, Endogenous Regressor

Mahajan (2006),

Shiu (2015), Denteh et al. (2016), Ura (2016), Calvi et al. (2017)

Related Literature

Continuous Regressor

Lewbel (1997, 2012), Schennach (2004, 2007), Chen et al. (2005), Hu & Schennach (2008), Song (2015), Hu et al. (2015)...

Binary/Discrete, "Exogenous"

Aigner (1973), Bollinger (1996), Kane et al. (1999), Black et al. (2000), Frazis & Loewenstein (2003), Mahajan (2006), Lewbel (2007), Hu (2008), Molinari (2008)

Binary, Endogenous Regressor

Mahajan (2006),

Shiu (2015), Denteh et al. (2016), Ura (2016), Calvi et al. (2017)

Very large literature on measurement error. Can't summarize everything here. But I want to draw out a few themes that will be relevant later and point out one paper in particular:

- Continuous vs. discrete
- Classical vs. non-classical
- Exogenous vs. endogenous

Don't spend too much time on other papers. Just say there was almost no work on the endogenous case when we started this. Recently some work that complements but doesn't overlap with ours. Focus on Mahajan paper and give the capsule explanation.

# “Baseline” Assumptions I – Model & Instrument

Additively Separable Model

$y = c(\mathbf{x}) + \beta(\mathbf{x})T^* + \varepsilon, \quad \mathbb{E}[\varepsilon] = 0$

Valid & Relevant Instrument: $z \in \{0, 1\}$

- $\mathbb{P}(T^* = 1|\mathbf{x}, z = 1) \neq \mathbb{P}(T^* = 1|\mathbf{x}, z = 0)$
- $\mathbb{E}[\varepsilon|\mathbf{x}, z] = 0$
- $0 < \mathbb{P}(z = 1|\mathbf{x}) < 1$

If $T^*$ were observed, these conditions would identify $\beta$.

# “Baseline” Assumptions II – Measurement Error

## Notation: Mis-classification Rates

“↑” $\alpha_0(\mathbf{x}, z) \equiv \mathbb{P}\left(T = 1 \middle| T^* = 0, \mathbf{x}, z\right)$

“↓” $\alpha_1(\mathbf{x}, z) \equiv \mathbb{P}\left(T = 0 \middle| T^* = 1, \mathbf{x}, z\right)$

## Mis-classification unaffected by $z$

$$\alpha_0(\mathbf{x}, z) = \alpha_0(\mathbf{x}), \quad \alpha_1(\mathbf{x}, z) = \alpha_1(\mathbf{x})$$

## Extent of Mis-classification

$\alpha_0(\mathbf{x}) + \alpha_1(\mathbf{x}) < 1$ ($T$ is positively correlated with $T^*$)

## Non-differential Mis-classification

$$\mathbb{E}[\varepsilon|\mathbf{x}, z, T^*, T] = \mathbb{E}[\varepsilon|\mathbf{x}, z, T^*]$$

# Identification Results from the Literature

## Mahajan (2006) Theorem 1, Frazis & Loewenstein (2003)

$\mathbb{E}[\varepsilon|\mathbf{x}, z, T^*] = 0$, plus "Baseline" $\implies \beta(\mathbf{x})$ identified

Requires $(T^*, z)$ jointly exogenous.

## Mahajan (2006) A.2

$\mathbb{E}[\varepsilon|\mathbf{x}, z, T^*, T] = \mathbb{E}[\varepsilon|\mathbf{x}, T^*]$, plus "Baseline" $\implies \beta(\mathbf{x})$ identified

Allows $T^*$ endogenous, but we prove this claim is false.

## Open Question

Can we identify $\beta(\mathbf{x})$ when $T^*$ is endogenous? If so, how?

Identification Results from the Literature

Mahajan (2006) Theorem 1, Frazis & Loewenstein (2003)

$\mathbb{E}[\varepsilon|\mathbf{x}, z, T^*] = 0$, plus "Baseline" $\implies$ $\beta(\mathbf{x})$ identified

Requires $(T^*, z)$ jointly exogenous.

Mahajan (2006) A.2

$\mathbb{E}[\varepsilon|\mathbf{x}, z, T^*, T] = \mathbb{E}[\varepsilon|\mathbf{x}, T^*]$, plus "Baseline" $\implies$ $\beta(\mathbf{x})$ identified

Allows $T^*$ endogenous, but we prove this claim is false.

Open Question

Can we identify $\beta(\mathbf{x})$ when $T^*$ is endogenous? If so, how?

Next: show you two known results that will play a role in what comes later: (1) simple bounds for mis-classification using only assumption that $z$ doesn't change $\alpha_0, \alpha_1$; (2) effect of mis-classification on IV estimator.

# First-stage Probabilities & Mis-classification Bounds

| Unobserved | Observed |
|---|---|
| $p_k^*(\mathbf{x}) \equiv \mathbb{P}(T^* = 1 \vert \mathbf{x}, z = k)$ | $p_k(\mathbf{x}) \equiv \mathbb{P}(T = 1 \vert \mathbf{x}, z = k)$ |

## Relationship

$$p_k^*(\mathbf{x}) = \frac{p_k(\mathbf{x}) - \alpha_0(\mathbf{x})}{1 - \alpha_0(\mathbf{x}) - \alpha_1(\mathbf{x})}, \quad k = 0, 1$$

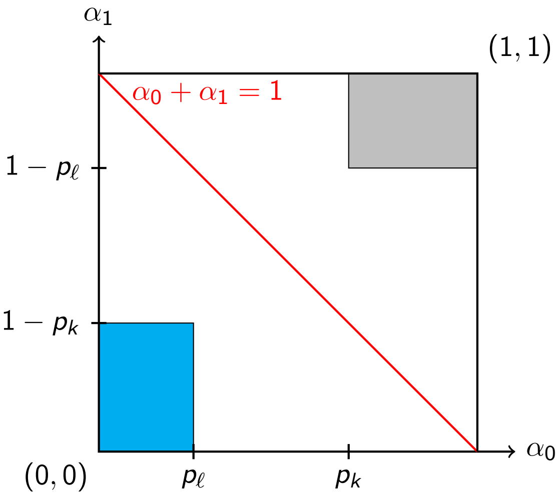
$z$ does not affect $(\alpha_0, \alpha_1)$; denominator $\neq 0$

## Bounds for Mis-classification

$$\alpha_0(\mathbf{x}) \leq p_k(\mathbf{x}) \leq 1 - \alpha_1(\mathbf{x}), \quad k = 0, 1$$

