Negative Control Falsification Tests for Instrumental Variable Designs

Oren Danieli, Daniel Nevo, Itai Walk, Bar Weinstein, Dan Zeltzer

NBER SI 2024 Labor Studies
July, 2024

What Is a Negative Control Falsification Test?

IV assumptions can never be tested directly

Falsification (AKA placebo) tests indirectly test the design validity

- 51% of highly cited papers with IV since 2013 use some falsification test

What Is a Negative Control Falsification Test?

IV assumptions can never be tested directly

Falsification (AKA placebo) tests indirectly test the design validity

- 51% of highly cited papers with IV since 2013 use some falsification test

Most falsification tests fall into two categories. Borrowing terminology from other disciplines we call these

1. Negative Control Outcome (**NCO**)
2. Negative Control Instrument (**NCI**)

Category 1 - Negative Control Outcome - NCO

IV is not associated with variables (NCOs) it should not be associated with.

- Implemented in 72% of papers with falsification test.

Category 1 - Negative Control Outcome - NCO

IV is not associated with variables (NCOs) it should not be associated with.

- Implemented in 72% of papers with falsification test.

Martin & Yurukoglu (2017, AER)

- IV: Channel position $\rightarrow$ X: Fox News Viewership $\rightarrow$ Y: Republican vote 2008
- Falsification: Channel position $\not\rightarrow$ Republican vote pre-Fox (1996)

Category 2 - Negative Control Instrument - NCI

Outcome is not associated with variables (NCIs) it should not be associated with.

- Implemented in 25% of papers with falsification test.

Category 2 - Negative Control Instrument - NCI

Outcome is not associated with variables (NCIs) it should not be associated with.

- Implemented in 25% of papers with falsification test.

Nunn & Qian (2014, AER)

- IV: US wheat production $\rightarrow$ X: Food aid supply $\rightarrow$ Y: Conflict in recipient country
- Falsification: Other corps production $\not\rightarrow$ Conflict in recipient country

Category 2 - Negative Control Instrument - NCI

Outcome is not associated with variables (NCIs) it should not be associated with.

- Implemented in 25% of papers with falsification test.

Nunn & Qian (2014, AER)

- IV: US wheat production $\rightarrow$ X: Food aid supply $\rightarrow$ Y: Conflict in recipient country
- Falsification: Other corps production $\nrightarrow$ Conflict in recipient country

Ashraf & Galor (2013, AER)

- IV: Dist. from Addis Ababa $\rightarrow$ X: Genetic diversity $\rightarrow$ Y: Economic development
- Falsification: Distance from other cities $\nrightarrow$ Economic development

Category 2 - Negative Control Instrument - NCI

Outcome is not associated with variables (NCIs) it should not be associated with.

- Implemented in 25% of papers with falsification test.

Nunn & Qian (2014, AER)

- IV: US wheat production $\rightarrow$ X: Food aid supply $\rightarrow$ Y: Conflict in recipient country
- Falsification: Other corps production $\nrightarrow$ Conflict in recipient country

Ashraf & Galor (2013, AER)

- IV: Dist. from Addis Ababa $\rightarrow$ X: Genetic diversity $\rightarrow$ Y: Economic development
- Falsification: Distance from other cities $\nrightarrow$ Economic development

Madestam et. al (2013, QJE)

- IV: Rain on 4/15/2009 $\rightarrow$ X: Tea party protests $\rightarrow$ Y: Republican vote in 2010
- Falsification: Rain on other dates $\nrightarrow$ Republican vote in 2010

This Paper

We develop a theory for negative control tests.

- Model negative controls as proxies for unobserved threats.

1. Correct mistakes: some implementations find "problems" in exogenous IVs.

This Paper

We develop a theory for negative control tests.

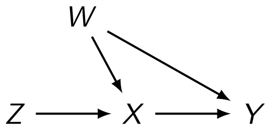
- Model negative controls as proxies for unobserved threats.

1. Correct mistakes: some implementations find "problems" in exogenous IVs.
2. Propose ways to extend the use of negative controls.

Negative Control Outcomes (NCO)

Notation

Assume there exists an IV design (Z, X, Y) such that:

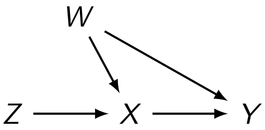


IV (Z) is affecting an endogenous variable (X) which affects the outcome (Y)

- W is a confounder that motivates the usage of an IV

Notation

Assume there exists an IV design (Z, X, Y) such that:

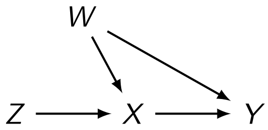


IV (Z) is affecting an endogenous variable (X) which affects the outcome (Y)

- W is a confounder that motivates the usage of an IV
- Use DAGs only for intuition, all proofs are with potential outcomes

Notation

Assume there exists an IV design (Z, X, Y) such that:

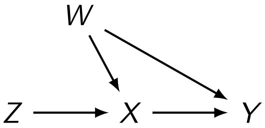


IV (Z) is affecting an endogenous variable (X) which affects the outcome (Y)

- W is a confounder that motivates the usage of an IV
- Use DAGs only for intuition, all proofs are with potential outcomes
- When outcome independence $(Z \perp\!\!\!\perp Y(z,x))$ and exclusion $(Y(z,x) = Y(z',x))$ hold, the IV is "exogenous":

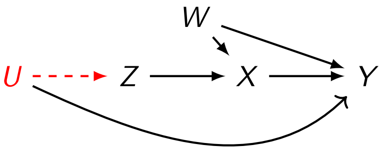
$$Z \perp\!\!\!\perp Y(x)$$

Negative Control Outcome Example - Martin and Yurukoglu (2017)



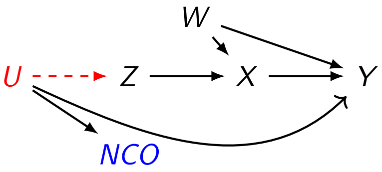
- Impact of Fox News (X) on 2008 Republican vote (Y)
- IV is Fox News channel position (Z)

Negative Control Outcome Example - Martin and Yurukoglu (2017)



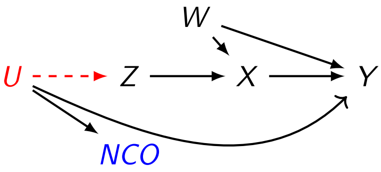
- Impact of Fox News (X) on 2008 Republican vote (Y)
- IV is Fox News channel position (Z)
- **Alternative Path Outcome Variable**: Unobserved conservativeness (U)

Negative Control Outcome Example - Martin and Yurukoglu (2017)



