Instrumented Common Confounding

A Novel Identification Strategy with Partially Endogenous Instruments

Christian Tien

EEA ESEM 2022

August 22, 2022

Causal Effect

$$J := \mathbb{E}\left[\int_{\mathcal{A}} Y(a)\pi(a)d\mu_A(a)\right]. \tag{1}$$

Causal Effect

$$J := \mathbb{E}\left[\int_{\mathcal{A}} Y(a)\pi(a)d\mu_A(a)\right]. \tag{1}$$

- Base measure of $\mathcal{A}$ ($A \in \mathcal{A}$): μ_A
- Contrast function $\pi : \mathcal{A} \to \mathbb{R}$ (ATE: $\pi(a) = 2a - 1$)

Causal Effect

$$J := \mathbb{E}\left[\int_{\mathcal{A}} Y(a)\pi(a)d\mu_A(a)\right]. \quad (1)$$

- Base measure of $\mathcal{A}$ ($A \in \mathcal{A}$): μ_A
- Contrast function $\pi : \mathcal{A} \to \mathbb{R}$ (ATE: $\pi(a) = 2a - 1$)

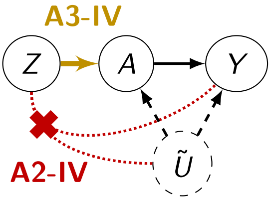
Assumption 1.1 (Outcome model linearly separable in error)

There exists some function $k_0 \in L_2(A, U)$ such that

$$Y = k_0(A, U) + \varepsilon_Y, \qquad \mathbb{E}\left[\varepsilon_Y | Z, U\right] = 0. \quad (2)$$

IV

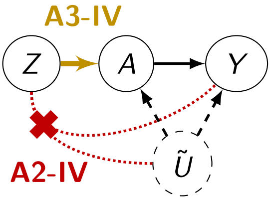
Figure: DAG of an IV model



Assumption 1.2 (IV Model)

IV

Figure: DAG of an IV model

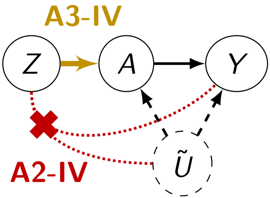


Assumption 1.2 (IV Model)

1. *Consistency:* $Y = Y(A, Z)$

IV

Figure: DAG of an IV model

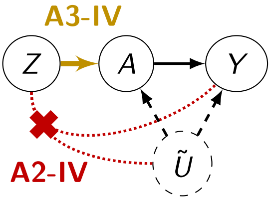


Assumption 1.2 (IV Model)

1. *Consistency:* $Y = Y(A, Z)$
2. *IV exogeneity:* $Y \perp\!\!\!\perp Z \mid A.$

IV

Figure: DAG of an IV model



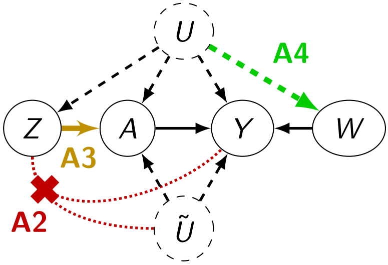
Assumption 1.2 (IV Model)

1. *Consistency:* $Y = Y(A, Z)$
2. *IV exogeneity:* $Y \perp\!\!\!\perp Z \mid A$.
3. *IV relevance:* $\mathbb{E}\left[g(A)|Z\right] = 0$ *only if* $g(A) = 0$.

ICC

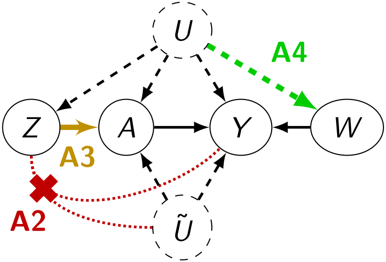
Assumption 1.3 (ICC Model)

Figure: DAG of an ICC model



ICC

Figure: DAG of an ICC model

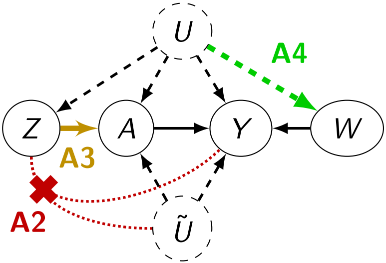


Assumption 1.3 (ICC Model)

1. *Consistency:* $Y = Y(A, Z)$, $W = W(A, Z)$

ICC

Figure: DAG of an ICC model

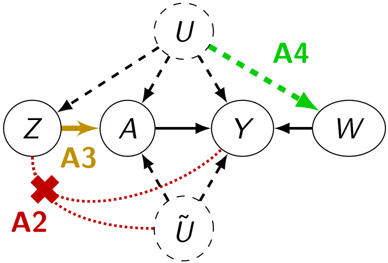


Assumption 1.3 (ICC Model)

1. *Consistency:* $Y = Y(A, Z)$, $W = W(A, Z)$
2. *Cond. IV exogeneity:* $(Y, W) \perp\!\!\!\perp Z \mid (A, U)$.

ICC

Figure: DAG of an ICC model

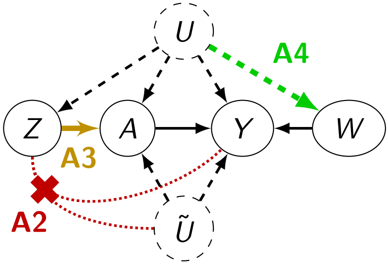


Assumption 1.3 (ICC Model)

1. *Consistency:* $Y = Y(A, Z)$, $W = W(A, Z)$
2. *Cond. IV exogeneity:* $(Y, W) \perp\!\!\!\perp Z \mid (A, U)$.
3. *IV relevance:* $\mathbb{E}\left[g(A, U)|Z\right] = 0$ *only if* $g(A, U) = 0$.

ICC

Figure: DAG of an ICC model



Assumption 1.3 (ICC Model)

1. *Consistency:* $Y = Y(A, Z)$, $W = W(A, Z)$
2. *Cond. IV exogeneity:* $(Y, W) \perp\!\!\!\perp Z \mid (A, U)$.
3. *IV relevance:* $\mathbb{E}\left[g(A, U)|Z\right] = 0$ *only if* $g(A, U) = 0$.
4. *NC outcomes:* $W \perp\!\!\!\perp (Z, A) \mid U$; *and* $\mathbb{E}\left[h_0(A, W)|A, U\right] = k_0(A, U)$ *for some* $h_0(A, W) \in L_2(A, W)$.

Related literature

- Nonparametric IV
  - Identification with outcome model restrictions [Chernozhukov et al., 2007, Horowitz and Lee, 2007, Newey and Powell, 2003]
  - Average structural function identification with control function [Blundell and Powell, 2004, Imbens and Newey, 2009]

Related literature

- Nonparametric IV
  - Identification with outcome model restrictions [Chernozhukov et al., 2007, Horowitz and Lee, 2007, Newey and Powell, 2003]
  - Average structural function identification with control function [Blundell and Powell, 2004, Imbens and Newey, 2009]
- Negative control/noisy conditioning variables
  - Proximal learning identification and robustness [Tchetgen Tchetgen et al., 2020, Cui et al., 2020]
  - Control function in negative control setting [Nagasawa, 2018]

Related literature

- Nonparametric IV
  - Identification with outcome model restrictions [Chernozhukov et al., 2007, Horowitz and Lee, 2007, Newey and Powell, 2003]
  - Average structural function identification with control function [Blundell and Powell, 2004, Imbens and Newey, 2009]
- Negative control/noisy conditioning variables
  - Proximal learning identification and robustness [Tchetgen Tchetgen et al., 2020, Cui et al., 2020]
  - Control function in negative control setting [Nagasawa, 2018]
- Mixture models
  - Identification with three independent measurements [Kruskal, 1977, Allman et al., 2009, Bonhomme et al., 2016b,a]

Related literature

- Nonparametric IV
  - Identification with outcome model restrictions [Chernozhukov et al., 2007, Horowitz and Lee, 2007, Newey and Powell, 2003]
  - Average structural function identification with control function [Blundell and Powell, 2004, Imbens and Newey, 2009]
- Negative control/noisy conditioning variables
  - Proximal learning identification and robustness [Tchetgen Tchetgen et al., 2020, Cui et al., 2020]
  - Control function in negative control setting [Nagasawa, 2018]
- Mixture models
  - Identification with three independent measurements [Kruskal, 1977, Allman et al., 2009, Bonhomme et al., 2016b,a]
- Panel data with fixed effects
  - Average effects with control function [Liu et al., 2021]