$\alpha_0(\mathbf{x}) + \alpha_1(\mathbf{x}) < 1$

$\alpha_0 \leq \min_k \{p_k\}, \quad \alpha_1 \leq \min_k \{1 - p_k\}$



# What does IV estimate under mis-classification?

## Unobserved

$$\beta(\mathbf{x}) = \frac{\mathbb{E}[y|\mathbf{x}, z=1] - \mathbb{E}[y|\mathbf{x}, z=0]}{p_1^*(\mathbf{x}) - p_0^*(\mathbf{x})}$$

## Wald (Observed)

$$\frac{\mathbb{E}[y|\mathbf{x}, z=1] - \mathbb{E}[y|\mathbf{x}, z=0]}{p_1(\mathbf{x}) - p_0(\mathbf{x})} = \beta(\mathbf{x}) \left[\frac{p_1^*(\mathbf{x}) - p_0^*(\mathbf{x})}{p_1(\mathbf{x}) - p_0(\mathbf{x})}\right] = \frac{\beta(\mathbf{x})}{1 - \alpha_0(\mathbf{x}) - \alpha_1(\mathbf{x})}$$

$$\boxed{p_1^*(\mathbf{x}) - p_0^*(\mathbf{x}) = \frac{p_1(\mathbf{x}) - \alpha_0(\mathbf{x})}{1 - \alpha_0 - \alpha_1(\mathbf{x})} - \frac{p_0(\mathbf{x}) - \alpha_0(\mathbf{x})}{1 - \alpha_0 - \alpha_1(\mathbf{x})} = \frac{p_1(\mathbf{x}) - p_0(\mathbf{x})}{1 - \alpha_0(\mathbf{x}) - \alpha_1(\mathbf{x})}}$$

# Partial Identification for $\beta(\mathbf{x})$

## Known Result

- $\beta(\mathbf{x})$ is between Wald and Reduced form; same sign as Wald.
- Doesn't rely on non-differential assumption or additive sep.
- Frazis & Loewenstein (2003), Ura (2016), ...

## Non-differential Assumption

- $\mathbb{E}[\varepsilon|\mathbf{x}, T^*, T, z] = \mathbb{E}[\varepsilon|\mathbf{x}, T^*, z]$
- Used in literature to identify $\beta(\mathbf{x})$ when $T^*$ is exogenous.
- Does it restrict the identified set when $T^*$ is endogenous?
- Is $\beta(\mathbf{x})$ identified under the baseline assumptions?

# Restrictions from Non-differential Mis-classification

Suppress $\mathbf{x}$ for simplicity

## Notation

- $F_{tk} \equiv$ Observed conditional CDF of $y|(T = t, z = k)$
- $F_{tk}^{t^*} \equiv$ Unobserved conditional CDF of $y|(T^* = t^*, T = t, z = k)$
- $r_{tk} \equiv \mathbb{P}(T^* = 1|T = t, z = k)$ observed given $(\alpha_0, \alpha_1)$

## Law of Total Probability

$$F_{tk} = (1 - r_{tk})F_{tk}^0 + r_{tk}F_{tk}^1$$

Given $(\alpha_0, \alpha_1)$ can we construct $(F_{tk}^0, F_{tk}^1)$ to satisfy the mixture model?

# Restrictions from Non-differential Mis-classification

## Notation

- $r_{tk} \equiv \mathbb{P}(T^* = 1 | T = t, z = k)$ observed given $(\alpha_0, \alpha_1)$
- $z_k$ as shorthand for $z = k$

## Iterated Expectations over $T^*$

$$\mathbb{E}(y|T = 0, z_k) = (1 - r_{0k})\mathbb{E}(y|T^* = 0, T = 0, z_k) + r_{0k}\mathbb{E}(y|T^* = 1, T = 0, z_k)$$
$$\mathbb{E}(y|T = 1, z_k) = (1 - r_{1k})\mathbb{E}(y|T^* = 0, T = 1, z_k) + r_{1k}\mathbb{E}(y|T^* = 1, T = 1, z_k)$$

- $(\alpha_0, \alpha_1)$ pin down $r_{tk}$

# Restrictions from Non-differential Mis-classification

## Notation

- $r_{tk} \equiv \mathbb{P}(T^* = 1 | T = t, z = k)$ observed given $(\alpha_0, \alpha_1)$
- $z_k$ as shorthand for $z = k$

## Iterated Expectations over $T^*$ and Non-diff.

$$\mathbb{E}(y|T = 0, z_k) = (1 - r_{0k})\mathbb{E}(y|T^* = 0, z_k) + r_{0k}\mathbb{E}(y|T^* = 1, z_k)$$
$$\mathbb{E}(y|T = 1, z_k) = (1 - r_{1k})\mathbb{E}(y|T^* = 0, z_k) + r_{1k}\mathbb{E}(y|T^* = 1, z_k)$$

- $(\alpha_0, \alpha_1)$ pin down $r_{tk}$
- Non-diff. $\implies$ $(\alpha_0, \alpha_1)$ pin down $\mathbb{E}(y|T^* = t^*, z = k)$
- $\mathbb{E}(y|T^*, z = k)$ are the means of $(F^0_{tk}, F^1_{tk})$
- Can we satisfy the mixture model?

# Restrictions from Non-differential Mis-classification

## Equivalent Problem

Given an observed CDF $F$ and a probability $p$, do there exist CDFs $(G, H)$ such that $F = (1-p)G + pH$ and the mean of $H$ is $\mu$?

## Necessary and Sufficient Condition if $F$ is Continuous

$$\int_{-\infty}^{F^{-1}(p)} x\, f(x)\, dx \leq \ p\mu \ \leq \int_{F^{-1}(1-p)}^{+\infty} x\, f(x)\, dx$$

## Sharp Identified Set

Includes only those $(\alpha_0, \alpha_1)$ at which the preceding condition is satisfied jointly for the mixtures $F_{tk} = (1 - r_{tk})F_{tk}^0 + r_{tk}F_{tk}^1$.