- Impact of Fox News (X) on 2008 Republican vote (Y)
- IV is Fox News channel position (Z)
- **Alternative Path Outcome Variable**: Unobserved conservativeness (U)
- **Negative Control Outcome**: 1996 Republican vote (NCO)

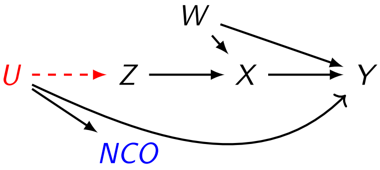
Negative Control Outcome Example - Martin and Yurukoglu (2017)



- Impact of Fox News (X) on 2008 Republican vote (Y)
- IV is Fox News channel position (Z)
- **Alternative Path Outcome Variable**: Unobserved conservativeness (U)
- **Negative Control Outcome**: 1996 Republican vote (NCO)
- Key idea: $Z \not\perp\!\!\!\perp NCO$ implies that the dashed line exists and IV not exogenous

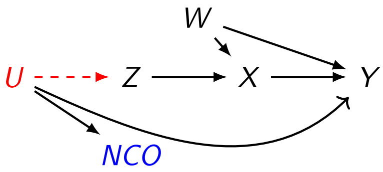
Negative Control Outcome Assumption

- In paper - Define APO variable - the threat to the identification Defnition
- APO U is often unobserved. Instead, we use an observed NCO.



Negative Control Outcome Assumption

- In paper - Define APO variable - the threat to the identification Defnition
- APO U is often unobserved. Instead, we use an observed NCO.



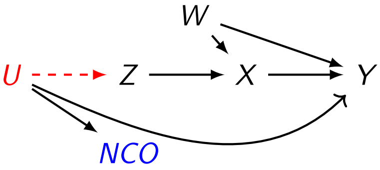
Definition

A random variable NCO satisfies the **negative control outcome assumption** if there exists an APO variable U such that

$$Z \perp\!\!\!\perp NCO | U$$

Negative Control Outcome Assumption

- In paper - Define APO variable - the threat to the identification Definition
- APO U is often unobserved. Instead, we use an observed NCO.



Definition

A random variable NCO satisfies the **negative control outcome assumption** if there exists an APO variable U such that

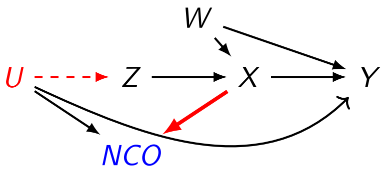
$$Z \perp\!\!\!\perp NCO | U$$

Example of violation: Guidetti et al. (2021)

- Is non-respiratory hospitalization an NCO when X is air pollution?

Negative Control Outcome Assumption

- In paper - Define APO variable - the threat to the identification Defnition
- APO U is often unobserved. Instead, we use an observed NCO.



Definition

A random variable NCO satisfies the **negative control outcome assumption** if there exists an APO variable U such that

$$Z \perp\!\!\!\perp NCO | U$$

Example of violation: Guidetti et al. (2021)

- Is non-respiratory hospitalization an NCO when X is air pollution?
- No. Air pollution causes hospital congestion and affects non-respiratory patients so

$$Z \not\perp\!\!\!\perp NCO | U$$

Negative Control Outcome - Theorem

Theorem

Assume that a random variable NCO satisfies the NCO assumption. If

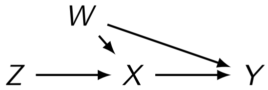
$$Z \not\perp\!\!\!\perp NCO$$

then IV design is not exogenous.

Proof

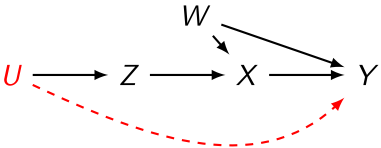
Negative Control Instruments (NCI)

Negative Control Instrument Example - Nunn & Qian 2014



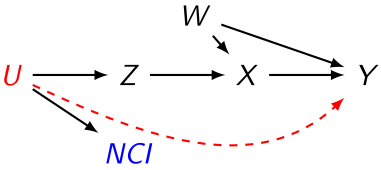
- Impact of U.S. food aid (X) on conflicts in recipient country (Y)
- U.S. wheat production (Z) affects aid amount

Negative Control Instrument Example - Nunn & Qian 2014



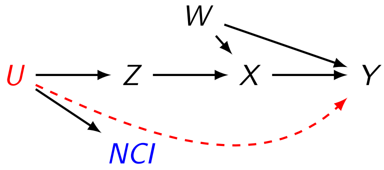
- Impact of U.S. food aid (X) on conflicts in recipient country (Y)
- U.S. wheat production (Z) affects aid amount
- **Alternative Path Instrument**: Global weather conditions (U)

Negative Control Instrument Example - Nunn & Qian 2014



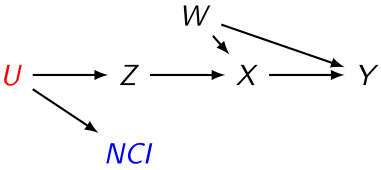
- Impact of U.S. food aid (X) on conflicts in recipient country (Y)
- U.S. wheat production (Z) affects aid amount
- **Alternative Path Instrument**: Global weather conditions (U)
- **Negative Control Instrument**: U.S. oranges production (NCI)

Negative Control Instrument Example - Nunn & Qian 2014



- Impact of U.S. food aid (X) on conflicts in recipient country (Y)
- U.S. wheat production (Z) affects aid amount
- **Alternative Path Instrument**: Global weather conditions (U)
- **Negative Control Instrument**: U.S. oranges production (NCI)
- $NCI \not\perp\!\!\!\perp Y | Z$ implies that the dashed line exists and the design is invalid
  - Note: $NCI \not\perp\!\!\!\perp Y$ always

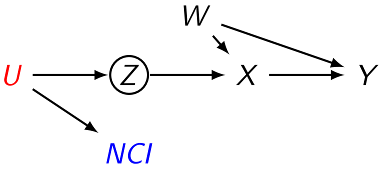
Controlling for the IV



Problem: NCI (orange production) is correlated with Z (wheat production)

- $NCI \not\perp\!\!\!\perp Y$ when IV is exogenous
  - Orange production correlated with wheat production which affects conflicts

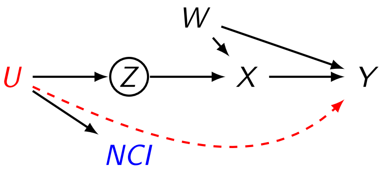
Controlling for the IV



Problem: NCI (orange production) is correlated with Z (wheat production)