ICC - Linear Model

- Linear Model Setup

$$Y = A\beta + U\gamma_Y + W\upsilon_Y + \varepsilon_Y, \qquad \mathbb{E}\left[\varepsilon_Y | Z\right] = 0$$
$$A = Z\zeta + U\gamma_A + \varepsilon_A, \qquad \mathbb{E}\left[\varepsilon_A | Z\right] = 0$$
$$W = U\gamma_W + \varepsilon_W, \qquad \mathbb{E}\left[\varepsilon_W | U, Z\right] = 0$$
$$Z = U\gamma_Z + \varepsilon_Z, \qquad \mathbb{E}\left[\varepsilon_Z | U, W\right] = 0$$

ICC - Linear Model

- Linear Model Setup

$$
\begin{aligned}
Y &= A\beta + U\gamma_Y + W\upsilon_Y + \varepsilon_Y, & \mathbb{E}\left[\varepsilon_Y | Z\right] &= 0 \\
A &= Z\zeta + U\gamma_A + \varepsilon_A, & \mathbb{E}\left[\varepsilon_A | Z\right] &= 0 \\
W &= U\gamma_W + \varepsilon_W, & \mathbb{E}\left[\varepsilon_W | U, Z\right] &= 0 \\
Z &= U\gamma_Z + \varepsilon_Z, & \mathbb{E}\left[\varepsilon_Z | U, W\right] &= 0
\end{aligned}
$$

- Linear Model Moment

$$
\begin{aligned}
\mathbb{E}\left[W|Z\right] = \mathbb{E}\left[U|Z\right]\gamma_W \implies \mathbb{E}\left[U|Z\right] &= \mathbb{E}\left[W|Z\right]\gamma_W^{-1}, \\
\mathbb{E}\left[A|Z\right] &= Z\zeta + \mathbb{E}\left[U|Z\right]\gamma_A = Z\zeta + \mathbb{E}\left[W|Z\right]\gamma_W^{-1}\gamma_A \\
\mathbb{E}\left[Y|Z\right] &= \mathbb{E}\left[A|Z\right]\beta + \mathbb{E}\left[W|Z\right]\upsilon_Y + \mathbb{E}\left[U|Z\right]\gamma_Y \\
&= Z\zeta\beta + \mathbb{E}\left[W|Z\right]\left(\gamma_W^{-1}\gamma_A\beta + \upsilon_Y + \gamma_W^{-1}\gamma_Y\right)
\end{aligned}
$$

ICC - Linear Model

$$\mathbb{E}\left[Y|Z\right] = Z\zeta\beta + \mathbb{E}\left[W|Z\right]\left(\gamma_W^{-1}\gamma_A\beta + \upsilon_Y + \gamma_W^{-1}\gamma_Y\right)$$

$$\mathbb{E}\left[A|Z\right] = Z\zeta + \mathbb{E}\left[W|Z\right]\gamma_W^{-1}\gamma_A$$

$$\beta = \frac{\mathbb{E}\left[(Z\zeta)\,Y\middle|\,\mathbb{E}\left[W|Z\right]\right]}{\mathbb{E}\left[(Z\zeta)\,A\middle|\,\mathbb{E}\left[W|Z\right]\right]}$$

ICC - Linear Model

$$\mathbb{E}\left[Y|Z\right] = Z\zeta\beta + \mathbb{E}\left[W|Z\right]\left(\gamma_W^{-1}\gamma_A\beta + \upsilon_Y + \gamma_W^{-1}\gamma_Y\right)$$

$$\mathbb{E}\left[A|Z\right] = Z\zeta + \mathbb{E}\left[W|Z\right]\gamma_W^{-1}\gamma_A$$

$$\beta = \frac{\mathbb{E}\left[\left(Z\zeta\right)Y\middle|\mathbb{E}\left[W|Z\right]\right]}{\mathbb{E}\left[\left(Z\zeta\right)A\middle|\mathbb{E}\left[W|Z\right]\right]}$$

When is this feasible? Relevance Requirements:

ICC - Linear Model

$$\mathbb{E}\left[Y|Z\right] = Z\zeta\beta + \mathbb{E}\left[W|Z\right]\left(\gamma_W^{-1}\gamma_A\beta + \upsilon_Y + \gamma_W^{-1}\gamma_Y\right)$$

$$\mathbb{E}\left[A|Z\right] = Z\zeta + \mathbb{E}\left[W|Z\right]\gamma_W^{-1}\gamma_A$$

$$\beta = \frac{\mathbb{E}\left[(Z\zeta)\,Y \middle| \mathbb{E}\left[W|Z\right]\right]}{\mathbb{E}\left[(Z\zeta)\,A \middle| \mathbb{E}\left[W|Z\right]\right]}$$

When is this feasible? Relevance Requirements:

1. $\mathbb{E}\left[U|Z\right]$ and $\mathbb{E}\left[W|Z\right]$ are proportional. Keep $\mathbb{E}\left[U|Z\right]$ fixed by keeping $\mathbb{E}\left[W|Z\right]$ fixed. Requires $d_Z \geq d_U$, $d_W \geq d_U$.

ICC - Linear Model

$$\mathbb{E}\left[Y|Z\right] = Z\zeta\beta + \mathbb{E}\left[W|Z\right]\left(\gamma_W^{-1}\gamma_A\beta + \upsilon_Y + \gamma_W^{-1}\gamma_Y\right)$$

$$\mathbb{E}\left[A|Z\right] = Z\zeta + \mathbb{E}\left[W|Z\right]\gamma_W^{-1}\gamma_A$$

$$\beta = \frac{\mathbb{E}\left[(Z\zeta)\,Y\middle|\,\mathbb{E}\left[W|Z\right]\right]}{\mathbb{E}\left[(Z\zeta)\,A\middle|\,\mathbb{E}\left[W|Z\right]\right]}$$

When is this feasible? Relevance Requirements:

1. $\mathbb{E}\left[U|Z\right]$ and $\mathbb{E}\left[W|Z\right]$ are proportional. Keep $\mathbb{E}\left[U|Z\right]$ fixed by keeping $\mathbb{E}\left[W|Z\right]$ fixed. Requires $d_Z \geq d_U$, $d_W \geq d_U$.
2. Use remaining variation in predictions $Z\zeta$ to instrument for A while $\mathbb{E}\left[U|Z\right]$ fixed. Requires $(d_Z - d_U) \geq d_A$.

Identification of J if U observed

$$J := \mathbb{E}\left[\int_{\mathcal{A}} Y(a)\pi(a)d\mu_A(a)\right]$$

$$Y(A) = k_0(A, U) + \varepsilon_Y, \qquad \mathbb{E}\left[\varepsilon_Y | Z, U\right] = 0$$

Identification of J if U observed

$$J := \mathbb{E}\left[\int_{\mathcal{A}} Y(a)\pi(a)d\mu_A(a)\right]$$

$$Y(A) = k_0(A, U) + \varepsilon_Y, \qquad \mathbb{E}\left[\varepsilon_Y | Z, U\right] = 0$$

Lemma 4.0.1

If A1.1 (conditional outcome moment) holds, then

$$J = \mathbb{E}\left[\phi_{IV}(U; k_0)\right]$$

$$\textit{where} \quad \phi_{IV}(u; k_0) = \int_{\mathcal{A}} k_0(a, u)\pi(a)d\mu_A(a)$$

Bridge Function

Let $\mathbb{H}_0$ be the nonempty (A1.3.4) set of valid outcome bridge functions defined by

$$\mathbb{H}_0 = \left\{ h \in L_2(A, W) : \mathbb{E}\left[k_0(A, U) - h(A, W) | A, U\right] = 0 \right\} \neq \emptyset. \tag{3}$$

Bridge Function

Let $\mathbb{H}_0$ be the nonempty (A1.3.4) set of valid outcome bridge functions defined by

$$\mathbb{H}_0 = \left\{ h \in L_2(A, W) : \mathbb{E}\left[k_0(A, U) - h(A, W) | A, U \right] = 0 \right\} \neq \emptyset. \tag{3}$$