# Point Identification: 1st Ingredient

## Reparameterization

$$\theta_1(\mathbf{x}) = \beta(\mathbf{x}) / \left[1 - \alpha_0(\mathbf{x}) - \alpha_1(\mathbf{x})\right]$$

$$\theta_2(\mathbf{x}) = \left[\theta_1(\mathbf{x})\right]^2 \left[1 + \alpha_0(\mathbf{x}) - \alpha_1(\mathbf{x})\right]$$

$$\theta_3(\mathbf{x}) = \left[\theta_1(\mathbf{x})\right]^3 \left[\left\{1 - \alpha_0(\mathbf{x}) - \alpha_1(\mathbf{x})\right\}^2 + 6\alpha_0(\mathbf{x})\left\{1 - \alpha_1(\mathbf{x})\right\}\right]$$

$$\boxed{\beta(\mathbf{x}) = 0 \iff \theta_1(\mathbf{x}) = \theta_2(\mathbf{x}) = \theta_3(\mathbf{x}) = 0}$$

## Lemma

Baseline Assumptions $\implies$ $\text{Cov}(y, z|\mathbf{x}) = \theta_1(\mathbf{x})\text{Cov}(z, T|\mathbf{x})$.

# Point Identification: 2nd Ingredient

## Assumption (II)

$\mathbb{E}[\varepsilon^2|\mathbf{x}, z] = \mathbb{E}[\varepsilon^2|\mathbf{x}]$

## Lemma

(Baseline) + (II) $\implies$

$$\text{Cov}(y^2, z|\mathbf{x}) = 2\text{Cov}(yT, z|\mathbf{x})\theta_1(\mathbf{x}) - \text{Cov}(T, z|\mathbf{x})\theta_2(\mathbf{x})$$

## Corollary

(Baseline) + (II) + $[\beta(\mathbf{x}) \neq 0] \implies [\alpha_1(\mathbf{x}) - \alpha_0(\mathbf{x})]$ is identified.

Hence, $\beta(\mathbf{x})$ is identified if mis-classification is one-sided.

# Point Identification: 1st Ingredient

### Assumption (III)

(i) $\mathbb{E}[\varepsilon^2|\mathbf{x}, z, T^*, T] = \mathbb{E}[\varepsilon^2|\mathbf{x}, z, T^*]$

(ii) $\mathbb{E}[\varepsilon^3|\mathbf{x}, z] = \mathbb{E}[\varepsilon^3|\mathbf{x}]$

### Lemma

(Baseline) + (II) + (III) $\implies$

$$\text{Cov}(y^3, z|\mathbf{x}) = 3\text{Cov}(y^2 T, z|\mathbf{x})\theta_1(\mathbf{x}) - 3\text{Cov}(yT, z|\mathbf{x})\theta_2(\mathbf{x}) + \text{Cov}(T, z|\mathbf{x})\theta_3(\mathbf{x})$$

# Point Identification Result

## Theorem

(Baseline) + (II) + (III) $\implies \beta(\mathbf{x})$ is point identified. If $\beta(\mathbf{x}) \neq 0$, then $\alpha_0(\mathbf{x})$ and $\alpha_1(\mathbf{x})$ are likewise point identified.

## Proof Sketch

1. $\beta(\mathbf{x}) = 0 \iff \theta_1(\mathbf{x}) = 0$ so suppose this is not the case.
2. Lemmas: full-rank linear system in $\theta_1(\mathbf{x}), \theta_2(\mathbf{x}), \theta_3(\mathbf{x})$ & observables.
3. Non-linear eqs. relating $\theta_1(\mathbf{x}), \theta_2(\mathbf{x}), \theta_3(\mathbf{x})$ to $\beta(\mathbf{x})$ and $\alpha_0(\mathbf{x}), \alpha_1(\mathbf{x})$. Show that solution exists and is unique.

## Sufficient Conditions for (II) and (III)

(i) $T$ is conditionally independent of $(\varepsilon, z)$ given $(T^*, \mathbf{x})$

(ii) $z$ is conditionally independent of $\varepsilon$ given $\mathbf{x}$

# Just-Identified System of Moment Equalities

Suppress dependence on $\mathbf{x}$ to simplify the notation from here on...

## Collect Lemmas from Above:

$$\begin{aligned}
\text{Cov}(y,z) - \text{Cov}(T,z)\theta_1 &= 0 \\
\text{Cov}(y^2,z) - 2\text{Cov}(yT,z)\theta_1 + \text{Cov}(T,z)\theta_2 &= 0 \\
\text{Cov}(y^3,z) - 3\text{Cov}(y^2T,z)\theta_1 + 3\text{Cov}(yT,z)\theta_2 - \text{Cov}(T,z)\theta_3 &= 0
\end{aligned}$$

## Notation: Observed Data Vector

$$\mathbf{w}_i' = (T_i, y_i, y_iT_i, y_i^2, y_i^2T_i, y_i^3)$$

# Just-Identified System of Moment Equalities

$$\boxed{\mathbb{E}\left[\left(\boldsymbol{\Psi}'(\boldsymbol{\theta})\mathbf{w}_i - \boldsymbol{\kappa}\right) \otimes \begin{pmatrix} 1 \\ z_i \end{pmatrix}\right] = \mathbf{0}}$$

$$\boldsymbol{\Psi} = \begin{bmatrix} \boldsymbol{\psi}_1 & \boldsymbol{\psi}_2 & \boldsymbol{\psi}_3 \end{bmatrix} \qquad \boldsymbol{\kappa} = (\kappa_1, \kappa_2, \kappa_3)' \equiv \text{“Intercepts”}$$

$$\boldsymbol{\psi}_1' = \begin{bmatrix} -\theta_1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \theta_1 = \beta/(1-\alpha_0-\alpha_1)$$

$$\boldsymbol{\psi}_2' = \begin{bmatrix} \theta_2 & 0 & -2\theta_1 & 1 & 0 & 0 \end{bmatrix} \qquad \theta_2 = \theta_1^2\left[1+\alpha_0-\alpha_1\right]$$

$$\boldsymbol{\psi}_3' = \begin{bmatrix} -\theta_3 & 0 & 3\theta_2 & 0 & -3\theta_1 & 1 \end{bmatrix} \qquad \theta_3 = \theta_1^3\left[(1-\alpha_0-\alpha_1)^2 + 6\alpha_0(1-\alpha_1)\right]$$

## Weak Identification Problem

Moment conditions are uninformative about $(\alpha_0, \alpha_1)$ when $\beta$ is small.

# Non-standard Inference Problem

- $\beta$ small $\Rightarrow$ moment equalities uninformative about $(\alpha_0, \alpha_1)$
- $(\alpha_0, \alpha_1)$ could be on the boundary of the parameter space
- Partial identification bounds remain informative even if $\beta$ is small or zero
- Same problem for other estimators from the literature but hasn't been pointed out...