- $NCI \not\perp\!\!\!\perp Y$ when IV is exogenous
  - Orange production correlated with wheat production which affects conflicts
- $NCI \perp\!\!\!\perp Y | Z$ when IV is exogenous

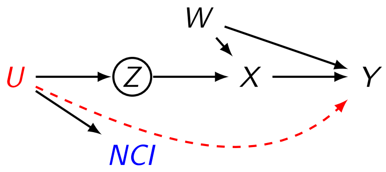
Controlling for the IV



Problem: *NCI* (orange production) is correlated with *Z* (wheat production)

- $NCI \not\perp\!\!\!\perp Y$ when IV is exogenous
  - Orange production correlated with wheat production which affects conflicts
- $NCI \perp\!\!\!\perp Y | Z$ when IV is exogenous
- $NCI \not\perp\!\!\!\perp Y | Z$ implies IV is not exogenous

Controlling for the IV



Problem: *NCI* (orange production) is correlated with *Z* (wheat production)

- $NCI \not\perp\!\!\!\perp Y$ when IV is exogenous
    - Orange production correlated with wheat production which affects conflicts
- $NCI \perp\!\!\!\perp Y | Z$ when IV is exogenous
- $NCI \not\perp\!\!\!\perp Y | Z$ implies IV is not exogenous

Failing to control for the IV in an NCI test can find false problems in valid IV designs.

- 81% of papers do not do this

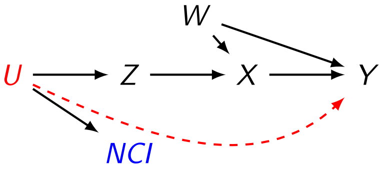
Negative Control Instrument Assumption

Definition

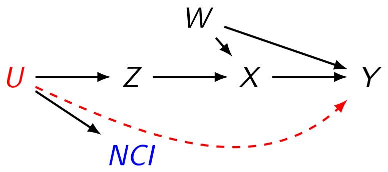
A random variable *NCI* satisfies the **negative control instrument assumption** if there exists an API variable U such that:

$$Y \perp\!\!\!\perp NCI | U, Z$$

API Definition



Negative Control Instrument Assumption



Definition

A random variable *NCI* satisfies the **negative control instrument assumption** if there exists an API variable U such that:

$$Y \perp\!\!\!\perp NCI | U, Z$$

Theorem (Negative-Control Instrument Test)

Assume that the random variable NCI satisfies the NCI assumption. If $Y \not\perp\!\!\!\perp NCI | Z$*, then the IV design is not exogenous.*

When control for IV is unnecessary

Functional Form

Including Control Variables

In most cases, the IV is only exogenous conditional on some controls ($Z \perp\!\!\!\perp Y(x) | C$).

- Ex: Judge allocation is only random within districts

Including Control Variables

In most cases, the IV is only exogenous conditional on some controls ($Z \perp\!\!\!\perp Y(x) | C$).

- Ex: Judge allocation is only random within districts

- Researchers typically assume some functional form for the reduced form

$$Y = \beta_0 Z + \gamma_0 C + \epsilon_0$$

Including Control Variables

In most cases, the IV is only exogenous conditional on some controls ($Z \perp\!\!\!\perp Y(x)|C$).

- Ex: Judge allocation is only random within districts
- Researchers typically assume some functional form for the reduced form

$$Y = \beta_0 Z + \gamma_0 C + \epsilon_0$$

- NCO tests typically implement "pseudo-outcome exercise" with similar structure

$$NCO = \color{red}{\beta_1} Z + \gamma_1 C + \epsilon_1$$

Including Control Variables

In most cases, the IV is only exogenous conditional on some controls ($Z \perp\!\!\!\perp Y(x)|C$).

- Ex: Judge allocation is only random within districts
- Researchers typically assume some functional form for the reduced form

$$Y = \beta_0 Z + \gamma_0 C + \epsilon_0$$

- NCO tests typically implement "pseudo-outcome exercise" with similar structure

$$NCO = \beta_1 Z + \gamma_1 C + \epsilon_1$$

- And proper NCI tests implement

$$Y = \beta_2 NCI + \gamma_2 C + \alpha_2 Z + \epsilon_2$$

Including Control Variables

In most cases, the IV is only exogenous conditional on some controls ($Z \perp\!\!\!\perp Y(x)|C$).

- Ex: Judge allocation is only random within districts
- Researchers typically assume some functional form for the reduced form

$$Y = \beta_0 Z + \gamma_0 C + \epsilon_0$$

- NCO tests typically implement "pseudo-outcome exercise" with similar structure

$$NCO = \beta_1 Z + \gamma_1 C + \epsilon_1$$

- And proper NCI tests implement

$$Y = \beta_2 NCI + \gamma_2 C + \alpha_2 Z + \epsilon_2$$

All these tests depend on functional form assumptions.

- Could find $\beta_1, \beta_2 \neq 0$ even when IV is exogenous ($Z \perp\!\!\!\perp Y(x)|C$)

Implications - NCO

Blandhol et al., (2022) define *rich covariates*

$$E[Z|C] = \gamma C$$

- Necessary assumption for 2SLS
- Blandhol et al. suggest solutions for this problem
- Typically easier problem than non-exogenous IV

Most NCO tests used in practice also test rich covariates

- Pro - necessary assumption worth checking
  - Reduces noise
- Con - want to separate functional form and exogeneity problems

Implications - NCI

Define *correctly specified reduced form*

$$E[Y|Z, C] = \beta Z + \delta C$$

- Not a necessary assumption
- LATE interpretation still valid without it

Most NCI tests used in practice test rely on this assumption.

- Pro - reduces noise
- Con - can reject null when design valid

Implications - NCI

Define *correctly specified reduced form*

$$E[Y|Z, C] = \beta Z + \delta C$$

- Not a necessary assumption
- LATE interpretation still valid without it

Most NCI tests used in practice test rely on this assumption.

