Categorical Instrumental Variable Model:

Characterization, Partial Identification and Inference

Yilin Song[1] F. Richard Guo[2]

January 13, 2026

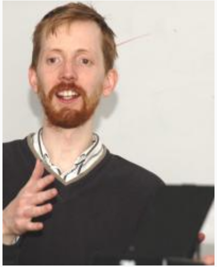
Online Causal Inference Seminar

[1]Department of Biostatistics, Columbia University

[2]Department of Statistics, University of Michigan

- Joint work with

Gary Chan (Univ. Washington)



Thomas Richardson (Univ. Washington)

Overview

1. Motivation
2. Categorical IV
3. Statistical inference
4. Data analysis

Motivation

Domestic violence calls

Background

Police responding to domestic violence calls:

1970s "hands-off" approach, "arrests only be made in cases of serious violence" in NYC

1970s/1980s women's advocacy groups were calling on police to take domestic violence more seriously and change intervention strategy

1980s $\hookrightarrow$ National Institute of Justice funds **Minneapolis Domestic Violence Experiment**

Minneapolis Domestic Violence Experiment

Domestic violence 911 calls. Both victim and offender must be at scene to be included in the study.

▶ When an officer believes that a domestic violence case has occurred, he/she is **instructed to** take one of the 3 courses of action by a lottery system:

- ARR : Arrest the offender.
- ADV : Advice and mediation of disputes.
- SEP : Separate the offender from the victim for 8 hours.

☞ This instruction is **randomized**.

A total of 314 cases were included in the study, and the suspects' **re-offence statuses** were followed up after a 6-month period through self-reports or from a police database.

Minneapolis Domestic Violence Experiment

- Z: action instructed
- X: action taken
- Y: no re-offence / re-offence in 6-month follow-up

	$X = \textsc{Arr}$	$X = \textsc{Adv}$	$X = \textsc{Sep}$
$Z = \textsc{Arr}$	81/10	0/0	1/0
$Z = \textsc{Adv}$	15/3	69/15	3/3
$Z = \textsc{Sep}$	21/5	4/1	62/20

Minneapolis Domestic Violence Experiment

- Z: action instructed
- X: action taken
- Y: no re-offence / re-offence in 6-month follow-up

	$X = \text{Arr}$	$X = \text{Adv}$	$X = \text{Sep}$	
$Z = \text{Arr}$	81/10	0/0	1/0	82/10
$Z = \text{Adv}$	15/3	69/15	3/3	87/21
$Z = \text{Sep}$	21/5	4/1	62/20	87/26

re-offence%	Arr	Adv	Sep
▶ Intent-to-treat (ITT)	11%	19%	23%

Minneapolis Domestic Violence Experiment

- Z: action instructed
- X: action taken
- Y: no re-offence / re-offence in 6-month follow-up

	$X = \textsc{Arr}$	$X = \textsc{Adv}$	$X = \textsc{Sep}$
$Z = \textsc{Arr}$	81/10	0/0	1/0
$Z = \textsc{Adv}$	15/3	69/15	3/3
$Z = \textsc{Sep}$	21/5	4/1	62/20

re-offence%	Arr	Adv	Sep
▶ Intent-to-treat (ITT)	11%	19%	23%
▶ Per Protocol (PP)	11%	18%	24%

☞ *The study found that the offenders assigned to be arrested had lower rates of re-offending than offenders assigned to counseling or temporarily sent away.* (Sherman & Berk, 1984)

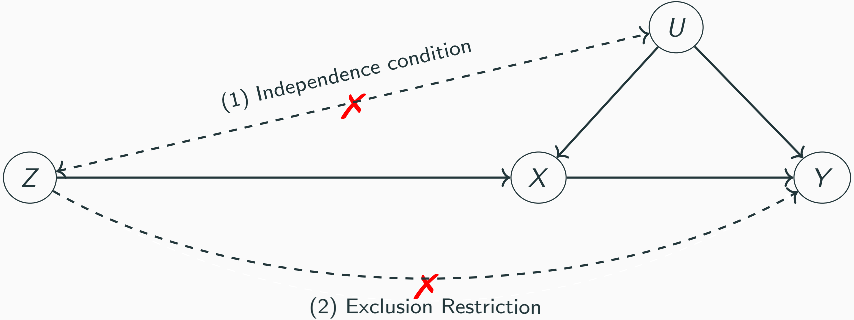
Minneapolis Domestic Violence Experiment

- Z: action instructed
- X: action taken
- Y: no re-offence / re-offence in 6-month follow-up

	$X = \textsc{Arr}$	$X = \textsc{Adv}$	$X = \textsc{Sep}$
$Z = \textsc{Arr}$	81/10	0/0	1/0
$Z = \textsc{Adv}$	15/3	69/15	3/3
$Z = \textsc{Sep}$	21/5	4/1	62/20

Z is (1) randomized and (2) has no effect on re-offence other than through the action taken.

☞ A categorical instrumental variable (IV) model with $|\mathcal{Z}| = |\mathcal{X}| = 3$ and $|\mathcal{Y}| = 2$.


U
(1) Independence condition
Z
X
Y
(2) Exclusion Restriction

Existing literature

Bounds for binary IV When $|\mathcal{Z}| = 2$, ATE is **partially identified** by the Manski–Robins bound or the (narrower) Balke–Pearl bound depending on the assumption. Generalized to $|\mathcal{Z}| \geq 2$ by Richardson and Robins (2014). A comprehensive discussion of the underlying assumptions and results is given in Swanson et al. (2018).