Lemma 4.0.2

Suppose A1.1 and A1.3.4 hold. For any $h_0 \in \mathbb{H}_0$,

$$J = \mathbb{E}\left[\tilde{\phi}_{IV}(w; h_0) \right]$$

$$\text{where} \quad \tilde{\phi}_{IV}(w; h_0) = \int_{\mathcal{A}} h_0(a, w)\pi(a) d\mu_A(a)$$

Proof of Lemma 4.0.2

Observable Bridge Function

Lemma 4.0.3

Suppose A1.1 and A1.3 hold. For any $h_0 \in \mathbb{H}_0$,

$$\mathbb{E}\left[Y - h_0(A, W)|Z\right] = 0 \tag{4}$$

Proof of Lemma 4.0.3

Observable Bridge Function

Lemma 4.0.3

Suppose A1.1 and A1.3 hold. For any $h_0 \in \mathbb{H}_0$*,*

$$\mathbb{E}\left[Y - h_0(A, W) | Z\right] = 0 \tag{4}$$

Proof of Lemma 4.0.3

Identify different set of observable bridge functions:

$$\mathbb{H}_0^{\text{obs}} = \{h \in L_2(A, W) : \mathbb{E}\left[Y - h(A, W) | Z\right] = 0\} \neq \emptyset. \tag{5}$$

Observable Bridge Function

Lemma 4.0.3

Suppose A1.1 and A1.3 hold. For any $h_0 \in \mathbb{H}_0$,

$$\mathbb{E}\left[Y - h_0(A, W) | Z\right] = 0 \tag{4}$$

Proof of Lemma 4.0.3

Identify different set of observable bridge functions:

$$\mathbb{H}_0^{\text{obs}} = \{h \in L_2(A, W) : \mathbb{E}\left[Y - h(A, W) | Z\right] = 0\} \neq \emptyset. \tag{5}$$

Lemma 4.0.4

Suppose A1.1 and A1.3 hold. Then,

$$\mathbb{H}_0 = \mathbb{H}_0^{obs}. \tag{6}$$

Proof of Lemma 4.0.4

Identification with Observable Bridge Function

Theorem 4.1

Suppose A1.1 and A1.3 hold. For any $h_0 \in \mathbb{H}_0^{obs}$,

$$J = \mathbb{E}\left[\tilde{\phi}_{IV}(w; h_0)\right]$$

$$\text{where} \quad \tilde{\phi}_{IV}(w; h_0) = \int_{\mathcal{A}} h_0(a, w)\pi(a)d\mu_A(a)$$

Proof of Theorem 8

NLS97 Data

- Y Household net worth at 35
- A BA degree

NLS97 Data

- Y Household net worth at 35
- A BA degree
- Z Pre-college GPA measures
- W Risky behaviour dummies

NLS97 Data

- Y Household net worth at 35
- A BA degree
- Z Pre-college GPA measures
- W Risky behaviour dummies
- U Ability
- $\tilde{U}$ Other biases (selection)

NLS97 Data

- Y Household net worth at 35
- A BA degree
- Z Pre-college GPA measures
- W Risky behaviour dummies
- U Ability
- $\tilde{U}$ Other biases (selection)
- X Covariates (sex, college GPA, parental education/net worth, siblings, region, etc)

Assumption 5.1 (Local Linear ICC Model)

Assumption 5.1 (Local Linear ICC Model)

1. *Model linearity:*

$$Y = \alpha_Y(X) + A\beta(X) + U\gamma_Y(X) + W\upsilon_Y(X) + \varepsilon_Y \tag{7}$$

$$A = \alpha_A(X) + Z\zeta(X) + U\gamma_A(X) + W\upsilon_A(X) + \varepsilon_A \tag{8}$$

$$W = \alpha_W(X) + U\gamma_W + \varepsilon_W, \qquad Z = \alpha_Z(X) + U\gamma_Z + \varepsilon_Z \tag{9}$$

Assumption 5.1 (Local Linear ICC Model)

1. *Model linearity:*

$$Y = \alpha_Y(X) + A\beta(X) + U\gamma_Y(X) + W\upsilon_Y(X) + \varepsilon_Y \tag{7}$$

$$A = \alpha_A(X) + Z\zeta(X) + U\gamma_A(X) + W\upsilon_A(X) + \varepsilon_A \tag{8}$$

$$W = \alpha_W(X) + U\gamma_W + \varepsilon_W, \qquad Z = \alpha_Z(X) + U\gamma_Z + \varepsilon_Z \tag{9}$$

2. *Cond. IV exogeneity:*

$$\mathbb{E}\left[\varepsilon_Y \middle| Z, U, X\right] = 0. \tag{10}$$

Assumption 5.1 (Local Linear ICC Model)

1. *Model linearity:*

$$Y = \alpha_Y(X) + A\beta(X) + U\gamma_Y(X) + W\upsilon_Y(X) + \varepsilon_Y \quad (7)$$

$$A = \alpha_A(X) + Z\zeta(X) + U\gamma_A(X) + W\upsilon_A(X) + \varepsilon_A \quad (8)$$

$$W = \alpha_W(X) + U\gamma_W + \varepsilon_W, \qquad Z = \alpha_Z(X) + U\gamma_Z + \varepsilon_Z \quad (9)$$

2. *Cond. IV exogeneity:*

$$\mathbb{E}\left[\varepsilon_Y \mid Z, U, X\right] = 0. \quad (10)$$

3. *IV relevance:*

$$rank(\zeta(x)) \geq d_A + d_U \text{ for any } x \in \mathcal{X}, \quad (11)$$

$$\mathbb{E}\left[\varepsilon_Z \mid W, U, X\right] = 0, \text{ and } rank(\gamma_Z) \geq d_U \quad (12)$$

Assumption 5.1 (Local Linear ICC Model)

1. *Model linearity:*

$$Y = \alpha_Y(X) + A\beta(X) + U\gamma_Y(X) + W\upsilon_Y(X) + \varepsilon_Y \tag{7}$$

$$A = \alpha_A(X) + Z\zeta(X) + U\gamma_A(X) + W\upsilon_A(X) + \varepsilon_A \tag{8}$$

$$W = \alpha_W(X) + U\gamma_W + \varepsilon_W, \qquad Z = \alpha_Z(X) + U\gamma_Z + \varepsilon_Z \tag{9}$$

2. *Cond. IV exogeneity:*

$$\mathbb{E}\left[\varepsilon_Y \middle| Z, U, X\right] = 0. \tag{10}$$

3. *IV relevance:*

$$rank(\zeta(x)) \geq d_A + d_U \text{ for any } x \in \mathcal{X}, \tag{11}$$

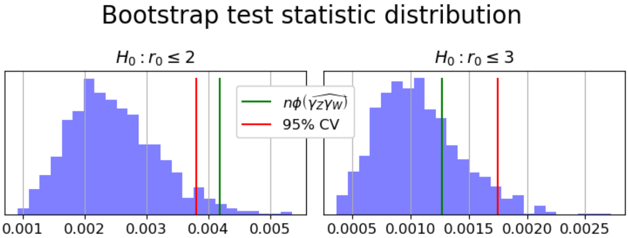
$$\mathbb{E}\left[\varepsilon_Z \middle| W, U, X\right] = 0, \text{ and } rank(\gamma_Z) \geq d_U \tag{12}$$

4. *NC outcomes:*

$$\mathbb{E}\left[\varepsilon_W \middle| Z, U, X\right] = 0, \text{ and } rank(\gamma_W) \geq d_U. \tag{13}$$

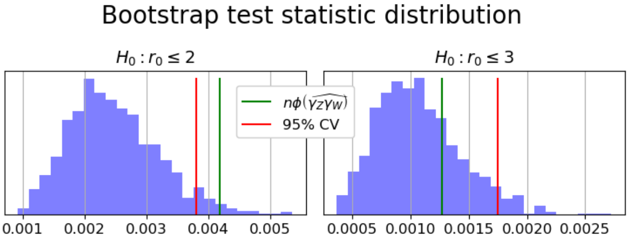
Test relevance of Z and W for U

- p-values for $H_0 : r_0 \leq r$: 2.1% ($r = 2$); 25.9% ($r = 3$) [Chen and Fang, 2019]
- 3-dim. U explains Z-W covariance.