- Pro - reduces noise
- Con - can reject null when design valid

Possible solution: If NC tests reject the null, test functional form and exogeneity separately

Practical Considerations

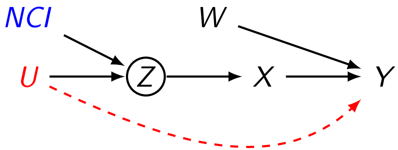
Practical Considerations

1. Detect all NC in the data

Details

Practical Considerations

1. Detect all NC in the data
   Ex: Anything that causes the IV is an NCI (since the IV is a collider)



- IV is quarter of birth (Angrist and Krueger, 1991), NCI is quarter of marriage

Details

Practical Considerations

1. Detect all NC in the data
2. Choose (conditional) independence test

Details

Practical Considerations

1. Detect all NC in the data
2. Choose (conditional) independence test
3. Post-mortem analysis

Details

Bias Correction

Beyond Bias Detection

In non-IV-settings, negative controls are also used to correct biases

- Simple example: Diff-in-Diff when lagged outcome is an NCO (Sofer et al., 2016)
- Proximal learning (Tchetgen Tchetgen et al., 2020; Shi et al., 2020)

Beyond Bias Detection

In non-IV-settings, negative controls are also used to correct biases

- Simple example: Diff-in-Diff when lagged outcome is an NCO (Sofer et al., 2016)
- Proximal learning (Tchetgen Tchetgen et al., 2020; Shi et al., 2020)

Can you do the same with negative controls for IV?

Beyond Bias Detection

In non-IV-settings, negative controls are also used to correct biases

- Simple example: Diff-in-Diff when lagged outcome is an NCO (Sofer et al., 2016)
- Proximal learning (Tchetgen Tchetgen et al., 2020; Shi et al., 2020)

Can you do the same with negative controls for IV?

Yes. We show this in a simple IV setting

- Yet requires stronger assumptions

Setting

Let (Z_1, X_1, Y_1) be an IV design

- Assume binary IV (Z_1) and treatment (X_1)

Setting

Let (Z_1, X_1, Y_1) be an IV design

- Assume binary IV (Z_1) and treatment (X_1)

Assume lagged outcome Y_0 is an observed NCO.

- Assume in the lagged period the IV and treatment do not exist ($Z_0 = X_0 = 0$).
  - In paper - more scenarios

Setting

Let (Z_1, X_1, Y_1) be an IV design

- Assume binary IV (Z_1) and treatment (X_1)

Assume lagged outcome Y_0 is an observed NCO.

- Assume in the lagged period the IV and treatment do not exist $(Z_0 = X_0 = 0)$.
    - In paper - more scenarios

Assume the NCO test failed $(Z_1 \not\perp\!\!\!\perp Y_0)$, indicating a violation of outcome independence.

Setting

Let (Z_1, X_1, Y_1) be an IV design

- Assume binary IV (Z_1) and treatment (X_1)

Assume lagged outcome Y_0 is an observed NCO.

- Assume in the lagged period the IV and treatment do not exist $(Z_0 = X_0 = 0)$.
  - In paper - more scenarios

Assume the NCO test failed $(Z_1 \not\perp\!\!\!\perp Y_0)$, indicating a violation of outcome independence.

Define a Diff-in-Wald Estimator

$$DiW = \frac{E\left[Y_1 - Y_0 | Z_1 = 1\right] - E\left[Y_1 - Y_0 | Z_1 = 0\right]}{E\left[X_1 | Z_1 = 1\right] - E\left[X_1 | Z_1 = 0\right]}$$

Bias Correction Assumptions

1. Remaining IV assumptions
   - Exclusion
   - Treatment independence
   - Relevance

Bias Correction Assumptions

1. Remaining IV assumptions
   - Exclusion
   - Treatment independence
   - Relevance

2. DiD assumptions
   - No anticipation (typically satisfied from NCO assumption)
   - Parallel trend: $E[Y_1(0,0) - Y_0(0,0)|Z_1 = 1] = E[Y_1(0,0) - Y_0(0,0)|Z_1 = 0]$

Bias Correction Assumptions

1. Remaining IV assumptions
   - Exclusion
   - Treatment independence
   - Relevance

2. DiD assumptions
   - No anticipation (typically satisfied from NCO assumption)
   - Parallel trend: $E[Y_1(0,0) - Y_0(0,0)|Z_1 = 1] = E[Y_1(0,0) - Y_0(0,0)|Z_1 = 0]$

3. Treatment effect homogeneity: for every observation $Y_1(1) - Y_1(0) = \tau$

Bias Correction Theorem

Theorem (No-treatment + No-IV)

Under the above assumptions,

$$DiW = \tau$$

Bias Correction Theorem

Theorem (No-treatment + No-IV)

Under the above assumptions,

$$DiW = \tau$$

With substantial treatment effect heterogeneity, bias can be arbitrarily large

Bias Correction Theorem

Theorem (No-treatment + No-IV)

Under the above assumptions,

$$DiW = \tau$$

With substantial treatment effect heterogeneity, bias can be arbitrarily large

Can drop some assumptions under different (less likely) assumptions on Z_0, X_0 Details

Summary

1. This paper develops a theory for a large share of IV falsification tests

Summary

1. This paper develops a theory for a large share of IV falsification tests

2. Correct mistakes
   - NCI tests typically require controlling for the IV
   - NC tests find (fixable/unimportant) functional form problems

Summary

1. This paper develops a theory for a large share of IV falsification tests

2. Correct mistakes
   - NCI tests typically require controlling for the IV
   - NC tests find (fixable/unimportant) functional form problems

3. Extend use of negative controls
   - Use NCOs for bias correction
   - Additional types of NCIs
   - Additional types of statistical tests

Theory Appendix

Alternative Path Outcome (APO) Variable

◂ Back

Simple case with only one potential threat to IV validity

In paper: general definition for the case of multiple threats Full Definition

Alternative Path Outcome (APO) Variable ◂ Back

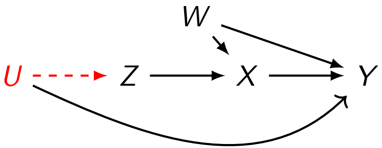
Simple case with only one potential threat to IV validity
In paper: general definition for the case of multiple threats Full Definition

Definition

A random variable U is an APO if the following conditions hold:

1. **Latent IV Validity**: $Z \perp\!\!\!\perp Y(x) | U$

2. **Path Indication** $Z \perp\!\!\!\perp Y(x) \rightarrow Z \perp\!\!\!\perp U$



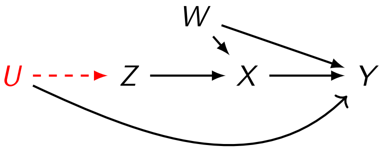
Alternative Path Outcome (APO) Variable ◂ Back

Simple case with only one potential threat to IV validity
In paper: general definition for the case of multiple threats Full Definition

Definition

A random variable U is an APO if the following conditions hold:

1. **Latent IV Validity**: $Z \perp\!\!\!\perp Y(x) | U$
   - IV is valid controlling for U
   - Implies U is the threat itself
2. **Path Indication** $Z \perp\!\!\!\perp Y(x) \rightarrow Z \perp\!\!\!\perp U$