Falsification Testing whether a given observed distribution is compatible with particular sets of IV assumptions (Pearl, 1995; Bonet, 2001; Wang et al., 2017; Kédagni & Mourifié, 2020; Bhadane et al., 2025).

Other related work Beresteanu et al. (2012), Russell (2021), and Luo and Wang (2017) based on "random set theory".

Existing Literature

LATE for binary IV Effect among *compliers* is point-identified (Imbens & Angrist, 1994), but need monotonicity assumption (i.e., *no defiers*).

Existing Literature

LATE for binary IV Effect among *compliers* is point-identified (Imbens & Angrist, 1994), but need monotonicity assumption (i.e., *no defiers*).

- requires determining whether a subject would have been a complier had they been in the experiment (Kennedy et al., 2020).

Existing Literature

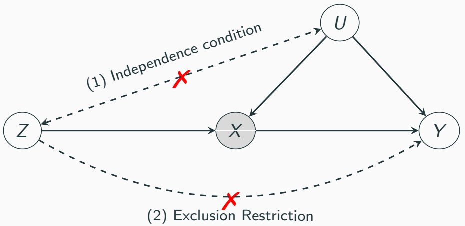
LATE for binary IV Effect among *compliers* is point-identified (Imbens & Angrist, 1994), but need monotonicity assumption (i.e., *no defiers*).

- requires determining whether a subject would have been a complier had they been in the experiment (Kennedy et al., 2020).
- e.g. requires judging that had a domestic violence incident happened during the course of the study, the responding officer would have judged it according to the assigned strategy, whatever that was.

Procrustean analysis: Pretending it is a binary IV?

- Z: action instructed
- X: action taken
- Y: no re-offence / re-offence in 6-month follow-up

	$X = \textsc{Arr}$	$X = \textsc{Adv}$	$X = \textsc{Sep}$
$Z = \textsc{Arr}$	81/10	0/0	1/0
$Z = \textsc{Adv}$	15/3	69/15	3/3
$Z = \textsc{Sep}$	21/5	4/1	62/20



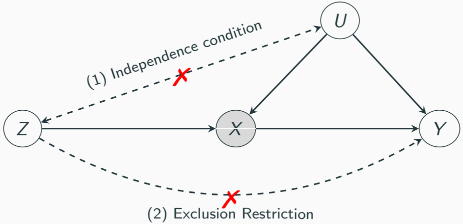
▶ Suppose we only want to contrast $\textsc{Adv}$ vs $\textsc{Sep}$. Can we just drop the first column?

☞ **No**, because that would be conditioning on $X \neq \textsc{Arr}$, making Z not independent of $Y(x)$ anymore.

Procrustean analysis: Pretending it is a binary IV?

- Z: action instructed
- X: action taken
- Y: no re-offence / re-offence in 6-month follow-up

	$X = \textsc{Arr}$	$X = \textsc{Adv}$	$X = \textsc{Sep}$
$Z = \textsc{Arr}$	81/10	0/0	1/0
$Z = \textsc{Adv}$	15/3	69/15	3/3
$Z = \textsc{Sep}$	21/5	4/1	62/20



▶ Suppose we only want to contrast $\textsc{Adv}$ vs $\textsc{Sep}$. Can we just drop the first column?

☞ **No**, because that would be conditioning on $X \neq \textsc{Arr}$, making Z not independent of $Y(x)$ anymore.

Our Goal ▶ Characterization and partial identification of a generic categorical IV model.

Categorical IV

Objective

Consider a categorical IV model with

$$\mathcal{Z} = [Q] \equiv \{1, \dots, Q\}, \quad \mathcal{X} = [K] \equiv \{1, \dots, K\}, \quad \mathcal{Y} = [M] \equiv \{1, \dots, M\}.$$

The model obeys the usual consistency assumption and
(1) independence/exogeneity condition, and
(2) exclusion restriction.

- **Characterize** the set of $P(Y(x_1), \dots, Y(x_K))$ that is compatible with $P(X, Y \mid Z)$.
- **Falsification test** of whether a given observed distribution is compatible with the IV assumptions.
- **Partially identify** ATEs such as $\mathbb{E}[Y(x_k) - Y(x_{k'})]$ or any linear functionals of $P(Y(x_1), \dots, Y(x_K))$.
- **Construct confidence intervals** for ATEs such as $\mathbb{E}[Y(x_k) - Y(x_{k'})]$ or any linear functionals of $P(Y(x_1), \dots, Y(x_K))$.