Test relevance of Z and W for U

- p-values for $H_0 : r_0 \leq r$: 2.1% ($r = 2$); 25.9% ($r = 3$) [Chen and Fang, 2019]
- 3-dim. U explains Z-W covariance.



$$\underset{(n\times 3)}{\hat{U}_Z} := \mathbb{E}\left[U|Z\right] = \underset{(n\times d_Z)}{Z}\underset{(d_Z\times 3)}{\tilde{\gamma}^{\intercal}}, \qquad \underset{(n\times d_W)}{\mathbb{E}\left[W|Z\right]} = \underset{(n\times d_Z)}{Z}\underset{(d_Z\times 3)}{\tilde{\gamma}^{\intercal}}\underset{(3\times d_W)}{\gamma_W}$$

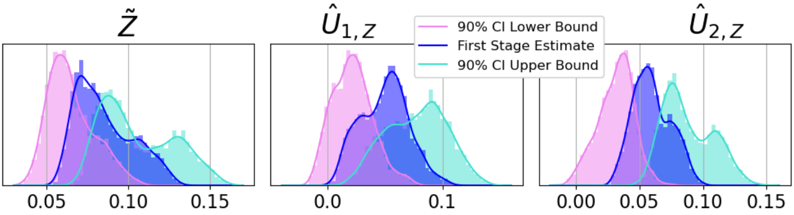
More on this Relevance Test | Understanding U

Test relevance of Z for A (given U)

1. Obtain residuals given covariates X: $v_A, v_{\tilde{Z}}, v_{\hat{U}_Z}$.
2. Regress v_A on $v_{\tilde{Z}}$ and $v_{\hat{U}_Z}$ in local linear regression.

$$v_A = v_{\tilde{Z}}\tilde{\zeta}(X) + v_{\hat{U}_Z}\tilde{\gamma}_A(X) + w_A \tag{14}$$

3. Test significance of $\tilde{\zeta}(X)$: All significant here.



More on this Relevance Test | Understanding U

Conditional Exogeneity of Z

- Cond. IV exogeneity:

$$\mathbb{E}\left[\varepsilon_Y \middle| Z, U, X\right] = 0.$$

- Exogeneity of Z given U cannot be tested.

Conditional Exogeneity of Z

- Cond. IV exogeneity:

$$\mathbb{E}\left[\varepsilon_Y \middle| Z, U, X\right] = 0.$$

- Exogeneity of Z given U cannot be tested.
- Pre-college GPA Z does not directly affect post-college earnings.
- $\hat{U}_Z$ accounts for ability (U) confounding.

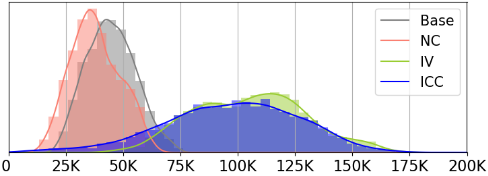
Conditional Exogeneity of Z

- Cond. IV exogeneity:

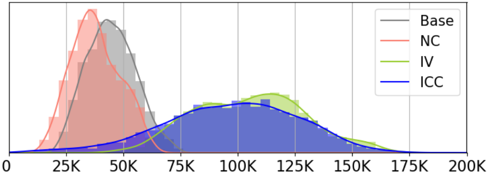
$$\mathbb{E}\left[\varepsilon_Y \middle| Z, U, X\right] = 0.$$

- Exogeneity of Z given U cannot be tested.
- Pre-college GPA Z does not directly affect post-college earnings.
- $\hat{U}_Z$ accounts for ability (U) confounding.
- Covariates:
  - College GPA
  - family net worth, net worth at 20
  - parental education, maternal age at first/subject's birth, # siblings
  - sex, citizenship status based on birth, residence region, urban residence (at 13-17 and 31-35)

Histograms of Estimates

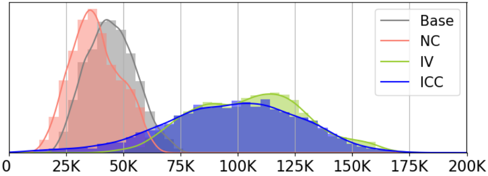


Histograms of Estimates



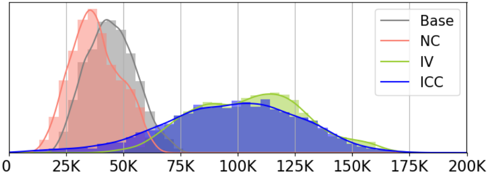
1. Baseline assumes exogeneity of A

Histograms of Estimates



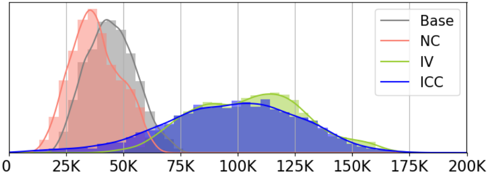
1. Baseline assumes exogeneity of A
2. NC estimates slightly below baseline
   - indicates small positive ability bias
   - omits selection bias

Histograms of Estimates



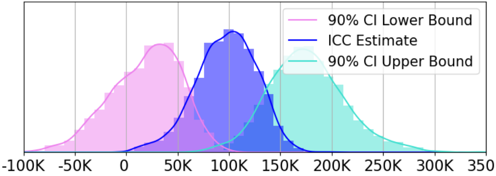
1. Baseline assumes exogeneity of A
2. NC estimates slightly below baseline
   - indicates small positive ability bias
   - omits selection bias
3. IV estimates much larger than baseline
   - indicates negative selection bias
   - omits ability bias

Histograms of Estimates



1. Baseline assumes exogeneity of A
2. NC estimates slightly below baseline
   - indicates small positive ability bias
   - omits selection bias
3. IV estimates much larger than baseline
   - indicates negative selection bias
   - omits ability bias
4. ICC estimates much larger than baseline but slightly below IV
   - indicates that selection bias was most relevant bias
   - ability bias has minor role

ICC Estimates with 90% Confidence Interval



- ICC estimates positive at significance level $\alpha = 5\%$ for 72% of individuals in sample
- ICC SEs about 54% wider than standard IV.
- ICC estimates larger for individuals with
  - ordinarily high college GPA,
  - less affluent/educated family background.

More on Example | CIs compared to IV

Other Use Cases

- Nonlinear (dynamic) panel data models
- Mismeasured observed confounders in IV
- Other microeconomic/medical causal inference problems

Conclusion

- Connects IV and negative control for a novel identification approach.
- Identifies causal effects with partially endogenous instruments.
- Economically and statistically significant results in the returns to education problem.

Next steps

- Generalise assumptions in nonparametric model to allow for $A \not\perp\!\!\!\perp W \mid U$ (using index sufficiency).
- Integrate returns to education application more with existing literature/estimates (MTE).
- Software
- Apply ICC to other causal questions.

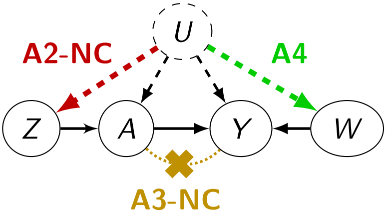
Thank You

arXiv:2206.12919
ct493@cam.ac.uk

Appendix

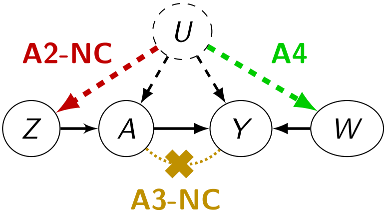
NC

Figure: DAG of a NC model



NC

Figure: DAG of a NC model



Assumption 7.1 (NC Model)

1. *Consistency:* $Y = Y(A, Z)$, $W = W(A, Z)$.
2. *NC action:* $(Y, W) \perp\!\!\!\perp Z \mid (A, U)$; *and* $\mathbb{E}\left[q_0(A, Z) | A, U\right] = \frac{1}{f(A|U)}$ *for some* $q_0(A, U) \in L_2(A, U)$.
3. *Latent unconfoundedness:* $(Y(a), W) \perp\!\!\!\perp (A, Z) \mid U, \ \forall a$
4. *NC outcomes:* $W \perp\!\!\!\perp (Z, A) \mid U$; *and* $\mathbb{E}\left[h_0(A, W) | A, U\right] = k_0(A, U)$ *for some* $h_0(A, W) \in L_2(A, W)$.