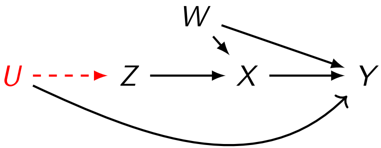
Alternative Path Outcome (APO) Variable Back

Simple case with only one potential threat to IV validity
In paper: general definition for the case of multiple threats Full Definition

Definition

A random variable U is an APO if the following conditions hold:

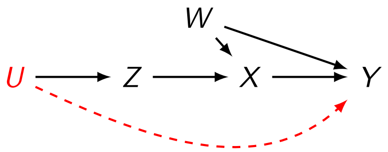
1. **Latent IV Validity**: $Z \perp\!\!\!\perp Y(x) | U$
   - IV is valid controlling for U
   - Implies U is the threat itself
2. **Path Indication** $Z \perp\!\!\!\perp Y(x) \rightarrow Z \perp\!\!\!\perp U$
   - When the dashed part of the path exist ($Z \not\perp\!\!\!\perp U$), the rest of the path exists ($U \not\perp\!\!\!\perp Y(x)$)
   - Loosely means that U is related to $Y(x)$



Alternative Path Instrument (API) Variable ◂ Back

A random variable U is an API if:

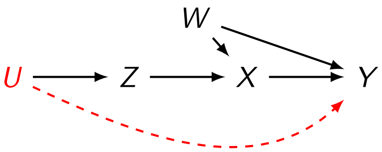
1. *Latent IV Validity*: $Z \perp\!\!\!\perp Y(x) | U$
2. *Path Indication*: $Z \perp\!\!\!\perp Y(x) \implies U \perp\!\!\!\perp Y(x)$
3. *No Via-Treatment Link*:
   $U \perp\!\!\!\perp Y(x) | Z \implies U \perp\!\!\!\perp Y | Z$



Alternative Path Instrument (API) Variable ◂ Back

A random variable U is an API if:

1. *Latent IV Validity*: $Z \perp\!\!\!\perp Y(x) | U$
2. *Path Indication*: $Z \perp\!\!\!\perp Y(x) \implies U \perp\!\!\!\perp Y(x)$
3. *No Via-Treatment Link*:
   $U \perp\!\!\!\perp Y(x) | Z \implies U \perp\!\!\!\perp Y | Z$



Similarities and differences from APO Variables:

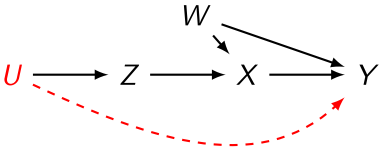
- Latent IV validity is the same as in APO

Alternative Path Instrument (API) Variable

Back

A random variable U is an API if:

1. *Latent IV Validity*: $Z \perp\!\!\!\perp Y(x) | U$
2. *Path Indication*: $Z \perp\!\!\!\perp Y(x) \implies U \perp\!\!\!\perp Y(x)$
3. *No Via-Treatment Link*:
   $U \perp\!\!\!\perp Y(x) | Z \implies U \perp\!\!\!\perp Y | Z$



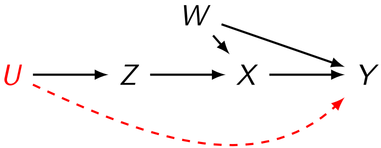
Similarities and differences from APO Variables:

- Latent IV validity is the same as in APO
- Path indication for API variables requires orthogonality to $Y(x)$ not Z

Alternative Path Instrument (API) Variable ◂ Back

A random variable U is an API if:

1. *Latent IV Validity*: $Z \perp\!\!\!\perp Y(x) | U$
2. *Path Indication*: $Z \perp\!\!\!\perp Y(x) \implies U \perp\!\!\!\perp Y(x)$
3. *No Via-Treatment Link*:
   $U \perp\!\!\!\perp Y(x) | Z \implies U \perp\!\!\!\perp Y | Z$



Similarities and differences from APO Variables:

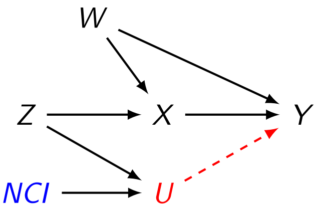
- Latent IV validity is the same as in APO
- Path indication for API variables requires orthogonality to $Y(x)$ not Z
- No via-treatment: link between U and Y is not through the treatment
  - Typically satisfied if

$$U \perp\!\!\!\perp X | Z$$

When Control for the IV is Unnecessary

Hypothetical example:

- Date of birth cutoff (Z) $\rightarrow$ Participation in schooling program A (X) $\rightarrow$ Wages (Y)
- Same cutoff used for schooling program B with no participation data (U)
- NCI: Program B availability by school



Can test whether $NCI \perp\!\!\!\perp Y$ without control to learn about the dashed line.

IV and NCI Independence

Back

Theorem

Assume that the random variable NCI satisfies the NCI assumption. If in addition

$$NCI \perp\!\!\!\perp Z,$$

then if $NCI \not\perp\!\!\!\perp Y$, the IV design is not exogenous.

Typically, this condition is not satisfied (couldn't find an application where it does).

- Oranges production correlated with wheat production
- Distance to Addis correlated with distance to London, etc.

IV and NCI Independence ◂ Back

Theorem

Assume that the random variable NCI satisfies the NCI assumption. If in addition

$$NCI \perp\!\!\!\perp Z,$$

then if $NCI \not\perp\!\!\!\perp Y$, *the IV design is not exogenous.*

Typically, this condition is not satisfied (couldn't find an application where it does).

- Oranges production correlated with wheat production
- Distance to Addis correlated with distance to London, etc.

Possible for violation of exclusion restriction

- Therefore, unique for IV design

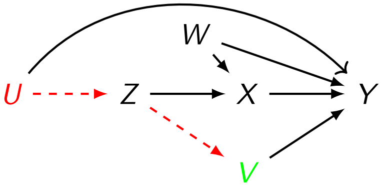
Alternative Path Variable - General Definition ◂ Back

In many cases, more than one threat might exist. In these cases, we can define an APV more generally:

Definition

A random variable U is an APV if there exists a random variable V such that the following four conditions hold:

1. **Latent IV Validity**: $Z \perp\!\!\!\perp Y(x) | U, V$
2. **Path Indication** $Z \perp\!\!\!\perp Y(x) | V \rightarrow Z \perp\!\!\!\perp U | V$



Alternative Path Variable - General Definition

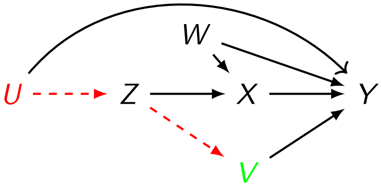
◂ Back

In many cases, more than one threat might exist. In these cases, we can define an APV more generally:

Definition

A random variable U is an APV if there exists a random variable V such that the following four conditions hold:

1. **Latent IV Validity**: $Z \perp\!\!\!\perp Y(x)|U, V$
2. **Path Indication** $Z \perp\!\!\!\perp Y\,(x)\,|V \rightarrow Z \perp\!\!\!\perp U|V$
3. **Direct IV Link**: $Z \perp\!\!\!\perp U|V \rightarrow Z \perp\!\!\!\perp U$
4. **V-Validity**: $Z \perp\!\!\!\perp Y\,(x) \rightarrow Z \perp\!\!\!\perp Y\,(x)\,|V$



Alternative Path Variable - General Definition

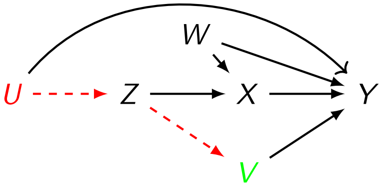
Back

In many cases, more than one threat might exist. In these cases, we can define an APV more generally:

Definition

A random variable U is an APV if there exists a random variable V such that the following four conditions hold:

1. **Latent IV Validity**: $Z \perp\!\!\!\perp Y(x) | U, V$
2. **Path Indication** $Z \perp\!\!\!\perp Y(x) | V \rightarrow Z \perp\!\!\!\perp U | V$
3. **Direct IV Link**: $Z \perp\!\!\!\perp U | V \rightarrow Z \perp\!\!\!\perp U$
   - The (potential) link between Z and U is not through V
4. **V-Validity**: $Z \perp\!\!\!\perp Y(x) \rightarrow Z \perp\!\!\!\perp Y(x) | V$



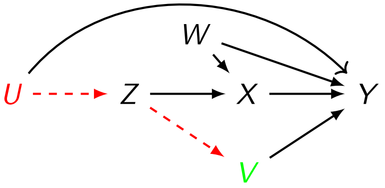
Alternative Path Variable - General Definition ◂ Back

In many cases, more than one threat might exist. In these cases, we can define an APV more generally:

Definition

A random variable U is an APV if there exists a random variable V such that the following four conditions hold:

1. **Latent IV Validity**: $Z \perp\!\!\!\perp Y(x) | U, V$
2. **Path Indication** $Z \not\perp\!\!\!\perp Y(x) | V \rightarrow Z \not\perp\!\!\!\perp U | V$
3. **Direct IV Link**: $Z \not\perp\!\!\!\perp U | V \rightarrow Z \not\perp\!\!\!\perp U$
   - The (potential) link between Z and U is not through V
4. **V-Validity**: $Z \perp\!\!\!\perp Y(x) \rightarrow Z \perp\!\!\!\perp Y(x) | V$
   - Conditioning on V does not "ruin" the IV



Sketch of Proof ◂ Back

- From the NC assumption $Z \perp\!\!\!\perp NC | U$, if $Z \not\perp\!\!\!\perp NC$ then $Z \not\perp\!\!\!\perp U$
- Given the definition of APV, this implies $Z \not\perp\!\!\!\perp Y(x)$
- This implies either independence assumption or exclusion restriction are violated
- Details

Negative Control Exposure

So far we discussed cases searching $Z \perp\!\!\!\perp NC$

- Less often, researchers search for a link with the outcome

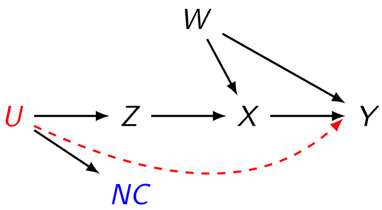
Negative Control Exposure

So far we discussed cases searching $Z \perp\!\!\!\perp NC$

- Less often, researchers search for a link with the outcome

Ex: Arteaga and Barone (2022)

- Opioid marketing (X) $\rightarrow$ addiction (Y)
- Use cancer mortality in region as IV (Z)
- Use mortality by other factors as NC (NC)



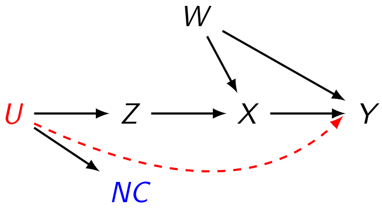
Negative Control Exposure

So far we discussed cases searching $Z \perp\!\!\!\perp NC$

- Less often, researchers search for a link with the outcome

Ex: Arteaga and Barone (2022)

- Opioid marketing (X) $\rightarrow$ addiction (Y)
- Use cancer mortality in region as IV (Z)
- Use mortality by other factors as NC (NC)



Theory requires modification. For instance, must test conditional independence

$$NC \perp\!\!\!\perp Y | Z$$

Bias in Bias Correction (B_1) Back

$$B_1 = \frac{\Pr\left[X_1(1) = 0\right]}{\Pr\left[X_1(1) > X_1(0)\right]} E\left[Y_1(1,0) - Y_1(0,0) | Z_1 = 1, X_1(1) = 0\right]$$
$$+ \frac{\Pr\left(X_1(0) = 1\right)}{\Pr\left[X_1(1) > X_1(0)\right]} \Big(E\left[Y_1(1,1) - Y_1(0,0) \mid Z_1 = 1, X_1(0) = 1\right]$$
$$- E\left[Y_1(0,1) - Y_1(0,0) \mid Z_1 = 0, X_1(0) = 1\right] \Big).$$

Scenario 2 - No IV ◂ Back

- Assume $Z_0 = 0$ - IV not available
- Treatment is available $X_0 = X_0(z_0) = X_0(0)$
- Assume also "same-type": $\forall z,\ X_0(z) = X_1(z)$
- Hence: Treatment exists only for always takers $X_0 = X_1(0)$.

Scenario 2 - No IV ◂ Back

- Assume $Z_0 = 0$ - IV not available
- Treatment is available $X_0 = X_0(z_0) = X_0(0)$
- Assume also "same-type": $\forall z,\ X_0(z) = X_1(z)$
- Hence: Treatment exists only for always takers $X_0 = X_1(0)$.

Theorem (No-IV)

Under the above assumptions

(i) $DiW = E[Y_1(1,1) - Y_1(0,0)|X_1 = 1, X_1(1) > X_1(0)] + B_2$

Scenario 2 - No IV ◂ Back

- Assume $Z_0 = 0$ - IV not available
- Treatment is available $X_0 = X_0(z_0) = X_0(0)$
- Assume also "same-type": $\forall z,\ X_0(z) = X_1(z)$
- Hence: Treatment exists only for always takers $X_0 = X_1(0)$.