What does a characterization mean? The case of binary X, Y

In an observational study with $|\mathcal{Z}| = 1$, we have

$$0 \leq \%HE \leq P(X=0, Y=0) + P(X=1, Y=1)$$

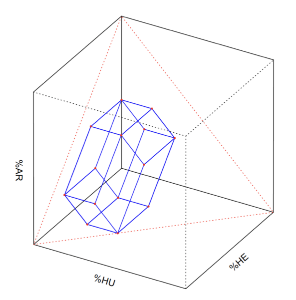
$$0 \leq \%HU \leq P(X=0, Y=1) + P(X=1, Y=0)$$

$$0 \leq \%NR \leq P(X=0, Y=0) + P(X=1, Y=0)$$

$$0 \leq \%AR \leq P(X=0, Y=1) + P(X=1, Y=1)$$

$$P(X=0, Y=1) \leq P\{Y(0)=1\} \leq 1 - P(X=0, Y=0)$$

$$P(X=1, Y=1) \leq P\{Y(1)=1\} \leq 1 - P(X=1, Y=0)$$



6 pairs of parallel planes

%HE: $P(Y(0)=0, Y(1)=1)$; %HU: $P(Y(0)=1, Y(1)=0)$;
%NR: $P(Y(0)=0, Y(1)=0)$; %AR: $P(Y(0)=1, Y(1)=1)$.

What does a characterization mean? The case of binary X, Y

In an observational study with $|\mathcal{Z}| = 1$, we have

$$0 \leq \%HE \leq P(X=0, Y=0) + P(X=1, Y=1)$$

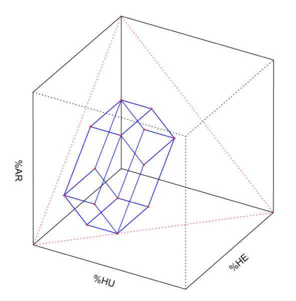
$$0 \leq \%HU \leq P(X=0, Y=1) + P(X=1, Y=0)$$

$$0 \leq \%NR \leq P(X=0, Y=0) + P(X=1, Y=0)$$

$$0 \leq \%AR \leq P(X=0, Y=1) + P(X=1, Y=1)$$

$$P(X=0, Y=1) \leq P\{Y(0)=1\} \leq 1 - P(X=0, Y=0)$$

$$P(X=1, Y=1) \leq P\{Y(1)=1\} \leq 1 - P(X=1, Y=0)$$



6 pairs of parallel planes

- Given the observed probabilities, what do I actually know about my counterfactual probabilities?

%HE: $P(Y(0)=0, Y(1)=1)$; %HU: $P(Y(0)=1, Y(1)=0)$;
%NR: $P(Y(0)=0, Y(1)=0)$; %AR: $P(Y(0)=1, Y(1)=1)$.

Simplest IV Model $\mathcal{M}_1$

The simplest IV model $\mathcal{M}_1$ is defined by the following assumptions:

1. **Consistency**

$$Y = Y(X, Z) \text{ and } X = X(Z)$$

2. **Individual-level exclusion**

$$Y(x_i, z) = Y(x_i, \widetilde{z}) \text{ for all } z, \widetilde{z} \in [Q],\ i \in [K], \text{ and } q \in [Q]$$

3. **Random assignment**

$$Z \perp\!\!\!\perp (Y(x, z), X(z) : x \in [K], z \in [Q])$$

Simplest IV Model $\mathcal{M}_1$

The simplest IV model $\mathcal{M}_1$ is defined by the following assumptions:

1. **Consistency**

$$Y = Y(X, Z) \text{ and } X = X(Z)$$

2. **Individual-level exclusion**

$$Y(x_i, z) = Y(x_i, \tilde{z}) \text{ for all } z, \tilde{z} \in [Q],\ i \in [K], \text{ and } q \in [Q]$$

3. **Random assignment**

$$Z \perp\!\!\!\perp (Y(x, z), X(z) : x \in [K], z \in [Q])$$

▶ Our paper also considers four other IV models, $\mathcal{M}_2, \ldots, \mathcal{M}_5$, defined by weaker versions of the Exclusion and Random assignment assumptions, to which our results also apply.

☞ These models may be of possible interest in testing causal quantum theories.

IV models $\mathcal{M}_1, \ldots, \mathcal{M}_5$

	Model Name	Exclusion	Independence
$\mathcal{M}_1$	**Randomization**	**Individual-level**	**Random assignment**
$\mathcal{M}_2$	*Joint Ind. & Indiv. Excl.*	Individual-level	Joint independence
$\mathcal{M}_3$	*Joint Ind. & Stoch. Excl.*	Joint stochastic exclusion	Joint independence
$\mathcal{M}_4$	*SWIG*	Individual-level	Single-world independence
$\mathcal{M}_5$	*Latent Model*	Latent exclusion	Latent-variable exogeneity

☞ $\mathcal{M}_1 \subseteq \mathcal{M}_2 \subseteq \mathcal{M}_3$, $\mathcal{M}_1 \subseteq \mathcal{M}_4$ and $\mathcal{M}_2 \subseteq \mathcal{M}_5$.

$\mathcal{M}_1$

$\rotatebox{}{\subseteq}$ $\quad$ $\subseteq$

$\mathcal{M}_2$ $\qquad$ $\mathcal{M}_4$

$\subseteq$ $\quad$ $\subseteq$

$\mathcal{M}_3$ $\qquad$ $\mathcal{M}_5$

Characterizing the IV model

Recall in our categorical IV,

$$\mathcal{Z} = [Q] \equiv \{1, \ldots, Q\}, \quad \mathcal{X} = [K] \equiv \{1, \ldots, K\}, \quad \mathcal{Y} = [M] \equiv \{1, \ldots, M\}.$$

Theorem 1 Under any IV model $\mathcal{M}_i$ ($i = 1, \ldots, 5$), the set of counterfactual distributions is characterized by the following set of inequalities: for each $z \in [Q]$, we have

$$P'\left(Y(x_1) \in \mathcal{V}^{(1)}, \ldots, Y(x_K) \in \mathcal{V}^{(K)}\right) \leq \sum_{i=1}^{K} P\left(X = i, Y \in \mathcal{V}^{(i)} \middle| Z = z\right), \; z \in [Q],$$

where $\mathcal{V}^{(k)}$ is a non-empty subset of $[M]$ for all $k \in [K]$ and a strict subset of $[M]$ for at least one k.