## Our Approach

Identification-robust inference combining equality and inequality moment conditions based on generalized moment selection (GMS)

# Inference With Moment Equalities and Inequalities

## Moment Conditions

$$\mathbb{E}\left[m_j(\mathbf{w}_i, \vartheta_0)\right] \geq 0, \quad j = 1, \cdots, J$$

$$\mathbb{E}\left[m_j(\mathbf{w}_i, \vartheta_0)\right] = 0, \quad j = J+1, \cdots, J+K$$

## Test Statistic

$$T_n(\vartheta) = \sum_{j=1}^{J} \left[\frac{\sqrt{n}\ \bar{m}_{n,j}(\vartheta)}{\widehat{\sigma}_{n,j}(\vartheta)}\right]_-^2 + \sum_{j=J+1}^{J+K} \left[\frac{\sqrt{n}\ \bar{m}_{n,j}(\vartheta)}{\widehat{\sigma}_{n,j}(\vartheta)}\right]^2$$

$$[x]_- = \min\{x, 0\}$$

$$\bar{m}_{n,j}(\vartheta) = n^{-1} \sum_{i=1}^{n} m_j(\mathbf{w}_i, \vartheta)$$

$$\widehat{\sigma}_{n,j}^2(\vartheta) = \text{consistent est. of AVAR}\left[\sqrt{n}\ \bar{m}_{n,j}(\vartheta)\right]$$

# Moment Inequalities: Part I

$\alpha_0(\mathbf{x}) \leq p_k \leq 1 - \alpha_1$ becomes $\mathbb{E}\left[m^I_{1k}(\mathbf{w}_i, \vartheta)\right] \geq \mathbf{0}$ for all $k$ where

$$m^I_{1k}(\mathbf{w}_i, \vartheta) \equiv \begin{bmatrix} \mathbf{1}(z_i = k)(T - \alpha_0) \\ \mathbf{1}(z_i = k)(1 - T_i - \alpha_1) \end{bmatrix}$$

# Moment Inequalities: Part II

For all $k$, we have $\mathbb{E}[m^{l}_{2k}\left(\mathbf{w}_i, \vartheta, \mathbf{q}_k)\right] \geq 0$ where

$$m^{l}_{2k}\left(\mathbf{w}_i, \vartheta, \mathbf{q}_k\right) \equiv \begin{bmatrix} y_i \mathbf{1}\left(z_i = k\right)\left\{(T_i - \alpha_0) - \mathbf{1}(y_i \leq \underline{q}_{0k})(1 - T_i)\left(\frac{1-\alpha_0-\alpha_1}{\alpha_1}\right)\right\} \\ -y_i \mathbf{1}(z_i = k)\left\{(T_i - \alpha_0) - \mathbf{1}(y_i > \overline{q}_{0k})(1 - T_i)\left(\frac{1-\alpha_0-\alpha_1}{\alpha_1}\right)\right\} \\ y_i \mathbf{1}\left(z_i = k\right)\left\{(T_i - \alpha_0) - \mathbf{1}(y_i \leq \underline{q}_{1k}) T_i \left(\frac{1-\alpha_0-\alpha_1}{1-\alpha_1}\right)\right\} \\ -y_i \mathbf{1}(z_i = k)\left\{(T_i - \alpha_0) - \mathbf{1}(y_i > \overline{q}_{1k}) T_i \left(\frac{1-\alpha_0-\alpha_1}{1-\alpha_1}\right)\right\} \end{bmatrix}$$

and $\mathbf{q}_k \equiv (\underline{q}_{0k}, \overline{q}_{0k}, \underline{q}_{1k}, \overline{q}_{1k})'$ defined by $\mathbb{E}[h^{l}_k(\mathbf{w}_i, \vartheta, \mathbf{q}_k)] = 0$ with

$$h^{l}_k(\mathbf{w}_i, \vartheta, \mathbf{q}_k) = \begin{bmatrix} \mathbf{1}(y_i \leq \underline{q}_{0k})\mathbf{1}(z_i = k)(1 - T_i) - \left(\frac{\alpha_1}{1-\alpha_0-\alpha_1}\right)\mathbf{1}(z_i = k)(T_i - \alpha_0) \\ \mathbf{1}(y_i \leq \overline{q}_{0k})\mathbf{1}(z_i = k)(1 - T_i) - \left(\frac{1-\alpha_0}{1-\alpha_0-\alpha_1}\right)\mathbf{1}(z_i = k)(1 - T_i - \alpha_1) \\ \mathbf{1}(y_i \leq \underline{q}_{1k})\mathbf{1}(z_i = k) T_i - \left(\frac{1-\alpha_1}{1-\alpha_0-\alpha_1}\right)\mathbf{1}(z_i = k)(T_i - \alpha_0) \\ \mathbf{1}(y_i \leq \overline{q}_{1k})\mathbf{1}(z_i = k) T_i - \left(\frac{\alpha_0}{1-\alpha_0-\alpha_1}\right)\mathbf{1}(z_i = k)(1 - T_i - \alpha_1) \end{bmatrix}$$

# Inference via Generalized Moment Selection

Andrews & Soares (2010)

## Moment Selection Step

If $\dfrac{\sqrt{n}\,\bar{m}_{n,j}(\vartheta_0)}{\widehat{\sigma}_{n,j}(\vartheta_0)} > \sqrt{\log n}$ then drop inequality $j$

## Critical Value

- $\sqrt{n}\,\bar{m}_n(\vartheta_0) \to_d$ normal limit with covariance matrix $\Sigma(\vartheta_0)$
- Use this to bootstrap the limit distribution of the test statistic.

## Theoretical Guarantees

Uniformly valid test of $H_0\colon \vartheta = \vartheta_0$ regardless of whether $\vartheta_0$ is identified.

Not asymptotically conservative.