ICC construction algorithm

1. Separate all confounders into U and $\tilde{U}$, such that
$$Z \perp\!\!\!\perp \tilde{U} \mid U.$$
2. Include in U any unobserved variables necessary to justify
$$(A, Z) \perp\!\!\!\perp W \mid U.$$
3. Check completeness of W wrt U given A, i.e.
$$\mathbb{E}\left[g(A,U)|A,W\right] = 0 \text{ only when } g(A,U) = 0 \text{ for any } g \in L_2(A,U).$$
4. Check completeness of Z wrt (A, U), i.e.
$$\mathbb{E}\left[g(A,U)|Z\right] = 0 \text{ only when } g(A,U) = 0 \text{ for any } g \in L_2(A,U).$$

Proof of Lemma 4.0.1.

Let $T_A(x) := \int_{\mathcal{A}} x(a')\pi(a')\,\mathrm{d}\mu_A(a')$.

$$
\begin{aligned}
J &= \mathbb{E}\left[\int_{\mathcal{A}} Y(a')\pi(a')\,\mathrm{d}\mu_A(a')\right] = \mathbb{E}\left[\int_{\mathcal{A}} \left(k_0(a',U)+\varepsilon_Y\right)\pi(a')\,\mathrm{d}\mu_A(a')\right] \\
&= \mathbb{E}\left[T_A\left(k_0(a',U)+\mathbb{E}\left[\varepsilon_Y|A=a',U\right]\right)\right] = \mathbb{E}\left[T_A\left(k_0(a',U)+\mathbb{E}\left[\varepsilon_Y|U\right]\right)\right] \\
&= \mathbb{E}\left[T_A\left(k_0(a',U)+\mathbb{E}\left[\mathbb{E}\left[\varepsilon_Y|Z,U\right]|U\right]\right)\right] = \mathbb{E}\left[T_A\left(k_0(a',U)\right)\right] \\
&= \mathbb{E}\left[\left(\int_{\mathcal{A}} k_0(a',U)\pi(a')\,\mathrm{d}\mu_A(a')\right)\right] = \mathbb{E}\left[\phi_{IV}(U;k_0)\right]
\end{aligned}
$$

The second line follows as for any change of a in $Y(a)$, ε_Y is unchanged by definition. The third line follows as $\mathbb{E}\left[\varepsilon_Y|Z,U\right]=0$. □

Back to Lemma 4.0.1

Proof of lemma 4.0.2.

$$
\begin{aligned}
\mathbb{E}\left[\int_{\mathcal{A}} h_0(a, W)\pi(a)\,\mathrm{d}\mu_A(a)\right] &= \mathbb{E}\left[\mathbb{E}\left[\int_{\mathcal{A}} h_0(a, W)\pi(a)\,\mathrm{d}\mu_A(a) | U\right]\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\int_{\mathcal{A}} h_0(a, W)\pi(a)\,\mathrm{d}\mu_A(a) | A = a, U\right]\right] \\
&= \mathbb{E}\left[\int_{\mathcal{A}} \mathbb{E}\left[h_0(a, W) | A = a, U\right]\pi(a)\,\mathrm{d}\mu_A(a)\right] \\
&= \mathbb{E}\left[\int_{\mathcal{A}} k_0(a, U)\pi(a)\,\mathrm{d}\mu_A(a)\right] \\
&= \mathbb{E}\left[\phi_{IV}(U; k_0)\right] = J
\end{aligned}
$$

From line one to two we used $W \perp\!\!\!\perp A \mid U$ (A1.3.4). From line three to four we used the definition of $\mathbb{H}_0$, where $h_0 \in \mathbb{H}_0$. On the last line we used lemma 4.0.1. □

Back to Lemma 4.0.2

Proof of lemma 4.0.3.

Any $h_0 \in \mathbb{H}_0$ satisfies

$$\mathbb{E}\left[k_0(A,U) - h_0(A,W)|A,U,Z\right] = \mathbb{E}\left[k_0(A,U) - h_0(A,W)|A,U\right] = 0.$$

The first equality holds by $(W,Y) \perp\!\!\!\perp Z|A,U$. Consequently,

$$\begin{aligned}
&\mathbb{E}\left[Y - h_0(A,W)|Z\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[k_0(A,U) - h_0(A,W)|U,Z\right]|Z\right] + \mathbb{E}\left[\mathbb{E}\left[\varepsilon|U,Z\right]|Z\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[k_0(A,U) - h_0(A,W)|U,Z\right]|Z\right] \\
&= \mathbb{E}\left[k_0(A,U) - h_0(A,W)|Z\right] = \mathbb{E}\left[\mathbb{E}\left[k_0(A,U) - h_0(A,W)|A,U,Z\right]|Z\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[k_0(A,U) - h_0(A,W)|A,U\right]|Z\right] = \mathbb{E}\left[0|Z\right] = 0
\end{aligned}$$

This proves that equation 4 of lemma 4.0.3 holds. □

Back to Lemma 4.0.3

Proof of lemma 4.0.4.

For any $h_0 \in \mathbb{H}_0^{\text{obs}}$,

$$\mathbb{E}\left[Y - h_0(A, W)|Z\right] = \mathbb{E}\left[k_0(A, U) - h_0(A, W)|Z\right]$$
$$= \mathbb{E}\left[\mathbb{E}\left[k_0(A, U) - h_0(A, W)|A, U\right]|Z\right] = 0.$$

Under completeness A1.3.3, the above can only be true if $\mathbb{E}\left[k_0(A, U) - h_0(A, W)|A, U\right] = 0$. Hence, any $h_0 \in \mathbb{H}_0^{\text{obs}}$ also satisfies $h_0 \in \mathbb{H}_0$, which implies $\mathbb{H}_0^{\text{obs}} \subseteq \mathbb{H}_0$. From lemma 4.0.3 it is known that $\mathbb{H}_0 \subseteq \mathbb{H}_0^{\text{obs}}$. Consequently, $\mathbb{H}_0^{\text{obs}} = \mathbb{H}_0$. □

Back to Lemma 4.0.4

Proof of theorem 4.1.

Identification of $\tilde{\phi}_{IV}$. For any $h_0 \in \mathbb{H}_0^{\text{obs}} = \mathbb{H}_0$,

$$
\begin{aligned}
\mathbb{E}\left[\tilde{\phi}_{IV}(W; h_0)\right] &= \mathbb{E}\left[\int_{\mathcal{A}} h_0(a, W)\pi(a)\, \mathrm{d}\mu_A(a)\right] \\
&= \mathbb{E}\left[\int_{\mathcal{A}} \mathbb{E}\left[h(a, W)|A = a, U\right] \pi(a)\, \mathrm{d}\mu_A(a)\right] \\
&= \mathbb{E}\left[\int_{\mathcal{A}} k_0(a, U)\pi(a)\, \mathrm{d}\mu_A(a)\right] \\
&= \mathbb{E}\left[\phi_{IV}(U; k_0)\right] = J
\end{aligned}
$$

We move from the first to the second equation by assumption $W \perp\!\!\!\perp (A, Z) \mid U$ of 1.3. The step from the second to the third line is by lemma 4.0.4 and 1.3.4. The last line holds by lemma 4.0.1. □

Back to Theorem 4.1

Option 2: Strict Monotonicity in First Stage

Assumption 9.1 (Strict Monotonicity)

$$A = h(Z, m(U, \eta)) \tag{15}$$

1. *The reduced form $h(Z, m)$ is strictly monotonic in m with probability 1,*
2. *$m(U, t)$ is strictly monotonic in t with probability 1,*
3. *and η is a continuously distributed scalar with a CDF that is strictly increasing on the support of η (conditional on U).*
4. *η and Z are independent (conditional on U).*

With Option 2: Enhanced Negative Control Information

Assumption 9.2 (Enhanced Negative Control Information)

1. *Reduced form NC instrument:* $A \perp\!\!\!\perp W_0 | (Z, U)$.
   There exists a bridge function $\kappa_0 \in L_2(Z, W_0)$ *such that*

$$\mathbb{E}\left[\kappa_0(Z, W_0) | Z, W_1, U\right] = \frac{f(U)}{f(U|Z)}. \tag{16}$$

2. *Reduced form NC action:* $W_1 \perp\!\!\!\perp Z | U$.
   There exists a bridge function $\tau_{A,0} \in L_2(Z, W)$ *for all* $A \in \mathcal{A}$ *such that*

$$\mathbb{E}\left[\tau_{A,0}(Z, W_1) | Z, W_0, U\right] = F(A|Z, U). \tag{17}$$

Control Bridge

$$V := F_{A|Z,U}(A, Z, U) = F_\eta(\eta)$$

The control bridge $\tilde{V}$ contains information about first stage disturbance η.