Example inequality when $|\mathcal{X}| = 2$, $|\mathcal{Y}| = 3$

Theorem 1 Under any IV model $\mathcal{M}_i$ ($i = 1, \ldots, 5$), the set of counterfactual distributions is characterized by the following set of inequalities: for each $z \in [Q]$, we have

$$P' \left(Y(x_1) \in \mathcal{V}^{(1)}, \ldots, Y(x_K) \in \mathcal{V}^{(K)} \right) \leq \sum_{i=1}^{K} P \left(X = i, Y \in \mathcal{V}^{(i)} \middle| Z = z \right), \; z \in [Q],$$

where $\mathcal{V}^{(k)}$ is a non-empty subset of $[M]$ for all $k \in [K]$ and a strict subset of $[M]$ for at least one k.

Taking $\mathcal{V}^{(1)} = \{1, 2, 3\}$ and $\mathcal{V}^{(2)} = \{1, 2\}$, gives

$$P'(Y(x_1) \in \mathcal{V}^{(1)}, Y(x_2) \in \mathcal{V}^{(2)}) \leq \sum_{i=1}^{2} P \left(X = i, Y \in \mathcal{V}^{(i)} \middle| Z = z \right),$$

which, upon subtracting from one on both sides, becomes

$$P'(Y(x_2) \neq 3) \leq 1 - P(X = 2, Y = 3 \mid Z = z). \quad (\blacktriangle)$$

Similarly, taking $\mathcal{V}^{(1)} = \{1, 2, 3\}$ and $\mathcal{V}^{(2)} = \{2, 3\}$, gives

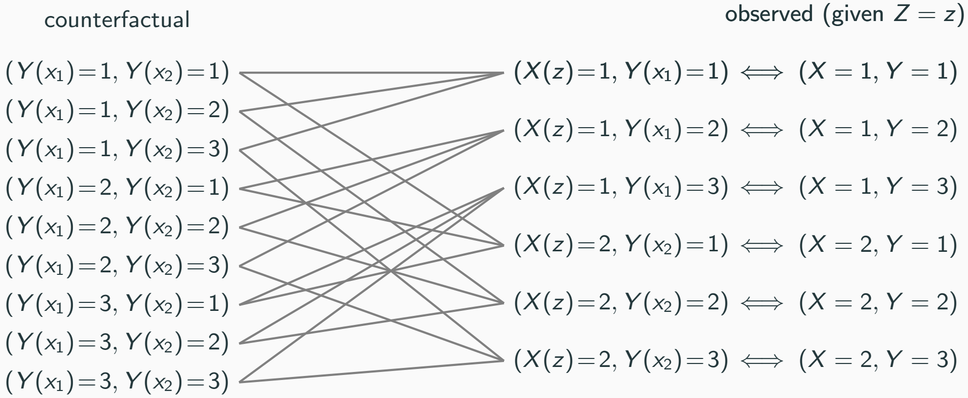
$$P'(Y(x_2) \neq 1) \leq 1 - P(X = 2, Y = 1 \mid Z = z). \quad (\bigstar)$$

Bipartite graph and compatible pairs

We can also represent each inequality graphically.

- Let us fix $Z = z$. Consider again $|\mathcal{X}| = 2$, $|\mathcal{Y}| = 3$.

☞ An edge is placed between every pair of (counterfactual, observed) values that are compatible under the IV model.

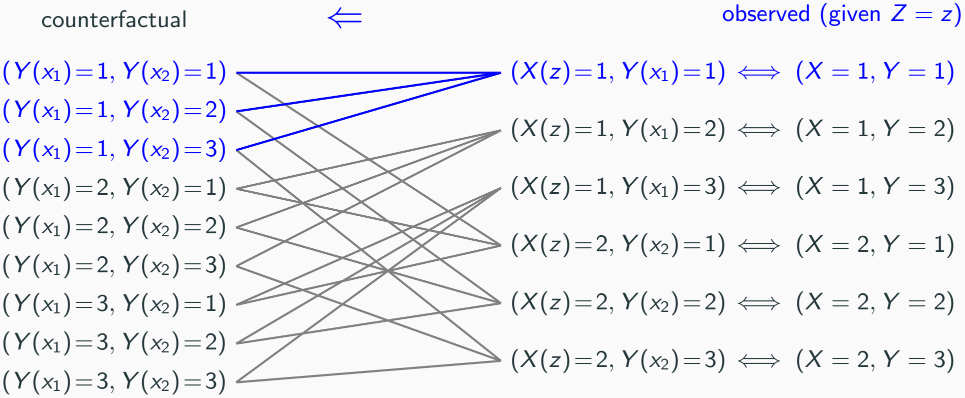


Bipartite graph and compatible pairs

We can also represent each inequality graphically.

- Let us fix $Z = z$. Consider again $|\mathcal{X}| = 2$, $|\mathcal{Y}| = 3$.