## Drawback

*Joint test* for the whole parameter vector but we're only interested in $\beta$

# Moment Equalities

Let $\boldsymbol{\vartheta} = (\alpha_0, \alpha_1)$ and $\boldsymbol{\gamma} = (\boldsymbol{\kappa}, \theta_1)$

$$\mathbb{E}[m^I(\mathbf{w}_i, \boldsymbol{\vartheta}_0, \mathbf{q}_0)] \geq \mathbf{0}, \quad \mathbb{E}[m^E(\mathbf{w}_i, \boldsymbol{\vartheta}_0, \boldsymbol{\gamma}_0)] = \mathbf{0} \tag{1}$$

where $m^I = (m_1^{I'}, m_2^{I'})'$ and

$$m^E(\mathbf{w}_i, \boldsymbol{\vartheta}_0, \boldsymbol{\gamma}_0) = \begin{bmatrix} \left\{\boldsymbol{\psi}_2'(\theta_1, \alpha_0, \alpha_1)\mathbf{w}_i - \kappa_2\right\} z_i \\ \left\{\boldsymbol{\psi}_3'(\theta_1, \alpha_0, \alpha_1)\mathbf{w}_i - \kappa_3\right\} z_i \end{bmatrix}. \tag{2}$$

$$h^E(\mathbf{w}_i, \boldsymbol{\vartheta}, \boldsymbol{\gamma}) = \begin{bmatrix} \boldsymbol{\Psi}'(\theta_1, \alpha_0, \alpha_1)\mathbf{w}_i - \boldsymbol{\kappa} \\ \left\{\boldsymbol{\psi}_1'(\theta_1)\mathbf{w}_i - \kappa_1\right\} z_i \end{bmatrix}. \tag{3}$$

# Bonferroni-Based Inference Procedure

## Leverage Special Structure of Model

- $\beta$ only enters MCs through $\theta_1 = \beta/(1-\alpha_0-\alpha_1)$
- Inference for $\theta_1$ is standard if $z$ is a strong IV.
- $(\boldsymbol{\kappa}, \mathbf{q})$ strongly identified under null for $(\alpha_0, \alpha_1)$

## Procedure

1. Concentrate out $(\theta_1, \boldsymbol{\kappa}, \boldsymbol{q}) \implies$ joint GMS test for $(\alpha_0, \alpha_1)$
2. Invert $\implies (1-\delta_1)\times 100\%$ confidence set for $(\alpha_0, \alpha_1)$
3. Project $\implies$ CI for $(1-\alpha_0-\alpha_1)$
4. Construct standard $(1-\delta_2)\times 100\%$ IV CI for $\theta_1$
5. Bonferroni $\implies (1-\delta-\delta_2)\times 100\%$ CI for $\beta$

Short empirical illustration using Burde & Linden.

# Conclusion

## Summary

- Endogenous, mis-classified binary treatment.
- Usual (1st moment) IV assumption fails to identify $\beta$
- Derive sharp identified set.
- Stronger assumptions point identify $\beta$
- Identification-Robust Inference incorportating equality and inequality moment conditions.

## Extensions / Future Work

- Arbitrary discrete $T^*$
- Relax additive separability in panel setting?

Talk about how this paper fits into a research agenda concerning measurement error: the beliefs paper, this paper, returns to lying (with Arthur), and biased measurements of displacement in the paper with Camilo.

# Simulation DGP: $y = \beta T^* + \varepsilon$

Sample Size = 1000; Simulation Replications = 2000

## Errors

$(\varepsilon, \eta) \sim$ jointly normal, mean 0, variance 1, correlation 0.5.

## First-Stage

- Half of observations have $z = 1$, the rest have $z = 0$.
- $T^* = \mathbf{1}\{\gamma_0 + \gamma_1 z + \eta > 0\}$
- $\mathbb{P}(T^* = 0|z = 1) = \mathbb{P}(T^* = 1|z = 0) = 0.15$

## Mis-classification

- $T|T^* = 0 \sim \text{Bernoulli}(\alpha_0)$
- $T|T^* = 1 \sim \text{Bernoulli}(1 - \alpha_1)$

| | | $\beta$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $\alpha_0$ | $\alpha_1$ | 0 | 0.25 | 0.5 | 0.75 | 1 | 1.5 | 2 | 3 |
| 0.0 | 0.0 | 27 | 33 | 30 | 14 | 1 | 0 | 0 | 0 |
| | 0.1 | 27 | 32 | 29 | 13 | 2 | 0 | 0 | 0 |
| | 0.2 | 26 | 33 | 32 | 15 | 4 | 0 | 0 | 0 |
| | 0.3 | 26 | 34 | 30 | 17 | 5 | 0 | 0 | 0 |
| 0.1 | 0.0 | 26 | 32 | 31 | 14 | 2 | 0 | 0 | 0 |
| | 0.1 | 26 | 36 | 32 | 16 | 4 | 0 | 0 | 0 |
| | 0.2 | 27 | 35 | 31 | 18 | 8 | 0 | 0 | 0 |
| | 0.3 | 25 | 35 | 32 | 21 | 11 | 1 | 0 | 0 |
| 0.2 | 0.0 | 26 | 33 | 30 | 15 | 3 | 0 | 0 | 0 |
| | 0.1 | 26 | 33 | 30 | 19 | 6 | 0 | 0 | 0 |
| | 0.2 | 26 | 35 | 33 | 22 | 12 | 1 | 0 | 0 |
| | 0.3 | 26 | 35 | 33 | 26 | 15 | 3 | 0 | 0 |
| 0.3 | 0.0 | 26 | 32 | 32 | 16 | 6 | 0 | 0 | 0 |
| | 0.1 | 24 | 35 | 33 | 21 | 11 | 1 | 0 | 0 |
| | 0.2 | 26 | 32 | 35 | 27 | 15 | 4 | 0 | 0 |
| | 0.3 | 26 | 35 | 35 | 28 | 21 | 7 | 2 | 0 |

Table: Percentage of simulation replications for which the standard GMM CI fails to exist.

| | | $\beta$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $\alpha_0$ | $\alpha_1$ | 0 | 0.25 | 0.5 | 0.75 | 1 | 1.5 | 2 | 3 |
| 0.0 | 0.0 | 72 | 62 | 62 | 80 | 92 | 95 | 94 | 95 |
| | 0.1 | 72 | 62 | 63 | 79 | 92 | 95 | 96 | 95 |
| | 0.2 | 73 | 61 | 61 | 77 | 90 | 96 | 96 | 96 |
| | 0.3 | 73 | 59 | 62 | 76 | 88 | 95 | 96 | 95 |
| 0.1 | 0.0 | 73 | 63 | 60 | 78 | 91 | 95 | 96 | 96 |
| | 0.1 | 73 | 58 | 59 | 77 | 90 | 95 | 95 | 94 |
| | 0.2 | 73 | 59 | 61 | 75 | 86 | 95 | 95 | 94 |
| | 0.3 | 74 | 59 | 58 | 71 | 82 | 94 | 96 | 96 |
| 0.2 | 0.0 | 74 | 62 | 60 | 78 | 91 | 95 | 96 | 96 |
| | 0.1 | 73 | 60 | 61 | 74 | 87 | 95 | 96 | 94 |
| | 0.2 | 73 | 58 | 57 | 70 | 81 | 93 | 95 | 95 |
| | 0.3 | 73 | 58 | 56 | 66 | 78 | 92 | 95 | 96 |
| 0.3 | 0.0 | 74 | 62 | 60 | 76 | 89 | 95 | 96 | 96 |
| | 0.1 | 75 | 59 | 58 | 71 | 82 | 93 | 96 | 95 |
| | 0.2 | 74 | 61 | 56 | 65 | 78 | 90 | 96 | 96 |
| | 0.3 | 73 | 58 | 55 | 64 | 71 | 88 | 93 | 96 |