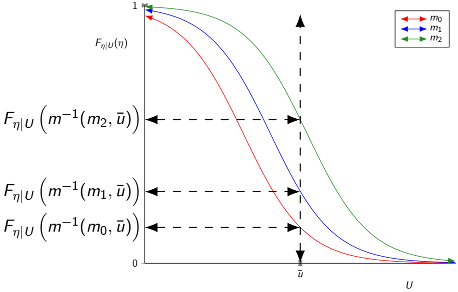
$$\begin{aligned}
\tilde{V} &:= \int_{\mathcal{U}} F_{A|Z,U}(A, Z, u)\, \mathrm{d}F(u) \\
&= \int_{\mathcal{U}} \Pr\left(h(Z, m(u, \eta)) \leq A\right) \mathrm{d}F(u) \\
&= \int_{\mathcal{U}} \Pr\left(\eta \leq m^{-1}\left(h^{-1}(A, Z), u\right)\right) \mathrm{d}F(u) \\
&= \int_{\mathcal{U}} F_{\eta|U}\left(m^{-1}\left(h^{-1}(A, Z), u\right)\right) \mathrm{d}F(u) \\
&= \int_{\mathcal{U}} F_{\eta|U}\left(\eta(u)\right) \mathrm{d}F(u)
\end{aligned}$$

Information Equivalence with Control Bridge

Lemma 9.0.1

Suppose 9.1 holds. Then,

$$\sigma(\tilde{V}, U) = \sigma(V, U) = \sigma(\eta, U).$$



First Stage Bridge Functions

First stage brige functions:

$$\mathbb{T}_0 = \left\{ \tau_A \in L_2(Z, W_1) : \mathbb{E}\left[\tau_A(Z, W_1) | Z, U, W_0\right] = F(A|Z, U) \right\} \neq \emptyset,$$

$$\mathbb{K}_0 = \left\{ \kappa \in L_2(Z, W_0) : \mathbb{E}\left[\kappa(Z, W_0) | Z, U, W_1\right] = \frac{f(U)}{f(U|Z)} \right\} \neq \emptyset.$$

Lemma 9.0.2

Under A1.3 (ICC), A9.1 and A9.2, any $\tau_{A,0} \in \mathbb{T}_0$ *and* $\kappa_0 \in \mathbb{K}_0$ *satisfy that*

$$\mathbb{E}\left[\tau_{A,0}(Z, W_1) | Z, W_0\right] = F(A|Z, W_0). \tag{18}$$

$$\mathbb{E}\left[\kappa_0(Z, W_0) | Z, W_1\right] = \frac{f(W_1)}{f(W_1|Z)}. \tag{19}$$

Observed 1st Stage Bridge Functions

Lemma 9.0.3

Under A1.3, A9.1 and A9.2, any $\tau_{A,0} \in \mathbb{T}_0$ *and* $\kappa_0 \in \mathbb{K}_0$ *satisfy that*

$$\mathbb{E}\left[\tau_{A,0}(Z, W_1)|Z, W_0\right] = F(A|Z, W_0). \tag{20}$$

$$\mathbb{E}\left[\kappa_0(Z, W_0)|Z, W_1\right] = \frac{f(W_1)}{f(W_1|Z)}. \tag{21}$$

Sets of observable bridge functions:

$$\mathbb{E}\left[\tau_{A,0}(Z, W_1)|Z, W_0\right] = F(A|Z, W_0). \tag{22}$$

$$\mathbb{E}\left[\kappa_0(Z, W_0)|Z, W_1\right] = \frac{f(W_1)}{f(W_1|Z)}. \tag{23}$$

1st Stage Identification with Observed Bridge Functions

Lemma 9.0.4

Suppose A1.3, A9.1 and A9.2 hold.
It follows that for any $\tau_A \in L_2(Z, W_1)$ *and* $\kappa_0 \in \mathbb{K}_0^{obs}$,

$$\int_{\mathcal{W}_1} \tau_A(Z, W_1)\,\mathrm{d}F(W_1) - \tilde{V} = \mathbb{E}\left[\kappa_0(Z, W_0)\,\mathbb{E}\left[\tau_A(Z, W_1) - F(A|Z, W_0)|Z, W_0\right]|Z\right].$$

Hence, for any $\tau_A \in \mathbb{T}_0^{obs}$ *as long as* $\mathbb{K}_0^{obs} \neq \emptyset$,

$$\tilde{V} = \int_{\mathcal{W}_1} \tau_A(Z, W_1)\,\mathrm{d}F(W_1).$$

Main step of both proofs for IPW and REG estimators, which requires $\mathbb{E}\left[\int_{\mathcal{A}} \mathbb{E}\left[Y|A=a,\tilde{V},U\right]\pi(a)\,\mathrm{d}\mu_A(a)\right] = \mathbb{E}\left[\int_{\mathcal{A}} \mathbb{E}\left[Y|A=a,V,U\right]\pi(a)\,\mathrm{d}\mu(a)\right]$.

$$
\begin{aligned}
&\mathbb{E}\left[\int_{\mathcal{A}} \mathbb{E}\left[Y|A=a,\tilde{V},U\right]\pi(a)\,\mathrm{d}\mu_A(a)\right] \\
&= \mathbb{E}\left[\int_{\mathcal{A}}\int_{\mathcal{V}} \mathbb{E}\left[Y|A=a,v,\tilde{V},U\right] f(v|A=a,\tilde{V},U)\,\mathrm{d}\mu_V(v)\pi(a)\,\mathrm{d}\mu_A(a)\right] \\
&= \mathbb{E}\left[\int_{\mathcal{A}}\int_{\mathcal{V}} \mathbb{E}\left[Y|A=a,v,U\right] f(v|\tilde{V},U)\,\mathrm{d}\mu_V(v)\pi(a)\,\mathrm{d}\mu_A(a)\right] \\
&= \mathbb{E}\left[\int_{\mathcal{V}}\int_{\mathcal{A}} \mathbb{E}\left[Y|A=a,v,U\right]\pi(a)\,\mathrm{d}\mu_A(a)\left(\int_{\tilde{\mathcal{V}}} f(v|\tilde{v},U)f(\tilde{v}|U)\,\mathrm{d}\mu_A(\tilde{v})\right)\mathrm{d}\mu_V(v)\right] \\
&= \mathbb{E}\left[\int_{\mathcal{V}}\int_{\mathcal{A}} \mathbb{E}\left[Y|A=a,v,U\right]\pi(a)\,\mathrm{d}\mu_A(a)f(v|U)\,\mathrm{d}\mu_V(v)\right] \\
&= \mathbb{E}\left[\int_{\mathcal{A}} \mathbb{E}\left[Y|A=a,V,U\right]\pi(a)\,\mathrm{d}\mu_A(a)\right]
\end{aligned}
$$

The third equality follows from Lemma 9.0.1.

Proof of equation 22. For any $\tau_{A,0} \in \mathbb{T}_0$,

$$\mathbb{E}\left[\tau_{A,0}(Z, W_1)|Z, U, W_0\right] = F(A|Z, U)$$
$$\mathbb{E}\left[\mathbb{E}\left[\tau_{A,0}(Z, W_1)|Z, U, W_0\right]|Z, W_0\right] = \mathbb{E}\left[F(A|Z, U)|Z, W_0\right]$$
$$\mathbb{E}\left[\tau_{A,0}(Z, W_1)|Z, W_0\right] = F(A|Z, W_0).$$

We move from the second to third equation by assumption $A \perp\!\!\!\perp W_0 \mid (Z, U)$ of 9.2.