☞ An edge is placed between every pair of (counterfactual, observed) values that are compatible under the IV model.

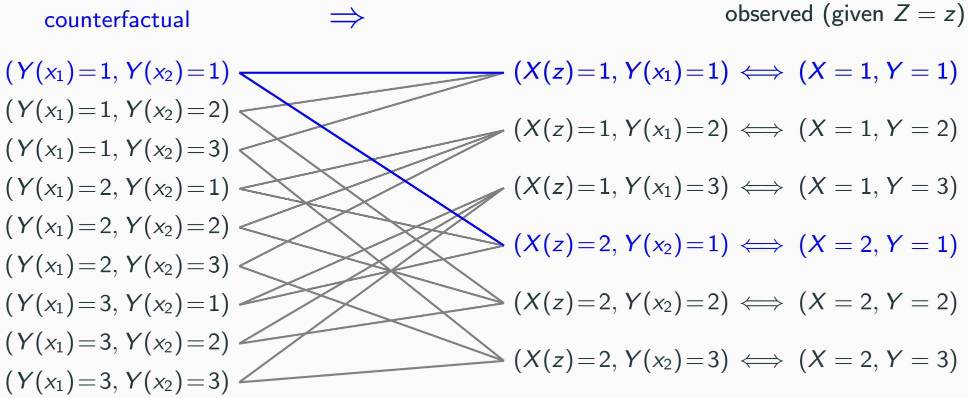


Bipartite graph and compatible pairs

We can also represent each inequality graphically.

- Let us fix $Z = z$. Consider again $|\mathcal{X}| = 2$, $|\mathcal{Y}| = 3$.

☞ An edge is placed between every pair of (counterfactual, observed) values that are compatible under the IV model.

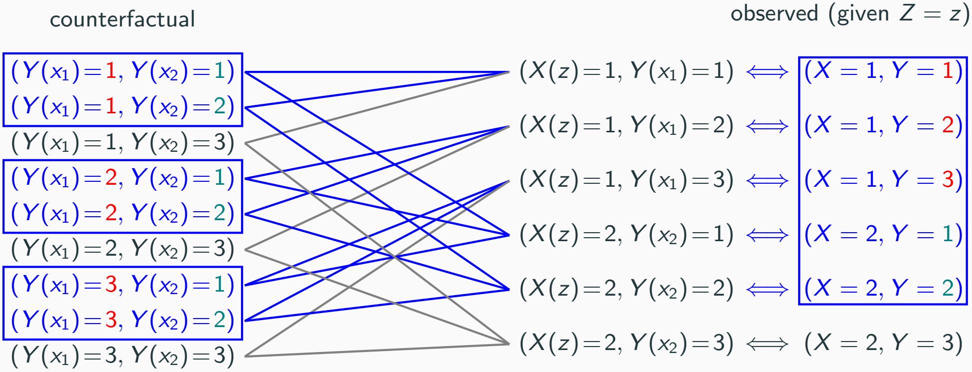


Example revisited

Our example: $\mathcal{V}^{(1)} = \{1, 2, 3\}$ and $\mathcal{V}^{(2)} = \{1, 2\}$ gives the inequality

$$P'(Y(x_1) \in \{1, 2, 3\}, Y(x_2) \in \{1, 2) \geq P(X = 1, Y \in \{1, 2, 3\} | Z = z) + P(X = 2, Y \in \{1, 2\} | Z = z),$$

$$\text{i.e.,} \quad P'(Y(x_2) \neq 3) \leq 1 - P(X = 2, Y = 3 \mid Z = z). \quad (\blacktriangle)$$



☞ This shows the necessity of the bound.

Theorem 1: proof sketch

Theorem 1 Under any IV model $\mathcal{M}_i$ $(i = 1, \ldots, 5)$, the set of counterfactual distributions is characterized by the following set of inequalities: for each $z \in [Q]$, we have

$$P'\left(Y(x_1) \in \mathcal{V}^{(1)}, \ldots, Y(x_K) \in \mathcal{V}^{(K)}\right) \leq \sum_{i=1}^{K} P\left(X = i, Y \in \mathcal{V}^{(i)} \middle| Z = z\right), \; z \in [Q],$$

where $\mathcal{V}^{(k)}$ is a non-empty subset of $[M]$ for all $k \in [K]$ and a strict subset of $[M]$ for at least one k.

Define $$\phi : \underbrace{P(Z, X, Y(x_1), \ldots, Y(x_K))}_{\in \mathcal{M}_i} \mapsto \left(\underbrace{P(Y(x_1), \ldots, Y(x_K))}_{\text{counterfactual}}, \quad \underbrace{P(X, Y \mid Z)}_{\text{observed}}\right).$$

Let $\mathcal{T}$ be the set of RHS pairs that obey the inequalities in Theorem 1.

☞ Theorem 1 is equivalent to the claim that $\phi(\mathcal{M}_i) = \mathcal{T}, \quad i = 1, \ldots, 5.$

Theorem 1: proof sketch

Theorem 1 Under any IV model $\mathcal{M}_i$ $(i = 1, \ldots, 5)$, the set of counterfactual distributions is characterized by the following set of inequalities: for each $z \in [Q]$, we have

$$P'\left(Y(x_1) \in \mathcal{V}^{(1)}, \ldots, Y(x_K) \in \mathcal{V}^{(K)}\right) \leq \sum_{i=1}^{K} P\left(X = i, Y \in \mathcal{V}^{(i)} \middle| Z = z\right), \; z \in [Q],$$

where $\mathcal{V}^{(k)}$ is a non-empty subset of $[M]$ for all $k \in [K]$ and a strict subset of $[M]$ for at least one k.