Table: Coverge of nominal 95% GMM CI, conditional on existence.

| $\alpha_0$ | $\alpha_1$ | $\beta$ = 0 | 0.25 | 0.5 | 0.75 | 1 | 1.5 | 2 | 3 |
|---|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.0 | 19.07 | 3.44 | 1.86 | 1.32 | 0.87 | 0.47 | 0.37 | 0.35 |
| | 0.1 | 17.52 | 3.47 | 1.92 | 1.41 | 1 | 0.61 | 0.51 | 0.46 |
| | 0.2 | 17.41 | 3.51 | 1.9 | 1.45 | 1.1 | 0.76 | 0.65 | 0.58 |
| | 0.3 | 18.23 | 3.34 | 1.92 | 1.48 | 1.24 | 0.91 | 0.79 | 0.7 |
| 0.1 | 0.0 | 17.13 | 3.51 | 1.86 | 1.38 | 0.97 | 0.61 | 0.51 | 0.46 |
| | 0.1 | 17.88 | 3.33 | 1.85 | 1.45 | 1.13 | 0.78 | 0.67 | 0.6 |
| | 0.2 | 17.37 | 3.36 | 1.95 | 1.54 | 1.24 | 0.97 | 0.85 | 0.75 |
| | 0.3 | 18.07 | 3.33 | 1.98 | 1.63 | 1.41 | 1.17 | 1.04 | 0.92 |
| 0.2 | 0.0 | 17.79 | 3.39 | 1.92 | 1.45 | 1.11 | 0.75 | 0.65 | 0.58 |
| | 0.1 | 18.98 | 3.43 | 1.96 | 1.54 | 1.26 | 0.97 | 0.84 | 0.75 |
| | 0.2 | 18.25 | 3.26 | 1.92 | 1.64 | 1.45 | 1.2 | 1.06 | 0.95 |
| | 0.3 | 19.03 | 3.31 | 2.02 | 1.75 | 1.66 | 1.49 | 1.33 | 1.19 |
| 0.3 | 0.0 | 18.27 | 3.48 | 1.87 | 1.5 | 1.25 | 0.9 | 0.79 | 0.7 |
| | 0.1 | 19.4 | 3.41 | 1.96 | 1.63 | 1.43 | 1.18 | 1.04 | 0.92 |
| | 0.2 | 18.22 | 3.56 | 1.96 | 1.74 | 1.67 | 1.49 | 1.35 | 1.19 |
| | 0.3 | 17.56 | 3.55 | 2.13 | 1.96 | 1.86 | 1.86 | 1.74 | 1.55 |

Table: Median width of nominal 95% GMM CI, conditional on existence.

| $\alpha_0$ | $\alpha_1$ | $\beta$ 0 | 0.25 | 0.5 | 0.75 | 1 | 1.5 | 2 | 3 |
|---|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.0 | 97.7 | 97.7 | 97.6 | 97.7 | 98.0 | 98.0 | 97.4 | 97.9 |
| | 0.1 | 98.0 | 98.7 | 98.8 | 99.1 | 98.8 | 98.4 | 97.1 | 96.4 |
| | 0.2 | 98.4 | 98.5 | 98.9 | 98.9 | 98.8 | 98.6 | 98.0 | 97.0 |
| | 0.3 | 98.5 | 98.8 | 98.8 | 99.0 | 98.7 | 98.4 | 97.8 | 97.5 |
| 0.1 | 0.0 | 98.1 | 98.5 | 98.3 | 98.8 | 98.8 | 98.4 | 96.8 | 95.7 |
| | 0.1 | 98.6 | 99.1 | 99.5 | 99.6 | 99.6 | 98.8 | 97.7 | 95.2 |
| | 0.2 | 99.0 | 99.3 | 99.7 | 99.8 | 99.7 | 98.9 | 97.5 | 95.7 |
| | 0.3 | 99.4 | 99.7 | 99.8 | 99.8 | 99.6 | 99.0 | 98.2 | 96.7 |
| 0.2 | 0.0 | 98.6 | 98.5 | 98.6 | 98.9 | 98.7 | 98.2 | 97.7 | 97.0 |
| | 0.1 | 99.0 | 99.5 | 99.7 | 99.7 | 99.4 | 99.0 | 98.1 | 96.5 |
| | 0.2 | 99.5 | 99.7 | 99.8 | 99.7 | 99.4 | 99.0 | 97.8 | 96.8 |
| | 0.3 | 99.7 | 99.8 | 99.8 | 99.8 | 99.5 | 99.0 | 98.7 | 97.7 |
| 0.3 | 0.0 | 98.7 | 98.7 | 98.8 | 98.7 | 98.7 | 98.2 | 98.1 | 97.6 |
| | 0.1 | 99.4 | 99.6 | 99.6 | 99.7 | 99.4 | 98.9 | 98.3 | 96.8 |
| | 0.2 | 99.8 | 99.8 | 99.7 | 99.8 | 99.5 | 99.1 | 98.5 | 97.8 |
| | 0.3 | 100.0 | 99.9 | 99.9 | 99.8 | 99.6 | 99.5 | 99.1 | 98.8 |

Table: Coverage (1 - size) of nominal 97.5% GMS joint test for $(\alpha_0, \alpha_1)$.

| $\alpha_0$ | $\alpha_1$ | $\beta$ = 0 | 0.25 | 0.5 | 0.75 | 1 | 1.5 | 2 | 3 |
|---|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.0 | 96 | 97 | 97 | 96 | 97 | 97 | 95 | 96 |
| | 0.1 | 97 | 99 | 99 | 99 | 99 | 100 | 100 | 99 |
| | 0.2 | 98 | 99 | 99 | 100 | 100 | 100 | 100 | 100 |
| | 0.3 | 97 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 0.1 | 0.0 | 97 | 99 | 99 | 99 | 100 | 100 | 100 | 98 |
| | 0.1 | 98 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | 0.2 | 98 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | 0.3 | 97 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 0.2 | 0.0 | 97 | 99 | 99 | 100 | 100 | 100 | 100 | 100 |
| | 0.1 | 98 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | 0.2 | 98 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | 0.3 | 98 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 0.3 | 0.0 | 97 | 99 | 100 | 100 | 100 | 100 | 100 | 100 |
| | 0.1 | 97 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | 0.2 | 98 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | 0.3 | 98 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