Proof of equation 23. For any $\kappa_0 \in \mathbb{K}_0$,

$$\mathbb{E}\left[\kappa_0(Z, W_0)|Z, U, W_1\right] = \frac{f(U)}{f(U|Z)}$$

$$\mathbb{E}\left[\mathbb{E}\left[\kappa_0(Z, W_0)|Z, U, W_1\right]|Z, W_1\right] = \mathbb{E}\left[\frac{f(U)}{f(U|Z)}|Z, W_1\right]$$

$$\mathbb{E}\left[\mathbb{E}\left[\kappa_0(Z, W_0)|Z, U, W_1\right]|Z, W_1\right] = f(Z)\,\mathbb{E}\left[\frac{1}{f(Z|U)}|Z, W_1\right]$$

$$\begin{aligned}\mathbb{E}\left[\kappa_0(Z, W_0)|Z, W_1\right] &= f(Z)\int_{\mathcal{U}} \frac{f(U|W_1, Z)}{f(Z|U)}\,\mathrm{d}\mu_U(U)\\ &= f(Z)\int_{\mathcal{U}} \frac{f(W_1, Z|U)f(U)}{f(Z|U)f(W_1, Z)}\,\mathrm{d}\mu_U(U)\\ &= f(Z)\int_{\mathcal{U}} \frac{f(W_1|U)f(U)}{f(W_1, Z)}\,\mathrm{d}\mu_U(U)\\ &= \frac{f(Z)}{f(Z|W_1)} = \frac{f(W_1)}{f(W_1|Z)}\end{aligned}$$

We move from the fifth to the sixth equation by assumption $W_1 \perp\!\!\!\perp Z \mid U$ of 9.2.

Proof of Lemma 9.0.4 I

First, note that for any $\tau_{A,0} \in \mathbb{T}_0$,

$$F(A|Z,U) = \mathbb{E}\left[\tau_{A,0}(Z,W_1)|Z,U,W_0\right] = \mathbb{E}\left[\tau_{A,0}(Z,W_1)|Z,U\right] = \int_{\mathcal{W}_1} \tau_{A,0}(Z,W_1)\,\mathrm{d}F(W_1|U)$$

$$\tilde{V} = \int_{\mathcal{U}} F(A|Z,U)\,\mathrm{d}F(U) = \int_{\mathcal{W}_1} \tau_{A,0}(Z,W_1)\,\mathrm{d}F(W_1)$$

We move from the second to the third line by assumption $Z \perp\!\!\!\perp W_1 \mid U$ of 9.2.

Proof of Lemma 9.0.4 II

Now, note that for any $\tau_A \in L_2(Z, W_1)$ and $\kappa_0 \in \mathbb{K}_0^{\text{obs}}$,

$$\begin{aligned}
\mathbb{E}\left[\kappa_0(Z, W_0)\tau_A(Z, W_1)|Z\right] &= \mathbb{E}\left[\mathbb{E}\left[\kappa_0(Z, W_0)\tau_A(Z, W_1)|Z, W_1\right]|Z\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\kappa_0(Z, W_0)|Z, W_1\right]\tau_A(Z, W_1)|Z\right] \\
&= \mathbb{E}\left[\frac{f(W_1)}{f(W_1|Z)}\tau_A(Z, W_1)|Z\right] \\
&= \int_{\mathcal{W}_1} \frac{f(W_1)}{f(W_1|Z)}\tau_A(Z, W_1)\,\mathrm{d}F(W_1|Z) \\
&= \int_{\mathcal{W}_1} \tau_A(Z, W_1)\,\mathrm{d}F(W_1)
\end{aligned}$$

Proof of Lemma 9.0.4 III

For any $\tau_A \in L_2(Z, W_1)$ and $\kappa_0 \in \mathbb{K}_0^{\text{obs}}$, we write

$$\int_{\mathcal{W}_1} \tau_A(Z, W_1)\, \mathrm{d}F(W_1) = \mathbb{E}\left[\kappa_0(Z, W_0)\tau_A(Z, W_1)|Z\right] = \mathbb{E}\left[\kappa_0(Z, W_0)\,\mathbb{E}\left[\tau_A(Z, W_1)|Z, W_0\right]|Z\right].$$

For any $\tau_{A,0} \in \mathbb{T}_0$ and $\kappa_0 \in \mathbb{K}_0^{\text{obs}}$, we write

$$\begin{aligned}\tilde{V} &= \int_{\mathcal{W}_1} \tau_{A,0}(Z, W_1)\, \mathrm{d}F(W_1) \\ &= \mathbb{E}\left[\kappa_0(Z, W_0)\tau_{A,0}(Z, W_1)|Z\right] \\ &= \mathbb{E}\left[\kappa_0(Z, W_0)\,\mathbb{E}\left[\tau_{A,0}(Z, W_1)|Z, W_0\right]|Z\right] \\ &= \mathbb{E}\left[\kappa_0(Z, W_0)F(A|Z, W_0)|Z\right].\end{aligned}$$

Proof of Lemma 9.0.4 IV

Now this implies that for any $\tau_A \in L_2(Z, W_1)$,

$$\int_{\mathcal{W}_1} \tau_A(Z, W_1)\, \mathrm{d}F(W_1) - \tilde{V} = \mathbb{E}\left[\kappa_0(Z, W_0)\, \mathbb{E}\left[\tau_A(Z, W_1) - F(A|Z, W_0)|Z, W_0\right.\right.$$

Hence, for any $\tau_A \in \mathbb{T}_0^{\text{obs}}$ as long as $\mathbb{K}_0^{\text{obs}} \neq \emptyset$,

$$\tilde{V} = \int_{\mathcal{W}_1} \tau_A(Z, W_1)\, \mathrm{d}F(W_1).$$

Test relevance of Z and W for U

- Let $r_0 := \text{rank}\left(\tilde{\gamma}_Z^{\intercal} \Sigma_U \gamma_W\right)$ and $m_{ZW} = \min\{d_Z, d_W\}$.
- Then $r_0 = \min\{d_U, d_Z, d_W\} \leq m_{ZW}$.
- If $r_0 < m_{ZW}$, then $d_U = r_0 < \min m_{ZW}$.
- Test

$$H_0 : r_0 \leq r \text{ vs } H_1 : r_0 > r \tag{24}$$

  for some $r < m_{ZW}$.
- If H_0 not rejected, this suggets $r_0 \leq r$.
- Bootstrap based test [Chen and Fang, 2019]

Back to Relevance Test of Z and W for U | Back to ICC CIs

Test relevance of Z and W for U

- Let $c_{WZ} := \tilde{\gamma}_Z^\intercal \Sigma_U \gamma_W$.
- Singular value decomposition:

$$c_{WZ} = \underset{d_Z \times d_Z}{P_0} \underset{d_Z \times d_W}{\Pi_0} \underset{d_W \times d_W}{Q_0^\intercal} \tag{25}$$

- Let $\phi_r(A) := \sum_{j=r+1}^{m_A} \pi_j^2(A)$
  - $\pi_j(A)$ is the j-th singular value of A.
  - $m_A := \min\left\{d_{A,row}, d_{A,col}\right\}$
- Equivalent test to 24:

$$H_0 : \phi_r\left(c_{WZ}\right) = 0 \text{ vs } H_1 : \phi_r\left(c_{WZ}\right) > 0 \tag{26}$$

  - $\pi_j(A)$ is the j-th singular value of A.
  - $m_A := \min\left\{d_{A,row}, d_{A,col}\right\}$.

Back to Relevance Test of Z and W for U | Back to ICC CIs

Test relevance of Z and W for U

- Test statistic distribution:

$$n\phi_r\left(\hat{c}_{WZ}\right) \xrightarrow{d} \sum_{j=r-r_0+1}^{m_{ZW}-r_0} \pi_j^2\left(P_{0,2}^{\intercal}\mathcal{M}Q_{0,2}\right), \tag{27}$$

  - $P_{0,2}$ corresponds to last $m_{ZW}-r$ columns of P_0.
  - $Q_{0,2}$ corresponds to last $m_{ZW}-r$ columns of Q_0.
  - $\mathcal{M}$ is asymptotic distribution of $\hat{c}_{WZ}$: $\sqrt{n}\left(\hat{c}_{WZ}-c_{WZ}\right)\xrightarrow{d}\mathcal{M}$.
- Bootstrap: est. $\mathcal{M}$ as distribution $\hat{\mathcal{M}}_n^* = \sqrt{n}\{\hat{c}_{WZ,n}^* - \hat{c}_{WZ,n}\}$
  - $\hat{c}_{WZ,n}^*$ is distr. of bootstrap sample OLS estimators $\hat{c}_{WZ,n,b}$.
  - $b \in \mathcal{B}$ is one bootstrap sample.