Define $$\phi : \underbrace{P(Z, X, Y(x_1), \ldots, Y(x_K))}_{\in \mathcal{M}_i} \mapsto \left(\underbrace{P(Y(x_1), \ldots, Y(x_K))}_{\text{counterfactual}}, \quad \underbrace{P(X, Y \mid Z)}_{\text{observed}}\right).$$

Let $\mathcal{T}$ be the set of RHS pairs that obey the inequalities in Theorem 1.

☞ Theorem 1 is equivalent to the claim that $\phi(\mathcal{M}_i) = \mathcal{T}, \quad i = 1, \ldots, 5$.

- $\phi(\mathcal{M}_i) \subseteq \mathcal{T}$: Relatively easy to show.

Theorem 1: proof sketch

Theorem 1 Under any IV model $\mathcal{M}_i$ ($i = 1, \ldots, 5$), the set of counterfactual distributions is characterized by the following set of inequalities: for each $z \in [Q]$, we have

$$P'\left(Y(x_1) \in \mathcal{V}^{(1)}, \ldots, Y(x_K) \in \mathcal{V}^{(K)}\right) \leq \sum_{i=1}^{K} P\left(X = i, Y \in \mathcal{V}^{(i)} \,\middle|\, Z = z\right), \; z \in [Q],$$

where $\mathcal{V}^{(k)}$ is a non-empty subset of $[M]$ for all $k \in [K]$ and a strict subset of $[M]$ for at least one k.

Define $$\phi : \underbrace{P(Z, X, Y(x_1), \ldots, Y(x_K))}_{\in \mathcal{M}_i} \mapsto \left(\underbrace{P(Y(x_1), \ldots, Y(x_K))}_{\text{counterfactual}}, \quad \underbrace{P(X, Y \mid Z)}_{\text{observed}} \right).$$

Let $\mathcal{T}$ be the set of RHS pairs that obey the inequalities in Theorem 1.

☞ Theorem 1 is equivalent to the claim that $\phi(\mathcal{M}_i) = \mathcal{T}, \quad i = 1, \ldots, 5.$

- $\phi(\mathcal{M}_i) \subseteq \mathcal{T}$: Relatively easy to show.
- $\phi(\mathcal{M}_i) \supseteq \mathcal{T}$: Much harder. We show this using a finite-space version of Strassen's theorem (Koperberg, 2024; Strassen, 1965). ▶ More user-friendly than the "random set theory".

Example continued

Recall that we obtained two inequalities

$$P'(Y(x_2) \neq 3) \leq 1 - P(X = 2, Y = 3 \mid Z = z), \quad (\blacktriangle)$$
$$P'(Y(x_2) \neq 1) \leq 1 - P(X = 2, Y = 1 \mid Z = z). \quad (\bigstar)$$

☞ $(\blacktriangle) + (\bigstar)$ leads to

$$P'(Y(x_2) = 2) \leq 1 - P(X = 2, Y = 1 \mid Z = z) - P(X = 2, Y = 3 \mid Z = z) \quad (\blacksquare),$$

which, however, is another inequality corresponding to $\mathcal{V}^{(1)} = \{1, 2, 3\}$ and $\mathcal{V}^{(2)} = \{2\}$ in Theorem 1.

Example continued

Recall that we obtained two inequalities

$$P'(Y(x_2) \neq 3) \leq 1 - P(X = 2, Y = 3 \mid Z = z), \quad (\blacktriangle)$$
$$P'(Y(x_2) \neq 1) \leq 1 - P(X = 2, Y = 1 \mid Z = z). \quad (\bigstar)$$

☞ $(\blacktriangle) + (\bigstar)$ leads to

$$P'(Y(x_2) = 2) \leq 1 - P(X = 2, Y = 1 \mid Z = z) - P(X = 2, Y = 3 \mid Z = z) \quad (\blacksquare),$$

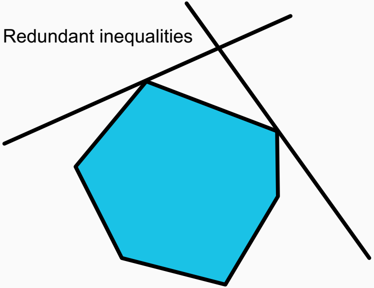
which, however, is another inequality corresponding to $\mathcal{V}^{(1)} = \{1, 2, 3\}$ and $\mathcal{V}^{(2)} = \{2\}$ in Theorem 1.

☞ The set of inequalities in Theorem 1 can be redundant.

A non-redundant set of inequalities

Given a set of inequalities, an individual inequality can be redundant if it is implied by other inequalities in the set.

☞ By characterizing the extreme points (and hence all the facets of the polytope), we further arrive at a set of non-redundant inequalities.