Table: Coverage of nominal $> 95\%$ Bonferroni CI for $\beta$

| $\alpha_0$ | $\alpha_1$ | $\beta$ = 0 | 0.25 | 0.5 | 0.75 | 1 | 1.5 | 2 | 3 |
|---|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.0 | 0.4 | 0.41 | 0.43 | 0.43 | 0.43 | 0.42 | 0.41 | 0.41 |
| | 0.1 | 0.45 | 0.47 | 0.54 | 0.59 | 0.63 | 0.7 | 0.75 | 0.86 |
| | 0.2 | 0.51 | 0.54 | 0.65 | 0.76 | 0.85 | 0.95 | 1.01 | 1.17 |
| | 0.3 | 0.58 | 0.62 | 0.79 | 0.95 | 1.07 | 1.17 | 1.24 | 1.48 |
| 0.1 | 0.0 | 0.45 | 0.47 | 0.54 | 0.59 | 0.63 | 0.7 | 0.76 | 0.88 |
| | 0.1 | 0.51 | 0.54 | 0.66 | 0.77 | 0.86 | 1.03 | 1.18 | 1.46 |
| | 0.2 | 0.58 | 0.63 | 0.8 | 0.98 | 1.12 | 1.38 | 1.55 | 1.88 |
| | 0.3 | 0.67 | 0.75 | 1 | 1.25 | 1.46 | 1.74 | 1.94 | 2.4 |
| 0.2 | 0.0 | 0.51 | 0.54 | 0.65 | 0.76 | 0.86 | 0.96 | 1.02 | 1.19 |
| | 0.1 | 0.58 | 0.63 | 0.81 | 0.99 | 1.14 | 1.42 | 1.64 | 2.08 |
| | 0.2 | 0.67 | 0.75 | 1.01 | 1.29 | 1.54 | 1.97 | 2.33 | 2.9 |
| | 0.3 | 0.81 | 0.91 | 1.3 | 1.7 | 2.09 | 2.73 | 3.13 | 3.9 |
| 0.3 | 0.0 | 0.58 | 0.62 | 0.8 | 0.95 | 1.09 | 1.18 | 1.25 | 1.5 |
| | 0.1 | 0.68 | 0.74 | 1.01 | 1.26 | 1.49 | 1.84 | 2.13 | 2.78 |
| | 0.2 | 0.81 | 0.91 | 1.3 | 1.7 | 2.11 | 2.8 | 3.4 | 4.48 |
| | 0.3 | 1.01 | 1.16 | 1.74 | 2.35 | 2.93 | 4.17 | 5.2 | 6.85 |

Table: Median width of nominal $> 95\%$ Bonferroni CI for $\beta$.

| $\alpha_0$ | $\alpha_1$ | $\beta$ 0 | 0.25 | 0.5 | 0.75 | 1 | 1.5 | 2 | 3 |
|---|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.0 | 0.4 | 0.41 | 0.43 | 0.43 | 0.43 | 0.42 | 0.41 | 0.41 |
| | 0.1 | 0.45 | 0.47 | 0.54 | 0.59 | 0.63 | 0.7 | 0.75 | 0.86 |
| | 0.2 | 0.51 | 0.54 | 0.65 | 0.76 | 0.85 | 0.95 | 1.01 | 1.17 |
| | 0.3 | 0.58 | 0.62 | 0.79 | 0.95 | 1.07 | 1.17 | 1.24 | 1.48 |
| 0.1 | 0.0 | 0.45 | 0.47 | 0.54 | 0.59 | 0.63 | 0.7 | 0.76 | 0.88 |
| | 0.1 | 0.51 | 0.54 | 0.66 | 0.77 | 0.86 | 1.03 | 1.18 | 1.46 |
| | 0.2 | 0.58 | 0.63 | 0.8 | 0.98 | 1.12 | 1.38 | 1.55 | 1.88 |
| | 0.3 | 0.67 | 0.75 | 1 | 1.25 | 1.46 | 1.74 | 1.94 | 2.4 |
| 0.2 | 0.0 | 0.51 | 0.54 | 0.65 | 0.76 | 0.86 | 0.96 | 1.02 | 1.19 |
| | 0.1 | 0.58 | 0.63 | 0.81 | 0.99 | 1.14 | 1.42 | 1.64 | 2.08 |
| | 0.2 | 0.67 | 0.75 | 1.01 | 1.29 | 1.54 | 1.97 | 2.33 | 2.9 |
| | 0.3 | 0.81 | 0.91 | 1.3 | 1.7 | 2.09 | 2.73 | 3.13 | 3.9 |
| 0.3 | 0.0 | 0.58 | 0.62 | 0.8 | 0.95 | 1.09 | 1.18 | 1.25 | 1.5 |
| | 0.1 | 0.68 | 0.74 | 1.01 | 1.26 | 1.49 | 1.84 | 2.13 | 2.78 |
| | 0.2 | 0.81 | 0.91 | 1.3 | 1.7 | 2.11 | 2.8 | 3.4 | 4.48 |
| | 0.3 | 1.01 | 1.16 | 1.74 | 2.35 | 2.93 | 4.17 | 5.2 | 6.85 |

Table: Median width of nominal $> 95\%$ Bonferroni CI for $\beta$.