Back to Relevance Test of Z and W for U | Back to ICC CIs

Understanding U

- $\hat{U}_Z$ explains unobserved characteristics that correlate pre-college GPA Z and risky behaviour W.

$\hat{U}_{0,Z}$ Academic - general ability
- less likely to engage in risky behaviour (4-10%)

$\hat{U}_{1,Z}$ General ability:
- less likely to engage in risky behaviour (2-8%)

$\hat{U}_{2,Z}$ Math/social sciences - English ability, & general ability:
- ambiguous effect on risky behaviour ((-2)-2%)

Back to Relevance Test of Z and W for U | Back to Relevance Test of Z for A (given U)
Estimates of $P_{0,1}\Pi_{0,1}$ | Back to ICC CIs

Construct $\widehat{\mathbb{E}\left[U|Z\right]}$

- Let

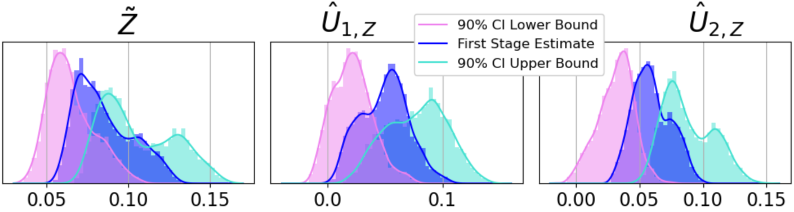
$$\hat{U}_Z := \widehat{\mathbb{E}\left[U|Z\right]} = \underset{n\times d_Z}{Z} \underset{d_Z\times r}{P_{0,1}} \underset{r\times r}{\Pi_{0,1}} \tag{28}$$

  - $P_{0,1}$ corresponds to first $r = 3$ columns of P_0.
  - $\Pi_{0,1}$ corresponds to first $r = 3$ rows and columns of Π_0.
  - U is normalised in scale as in SVD
- Want to understand and interpret $\hat{U}_Z$ to later support the theoretical argument about IV conditional exogeneity in assumption 5.1.2 10

Back to Relevance Test of Z and W for U | Back to Relevance Test of Z for A (given U)

Back to ICC CIs

Test relevance of Z for A (given U) II



- One SD increase in instrument $\tilde{Z}$ increases probability of obtaining BA degree by roughly 6-12%.
- Lower bound of 90% CI has minimum 5%.
- Reject $H_0 : \tilde{\zeta}(x) = 0$ for any $x \in \mathcal{X}$.
- Assumption 5.1.2 equation 11 (conditional relevance of Z for A given U) holds.
- Common confounders make obtaining BA degree more likely.

Back to Relevance Test of Z for A (given U)

Estimating $\mathbb{E}\left[U|Z\right]$

- $\mathbb{E}\left[U|Z\right] = ZP_{0,1}\Pi_{0,1}$
- Normalisation
  - $\bar{Z}_j = 0$, $\sigma(Z_j) = 1$, for all $j \in \{0, 1, \ldots, 7\}$
  - $\bar{\hat{U}}_{k,Z} = 0$, $\sigma(\hat{U}_{k,Z}) = 1$, for all $k \in \{0, 1, 2\}$
- Estimated $\widehat{P_{0,1}\Pi_{0,1}}^{\intercal}$

	GPA						ASVAB
	school	transcript	English	Math	SocSci	LifeSci	percentile
$\hat{U}_{0,Z}$	0.05	1.53	-0.10	-0.14	-0.09	-0.15	-0.46
$\hat{U}_{1,Z}$	0.33	1.84	-0.30	-0.56	-0.24	-0.50	0.34
$\hat{U}_{2,Z}$	-0.42	0.31	-0.87	0.29	0.31	-0.13	0.79

Back to Understanding U | Back to ICC CIs

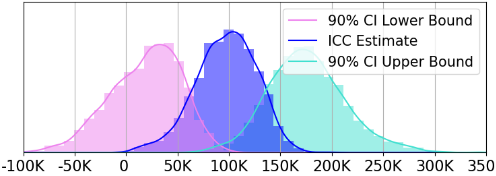
Estimating $Q_{0,1}$

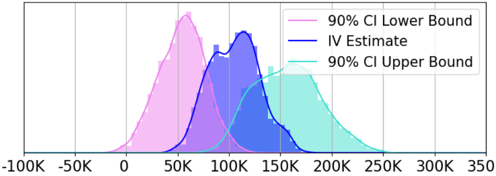
- $\mathbb{E}\left[W|Z\right] = ZP_{0,1}\Pi_{0,1}Q_{0,1} = \mathbb{E}\left[U|Z\right]Q_{0,1}$
- $W \in \{0,1\}$
- Estimated $Q_{0,1}$

	drink	smoke	try marijuana	run away	attack someone	sell drugs	destroy property	steal $< 50\$$	steal $> 50\$$
$\hat{U}_{0,Z}$	-0.06	-0.10	-0.11	-0.06	-0.10	-0.07	-0.08	-0.08	-0.05
$\hat{U}_{1,Z}$	-0.02	-0.06	-0.07	-0.06	-0.09	-0.05	-0.06	-0.04	-0.05
$\hat{U}_{2,Z}$	0.01	-0.01	-0.01	-0.03	-0.04	0.00	0.02	0.02	-0.01

Back to Understanding U | Back to ICC CIs

ICC vs IV Estimates with 90% Confidence Interval





- ICC confidence intervals on average 54% wider than IV

Back to ICC CIs

Elizabeth S Allman, Catherine Matias, John A Rhodes, et al. Identifiability of parameters in latent structure models with many observed variables. *The Annals of Statistics*, 37(6A):3099–3132, 2009.

Richard W Blundell and James L Powell. Endogeneity in semiparametric binary response models. *The Review of Economic Studies*, 71(3):655–679, 2004.

Stéphane Bonhomme, Koen Jochmans, and Jean-Marc Robin. Non-parametric estimation of finite mixtures from repeated measurements. *Journal of the Royal Statistical Society: Series B: Statistical Methodology*, pages 211–229, 2016a.

Stéphane Bonhomme, Koen Jochmans, Jean-Marc Robin, et al. Estimating multivariate latent-structure models. *The Annals of Statistics*, 44(2): 540–563, 2016b.

Qihui Chen and Zheng Fang. Improved inference on the rank of a matrix. *Quantitative Economics*, 10(4):1787–1824, 2019.

Victor Chernozhukov, Guido W Imbens, and Whitney K Newey. Instrumental variable estimation of nonseparable models. *Journal of Econometrics*, 139 (1):4–14, 2007.

Yifan Cui, Hongming Pu, Xu Shi, Wang Miao, and Eric Tchetgen Tchetgen. Semiparametric proximal causal inference. *arXiv preprint arXiv:2011.08411*, 2020.

Joel L Horowitz and Sokbae Lee. Nonparametric instrumental variables estimation of a quantile regression model. *Econometrica*, 75(4):1191–1208, 2007.

Guido W Imbens and Whitney K Newey. Identification and estimation of triangular simultaneous equations models without additivity. *Econometrica*, 77(5):1481–1512, 2009.

Joseph B Kruskal. Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear algebra and its applications*, 18(2):95–138, 1977.

Laura Liu, Alexandre Poirier, and Ji-Liang Shiu. Identification and estimation of average partial effects in semiparametric binary response panel models. *arXiv preprint arXiv:2105.12891*, 2021.

Kenichi Nagasawa. Treatment effect estimation with noisy conditioning variables. *arXiv preprint arXiv:1811.00667*, 2018.

Whitney K Newey and James L Powell. Instrumental variable estimation of nonparametric models. *Econometrica*, 71(5):1565–1578, 2003.

Eric J Tchetgen Tchetgen, Andrew Ying, Yifan Cui, Xu Shi, and Wang Miao. An introduction to proximal causal learning. *arXiv preprint arXiv:2009.10982*, 2020.