A non-redundant set of inequalities

Recall that **Theorem 1** gives inequalities

$$P'\left(Y(x_1) \in \mathcal{V}^{(1)}, \ldots, Y(x_K) \in \mathcal{V}^{(K)}\right) \le \sum_{i=1}^{K} P\left(X = i, Y \in \mathcal{V}^{(i)} \middle| Z = z\right), \quad z \in [Q],$$

where $\emptyset \neq \mathcal{V}^{(k)}$ is a subset of $[M]$ for all $k \in [K]$ and a strict subset of $[M]$ for at least one k.

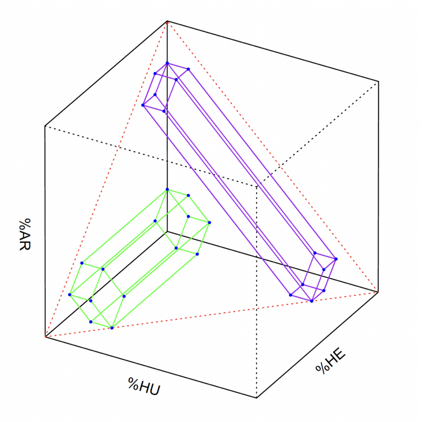
Theorem 2 The inequalities above can be reduced to a subset that only consists of inequalities that satisfy either

1. for at least two values $k \neq k^*$, we have $\mathcal{V}^{(k)} \neq [M]$ and $\mathcal{V}^{(k^*)} \neq [M]$, or
2. there exist k^* and $m \in [M]$ such that $\mathcal{V}^{(k^*)} = [M] \setminus \{m\}$ and $\mathcal{V}^{(k)} = [M]$ for every $k \neq k^*$.

This set of inequalities are non-redundant and characterize the IV model.

☞ $\#$ inequalities $= O(Q\,2^{MK}) \ll O(Q\,2^{M^K})$ Artstein's inequalities in random set theory.

Falsification of the IV model



☞ Empty intersection of the $|\mathcal{Z}|$ polytopes defining the joint counterfactual probability distribution given each instrument arm $Z = z$ implies falsification of the categorical IV model.

Statistical inference

Inference targets

1. Construct confidence intervals[1] for ATEs

$$\tau_{k,k'} := \mathbb{E}[Y(X = k) - Y(X = k')], \quad 1 \leq k < k' \leq K$$

that contrast all pairs of treatments.

☞ Every ATE $\tau_{k,k'}$ is a linear functional of the counterfactual distribution $P(Y(x_1), \ldots, Y(x_K))$. By projecting the polytope we characterized, we can get a tight lower and upper bound.

But this ignores the sampling variability in $\widehat{P}(X, Y \mid Z)$.

2. Falsification of IV model.

☞ If the data strongly suggests that the observed distribution $P(X, Y \mid Z)$ does not admit any underlying categorical IV, we should be able to raise an alarm.

[1] CIs that cover the population-level identified intervals with prescribed coverage.

Multinomial LRT

Consider a multinomial experiment over $N \geq 2$ categories

$$(X_1, \dots, X_N) \sim \text{Mult}(n; (p_1, \dots, p_N)),$$

where $(p_1, \dots, p_N) \in \Delta^{N-1}$. Let $(\widehat{p}_1, \dots, \widehat{p}_N) := (X_1, \dots, X_N)/n$.

☞ The Wilks' theorem states that the likelihood ratio test (LRT) statistic

$$2n\,\mathcal{D}_{\text{KL}}(\widehat{p} \| p) = 2n \sum_{i=1}^{k} \widehat{p}_i \log \frac{\widehat{p}_i}{p_i} \rightarrow_d \chi^2_{N-1}, \quad n \to \infty.$$

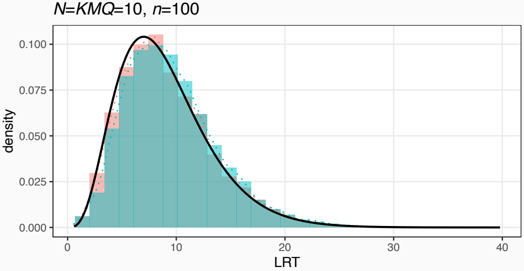
☞ Can we use this reliably for inference?

Wilks in the wild

☞ Simulate the null distribution of the LRT statistic under two 'typical' p's:

■ "sparse" $p \sim \text{Dir}(\alpha_1 = \cdots = \alpha_N = 0.5)$, ■ "dense" $p \sim \text{Dir}(\alpha_1 = \cdots = \alpha_N = 1)$.

We keep $n/N = 10$ and compare these to the Wilks' chi-squared under growing $N = KMQ$.

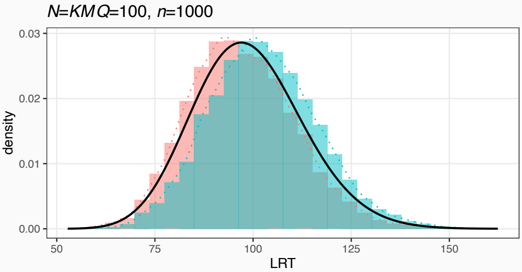


Wilks in the wild

☞ Simulate the null distribution of the LRT statistic under two 'typical' p's:

$$\blacksquare\text{“sparse”}\;\; p \sim \mathrm{Dir}(\alpha_1 = \cdots = \alpha_N = 0.5), \qquad \blacksquare\text{“dense”}\;\; p \sim \mathrm{Dir}(\alpha_1 = \cdots = \alpha_N = 1).$$

We keep $n/N = 10$ and compare these to the Wilks' chi-squared under growing $N = KMQ$.