| | | $\beta$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $\alpha_0$ | $\alpha_1$ | 0 | 0.25 | 0.5 | 0.75 | 1 | 1.5 | 2 | 3 |
| 0.0 | 0.0 | 19.07 | 3.44 | 1.86 | 1.32 | 0.87 | 0.47 | 0.37 | 0.35 |
| | 0.1 | 17.52 | 3.47 | 1.92 | 1.41 | 1 | 0.61 | 0.51 | 0.46 |
| | 0.2 | 17.41 | 3.51 | 1.9 | 1.45 | 1.1 | 0.76 | 0.65 | 0.58 |
| | 0.3 | 18.23 | 3.34 | 1.92 | 1.48 | 1.24 | 0.91 | 0.79 | 0.7 |
| 0.1 | 0.0 | 17.13 | 3.51 | 1.86 | 1.38 | 0.97 | 0.61 | 0.51 | 0.46 |
| | 0.1 | 17.88 | 3.33 | 1.85 | 1.45 | 1.13 | 0.78 | 0.67 | 0.6 |
| | 0.2 | 17.37 | 3.36 | 1.95 | 1.54 | 1.24 | 0.97 | 0.85 | 0.75 |
| | 0.3 | 18.07 | 3.33 | 1.98 | 1.63 | 1.41 | 1.17 | 1.04 | 0.92 |
| 0.2 | 0.0 | 17.79 | 3.39 | 1.92 | 1.45 | 1.11 | 0.75 | 0.65 | 0.58 |
| | 0.1 | 18.98 | 3.43 | 1.96 | 1.54 | 1.26 | 0.97 | 0.84 | 0.75 |
| | 0.2 | 18.25 | 3.26 | 1.92 | 1.64 | 1.45 | 1.2 | 1.06 | 0.95 |
| | 0.3 | 19.03 | 3.31 | 2.02 | 1.75 | 1.66 | 1.49 | 1.33 | 1.19 |
| 0.3 | 0.0 | 18.27 | 3.48 | 1.87 | 1.5 | 1.25 | 0.9 | 0.79 | 0.7 |
| | 0.1 | 19.4 | 3.41 | 1.96 | 1.63 | 1.43 | 1.18 | 1.04 | 0.92 |
| | 0.2 | 18.22 | 3.56 | 1.96 | 1.74 | 1.67 | 1.49 | 1.35 | 1.19 |
| | 0.3 | 17.56 | 3.55 | 2.13 | 1.96 | 1.86 | 1.86 | 1.74 | 1.55 |

Table: Median width of nominal 95% GMM CI, conditional on existence.

| | | $\beta$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $\alpha_0$ | $\alpha_1$ | 0 | 0.25 | 0.5 | 0.75 | 1 | 1.5 | 2 | 3 |
| 0.0 | 0.0 | 96 | 97 | 97 | 96 | 97 | 97 | 95 | 93 |
| | 0.1 | 97 | 99 | 99 | 99 | 99 | 98 | 96 | 95 |
| | 0.2 | 98 | 99 | 99 | 100 | 100 | 97 | 96 | 96 |
| | 0.3 | 97 | 100 | 100 | 100 | 99 | 96 | 96 | 96 |
| 0.1 | 0.0 | 97 | 99 | 99 | 99 | 100 | 98 | 97 | 95 |
| | 0.1 | 98 | 100 | 100 | 100 | 100 | 96 | 96 | 96 |
| | 0.2 | 98 | 100 | 100 | 100 | 99 | 96 | 96 | 95 |
| | 0.3 | 97 | 100 | 100 | 100 | 97 | 95 | 96 | 96 |
| 0.2 | 0.0 | 97 | 99 | 99 | 100 | 100 | 96 | 96 | 96 |
| | 0.1 | 98 | 100 | 100 | 100 | 99 | 96 | 96 | 96 |
| | 0.2 | 98 | 100 | 100 | 100 | 96 | 95 | 95 | 96 |
| | 0.3 | 98 | 100 | 100 | 98 | 95 | 95 | 95 | 96 |
| 0.3 | 0.0 | 97 | 99 | 100 | 100 | 100 | 95 | 96 | 97 |
| | 0.1 | 97 | 100 | 100 | 100 | 97 | 94 | 96 | 96 |
| | 0.2 | 98 | 100 | 100 | 98 | 94 | 94 | 96 | 96 |
| | 0.3 | 98 | 100 | 99 | 96 | 92 | 94 | 95 | 96 |

Table: Coverage of hybrid CI constructed from nominal 95% GMM and $>$ 95% Bonferroni intervals.

| $\alpha_0$ | $\alpha_1$ | $\beta$ 0 | 0.25 | 0.5 | 0.75 | 1 | 1.5 | 2 | 3 |
|---|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.0 | 0.4 | 0.41 | 0.43 | 0.43 | 0.43 | 0.42 | 0.4 | 0.35 |
| | 0.1 | 0.45 | 0.47 | 0.54 | 0.59 | 0.63 | 0.67 | 0.52 | 0.46 |
| | 0.2 | 0.51 | 0.54 | 0.65 | 0.76 | 0.84 | 0.82 | 0.65 | 0.58 |
| | 0.3 | 0.58 | 0.62 | 0.79 | 0.95 | 1.05 | 0.96 | 0.79 | 0.7 |
| 0.1 | 0.0 | 0.45 | 0.47 | 0.54 | 0.59 | 0.63 | 0.67 | 0.51 | 0.46 |
| | 0.1 | 0.51 | 0.54 | 0.66 | 0.77 | 0.86 | 0.92 | 0.69 | 0.61 |
| | 0.2 | 0.58 | 0.63 | 0.8 | 0.97 | 1.11 | 1.17 | 0.87 | 0.75 |
| | 0.3 | 0.67 | 0.75 | 1 | 1.25 | 1.4 | 1.4 | 1.06 | 0.92 |
| 0.2 | 0.0 | 0.51 | 0.54 | 0.65 | 0.76 | 0.85 | 0.83 | 0.65 | 0.58 |
| | 0.1 | 0.58 | 0.63 | 0.81 | 0.99 | 1.12 | 1.18 | 0.86 | 0.75 |
| | 0.2 | 0.67 | 0.75 | 1.01 | 1.29 | 1.48 | 1.56 | 1.08 | 0.95 |
| | 0.3 | 0.81 | 0.91 | 1.3 | 1.67 | 1.95 | 1.77 | 1.35 | 1.2 |
| 0.3 | 0.0 | 0.58 | 0.62 | 0.8 | 0.95 | 1.07 | 0.95 | 0.8 | 0.7 |
| | 0.1 | 0.68 | 0.74 | 1.01 | 1.26 | 1.43 | 1.48 | 1.06 | 0.93 |
| | 0.2 | 0.81 | 0.91 | 1.3 | 1.66 | 1.98 | 1.94 | 1.37 | 1.19 |
| | 0.3 | 1.01 | 1.16 | 1.73 | 2.24 | 2.71 | 2.33 | 1.78 | 1.55 |

Table: Median width of hybrid CI constructed from nominal 95% GMM and $>$ 95% Bonferroni intervals.

Figure: Coverage of hybrid vs. $> 95\%$ Bonferroni CIs: $\beta = 1$

Figure: Coverage of hybrid vs. $> 95\%$ Bonferroni CIs: $\beta = 2$

Figure: Coverage of hybrid vs. $> 95\%$ Bonferroni CIs: $\beta = 3$