Theorem (No-IV)

Under the above assumptions

(i) $DiW = E[Y_1(1,1) - Y_1(0,0)|X_1 = 1, X_1(1) > X_1(0)] + B_2$

(ii) *If, in addition, Exclusion holds DiW is the causal effect on* ***treated*** *compliers*

$$DiW = E[Y_1(x_1 = 1) - Y_1(x_1 = 0)|X_1 = 1, X_1(1) > X_1(0)]$$

Scenario 3 - No IV-treatment link ◂ Back

- The IV is identical in both settings $Z_0 = Z_1$
- "Same-type": $\forall z$, $X_0(z) = X_1(z)$
- But the IV is not affecting the treatment
  - Only always takers treated $X_0 = X_0(0) = X_1(0)$
- So the NCO is $Y_0 = Y_0(Z_0, X_0(0)) = Y_0(Z_1, X_1(0))$

Scenario 3 - No IV-treatment link Back

- The IV is identical in both settings $Z_0 = Z_1$
- "Same-type": $\forall z,\ X_0(z) = X_1(z)$
- But the IV is not affecting the treatment
  - Only always takers treated $X_0 = X_0(0) = X_1(0)$
- So the NCO is $Y_0 = Y_0(Z_0, X_0(0)) = Y_0(Z_1, X_1(0))$

Theorem (No IV-treatment link)

Under the above assumptions,

$$DiW = [Y_1(1,1) - Y_1(1,0) | X_1 = 1, X_1(1) > X_1(0)]$$

DiW identifies the causal effects on the **treated** compliers.

Testing Appendix

F-Test

In many data sets multiple negative control exists $\overline{NC} = (NC_1, ..., NC_M)$

- Can test jointly $Z \perp\!\!\!\perp \overline{NCO}$ or $Y \perp\!\!\!\perp \overline{NCI} | Z$ Conditions

F-Test

In many data sets multiple negative control exists $\overline{NC} = (NC_1, ..., NC_M)$

- Can test jointly $Z \perp\!\!\!\perp \overline{NCO}$ or $Y \perp\!\!\!\perp \overline{NCI}|Z$ Conditions

We can write the linear model

$$Z = \beta_0 + \beta_C^T C + \beta_{NCO}^T \overline{NCO} + \epsilon_Z$$

where $E\left[\epsilon_Z | C, \overline{NCO}\right] = 0$.

$$H_0 : \beta_{NCO} = 0$$

- Can use the standard F-test for the null hypothesis Details
- Rejection of $H_0 \Rightarrow$ violation of outcome independence, exclusion, or rich covariates
- Similarly for NCI

F-Test

In many data sets multiple negative control exists $\overline{NC} = (NC_1, ..., NC_M)$

- Can test jointly $Z \perp\!\!\!\perp \overline{NCO}$ or $Y \perp\!\!\!\perp \overline{NCI}|Z$ Conditions

We can write the linear model

$$Z = \beta_0 + \beta_C^T C + \beta_{NCO}^T \overline{NCO} + \epsilon_Z$$

where $E\left[\epsilon_Z | C, \overline{NCO}\right] = 0$.

$$H_0 : \beta_{NCO} = 0$$

- Can use the standard F-test for the null hypothesis Details
- Rejection of $H_0 \Rightarrow$ violation of outcome independence, exclusion, or rich covariates
- Similarly for NCI

Advantage: Does not require multiple hypotheses testing

Non-Linear Methods

Sometimes the researcher may not want a linear test:

- Does not assume rich covariates
  - For instance, when not using 2SLS (e.g Abadie, 2003)
- Non-linear link with negative controls

More general non parametric methods exist. For example:

1. Generalized Additive Models (GAM)
2. Invariant target prediction (Heinze-Deml et al., 2018)
   - Compare OOS predictions with only (C) compared to both (C, NC)
   - Similar to Ludwig et al. (2017) with control variables

Non-Linear Methods

Sometimes the researcher may not want a linear test:

- Does not assume rich covariates
  - For instance, when not using 2SLS (e.g Abadie, 2003)
- Non-linear link with negative controls

More general non parametric methods exist. For example:

1. Generalized Additive Models (GAM)
2. Invariant target prediction (Heinze-Deml et al., 2018)
   - Compare OOS predictions with only (C) compared to both (C, NC)
   - Similar to Ludwig et al. (2017) with control variables
3. Kernel-based conditional independence test (Zhang et al., 2011)
   - Beyond mean-independence

Under construction: package for all methods

Simulations

Simulations

Conditions ◂ Back

To test for independence with multiple negative control, one need to assume that

$$\forall i : Z \perp\!\!\!\perp NC_i \Rightarrow Z \perp\!\!\!\perp \overline{NC}$$

Theoretically, this might not hold as pairwise independence does not imply mutual independence.
However in practice, this only rules out knife edge cases where there exists some function g such that

$$Z \not\perp\!\!\!\perp g\left(NC_1, ..., NC_M\right)$$

while $Z \perp\!\!\!\perp NC_i$.
In paper we show example for violation. But also in example, small changes in the DGP will generate dependency with one of the negative controls.

Generalized Additive Models ◂ Back

We can generally estimate

- $$Z = \sum_{j=1}^{J} g_j^{(C)}(C_j) + \sum_{k=1}^{K} g_k^{(NC)}(NC_k) + \epsilon_Z$$
- With rich covariates $g_j^{(C)}$ are linear functions
- $g_k^{(NC)}$ are smoothed functions typically estimated with splines (Wood, 2006)
- Can also include interactions
- Can test $H_0 : g_k^{(NC)} \equiv 0$ with GLRT
- Need to assume $\epsilon_Z \sim N(0, \sigma^2)$
- In linear case, can use F-test

Applications

1. Detect All Negative Controls in the Data

In highly-cited econ papers, the median number of NC used is 4.

- In many cases multiple unused negative controls exist
- Use theory+domain knowledge to detect all NC in data

2. Choose (Conditional) Independence Test

When multiple negative controls exist, they can be combined into one test

- F-test is a good option for 2SLS

$$Z = \beta_0 + \beta_C^T C + \beta_{NCO}^T NCO + \epsilon_Z$$

when NCO is a vector of NCOs. $H_0 : \beta_{NCO} = 0$

2. Choose (Conditional) Independence Test

When multiple negative controls exist, they can be combined into one test

- F-test is a good option for 2SLS

$$Z = \beta_0 + \beta_C^T C + \beta_{NCO}^T NCO + \epsilon_Z$$

when NCO is a vector of NCOs. $H_0 : \beta_{NCO} = 0$

More sophisticated options that require more data are also valid

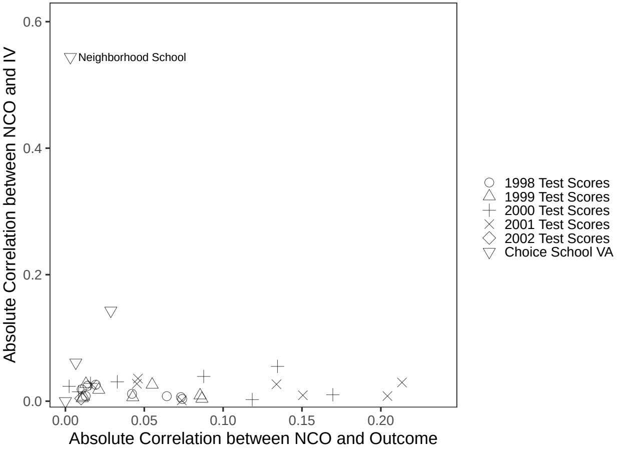
- GAM can be used for non-linear dependencies
- Non-parametric methods avoid testing functional form assumptions

3. Post-mortem Analysis ◂ Back

We recommend testing which NC are correlated with IV and outcome

3. Post-mortem Analysis ◂ Back

We recommend testing which NC are correlated with IV and outcome



Ex: Deming (2014, AER). IV is school lottery interacted with schools value-added

Autor, Dorn & Hanson (2013)

A shift-share IV to study the local impact of Chinese import penetration

- Use their original data
  - (Cleaned version of the census)

Identify all potential negative controls

- ADH use change in mfg. emp 1990-2000 for falsification
- We use everything that occurred before 2000
  - Trends in other industries, sub-populations (gender, education), etc.

Results

Deming (2014)

Uses IV based on school lotteries to study bias in school VA models

Use the same data and IV as the original paper

- NC - everything that happened before the lottery

Allegedly, a perfect IV

- In practice it fails the test ($p<0.01$)
- Using multiple negative controls found an unexpected problem
- Fixable error in implementation

Explanation

Deming defines the following IV, when L is an indicator for winning the lottery

$$Z = \begin{cases} \text{VA 1st choice} & L = 1 \\ \text{VA in neigh.} & L = 0 \end{cases}$$

Similar to interacting L with potential difference in VA

- But requires controlling for VA in neighborhood (not random)
- Adding controls fixes the problem Plot

Summary

1. Theory of negative controls for instrumental variables
   - Z correlated with $NC \Rightarrow Z$ correlated with unobserved APV
   - Main assumption: correlation between Z, NC is via an APV

2. Our approach for negative control testing
   - Exploits more information
   - Uses information more efficiently

Generalizability

A setting that does not satisfy generalizability.

- $Z \sim N(0,1)$
- $Y_x = \begin{cases} x + Z & U \in \{u_0, u_1\} \\ x - Z & else \end{cases}$
- $P(U \in \{u_0, u_1\}) = \frac{1}{2}$

When $U \in \{u_0, u_1\}$ $\rho_{Y_x, Z} = 1$.

- But unconditionally $\rho_{Y_x, Z} = 0$

Return

Proof

- Since $Z \perp NC|U$, if $Z \not\perp NC$ then $Z \not\perp U$

Lemma

If $A \perp B|C$ *and* $B \perp C$ *then* $A \perp B$

Proof.

$$f(A|C)\,f(B|C) = f(A \cap B|C)$$

$$f(A|c)\,f(B) = f(A \cap B|c)$$

$$\int f(A|c)\,f(B)\,f(c)\,dz = \int f(A \cap B|c)\,f(c)\,dc$$

$$f(A)\,f(B) = f(A \cap B)$$

□

Proof

Lemma

If $A \perp B|C$ *and* $A \not\perp B$ *then* $A \not\perp C$ *and* $B \not\perp C$

Proof.

Assume that $A \perp B|C$ and $B \perp C$. Then by previous Lemma $A \perp B$ which contradicts the assumption. Similarly for $A \perp C$.

$$f(A)\,f(B) = f(A \cap B)$$

□

Proof

- Given the definition of APV, this implies $Z \not\perp Y_x$

Lemma

If U is an APV then $Z \not\perp U \rightarrow Z \not\perp Y_x$

Proof.

Following indivisibility $P(Z|u) \neq P(Z|u')$. This means that $\exists z_0, z_1$ such that $P(z_0/u)/P(z_1/u) \neq P(z_0/u')/P(z_1/u')$ and therefore $P(u|z_0)/P(u'/z_0) \neq P(u|z_1)/P(u'/z_1)$.
Given causality $\exists y_x$ such that $P(y_x|u) \neq P(y_x|u')$.
Marking by A the event where $U \in u, u'$.
$P(y_x|z_i, A) = P(y_x|z_i, u) * P(u|z_i, A) + P(y_x|z_i, u') * P(u'|z_i, A)$.
Assuming independence we can write
$P(y_x|z_i, A) = P(y_x|u) * P(u|z_i, A) + P(y_x|u') * P(u'|z_i, A)$. Since $P(u|z_i, A)/P(u'|z_i, A) \neq P(u|z_i, A)/P(u'|z_i, A)$ then $P(y_x|z_0, A) \neq P(y_x|z_1, A)$. Finally, from generlaizability $Z \not\perp Y_x$. □

Proof

- This occurs only if either independence assumption or exclusion restriction are violated

Lemma

If exclusion and independence both hold, $Y_x \perp Z$.

Proof.

From ER $Y_{z,x} = Y_x$ for all z.
From independence $Y_x \perp Z$. □

Return

Counter Example

- $U_1, U_2 \sim Bernoulli\,(0.5)$
  - e.g. coin flips
- $NC_i = U_i + \varepsilon_i$
- $Z = U_1 \oplus U_2$
- $Y_x = x + b * U_1 + (1 - b) * U_2$
  - where $b \sim Bernoulli\,(0.5)$
- $Z \perp Y_x$ so IV is valid!
- $\overline{U}$ does not satisfy indivisability
  - Z correlated with $\overline{U}$ but only with the xor
  - the xor is not causal
- In this case Z is valid but $Z \not\perp \overline{NC}$

Return

Counter Example

- For simplicity assume $U_1 = U_2 = U$
- $NC_i = U + V_i$ with $V_i \sim Bernoulli\,(0.5)$
- $Z = V_1 \oplus V_2$
- $Z \perp NC_i | U$ since $Z \perp V_i$
- But $Z \not\perp \overline{NC} | U$
- In this case Z is valid but $Z \not\perp \overline{NC}$

Return