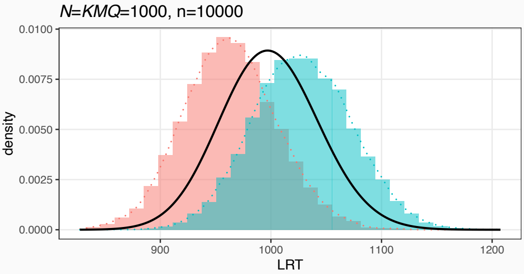


Wilks in the wild

☞ Simulate the null distribution of the LRT statistic under two 'typical' p's:

$$\text{■ “sparse” } p \sim \text{Dir}(\alpha_1 = \cdots = \alpha_N = 0.5), \quad \text{■ “dense” } p \sim \text{Dir}(\alpha_1 = \cdots = \alpha_N = 1).$$

We keep $n/N = 10$ and compare these to the Wilks' chi-squared under growing $N = KMQ$.



☞ $np_i > 5$ as a rule of thumb for CLT?

☞ Even ensuring solid coverage for p under $N = 2$ is non-trivial (Clopper & Pearson, 1934; Brown et al., 2001).

A finite-sample Chernoff bound

☞ For conducting inference, we use a finite-sample Chernoff bound (Guo & Richardson, 2021)

$$P\left(n\,\mathcal{D}_{\mathrm{KL}}(\widehat{p}\|p) > t\right) \leq \min_{\lambda\in[0,1]} \exp(-\lambda t) G_{N,n}(\lambda),$$

where $G_{N,n}$ is an explicit upper bound on the moment generating function of the LRT that only depends on the number of categories N and the sample size n.

A finite-sample Chernoff bound

☞ For conducting inference, we use a finite-sample Chernoff bound (Guo & Richardson, 2021)

$$P\left(n\,\mathcal{D}_{\mathrm{KL}}(\widehat{p}\|p) > t\right) \leq \min_{\lambda\in[0,1]} \exp(-\lambda t) G_{N,n}(\lambda),$$

where $G_{N,n}$ is an explicit upper bound on the moment generating function of the LRT that only depends on the number of categories N and the sample size n.

Finite-sample confidence region. With probability at least $1-\alpha$, it holds that

$$\sum_{z=1}^{Q} n_z \mathcal{D}_{\mathrm{KL}}\left(\widehat{p}_z \| p_z\right) \leq t_\alpha,$$

where for each arm $z \in [Q]$, n_z is its sample size and $p_z \equiv P(X, Y \mid Z = z) \in \Delta^{KM-1}$.

☞ a convex confidence region for $(p_z : z \in [Q])$ that shrinks at $n^{-1/2}$ rate.

The critical value t_α is determined from the Chernoff bound for Q independent multinomial trials

$$P\left(\sum_{z=1}^{Q} n_z\,\mathcal{D}_{\mathrm{KL}}(\widehat{p}_z\|p_z) > t\right) \leq \min_{\lambda\in[0,1]} \exp(-\lambda t) \prod_{z=1}^{Q} G_{KM,n_z}(\lambda).$$

Inference through a convex program

☞ Given a collection linear functionals $f_1, \dots, f_J$ (e.g, ATEs), we can construct their confidence intervals $[l_1, u_1], \dots, [l_J, u_J]$ by solving a convex program.

- Programming variables:

$$p_z := P(X, Y \mid Z = z) \in \mathbb{R}^{KM}, \quad z \in [Q]$$

$$p' := P'(Y(x_1), \dots, Y(x_K)) \in \mathbb{R}^{M^K}$$

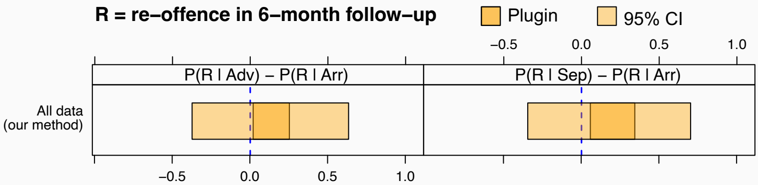
$$
\begin{aligned}
& l_j = \min f_j(p'), \quad u_j = \max f_j(p') \\
\text{s.t.} \quad & -Hp_z + H'p' \le 0, \quad z = 1, \dots, Q, \text{ (Theorem 2)} \\
& \sum_{z=1}^{Q} n_z \mathcal{D}_{\mathrm{KL}}(\widehat{p}_z \| p_z) \le t_\alpha, \\
& p_z \in \Delta^{KM-1}, \quad z = 1, \dots, Q, \\
& p' \in \Delta^{M^K-1}.
\end{aligned}
$$

Theorem 4 Under any IV model $\mathcal{M}_i$ ($i = 1, \dots, 5$), with probability at least $1 - \alpha$ it holds that $f_1 \in [l_1, u_1], \dots, f_J \in [l_J, u_J]$ with $-\infty < l_j < u_j < +\infty$ simultaneously. ☞ non-asymptotic

An alarm is raised when $l_j = +\infty, u_j = -\infty$, which indicates that the IV model is falsified by data. The probability of a false alarm is below α.

Data analysis

Revisiting Minneapolis Domestic Violence Experiment

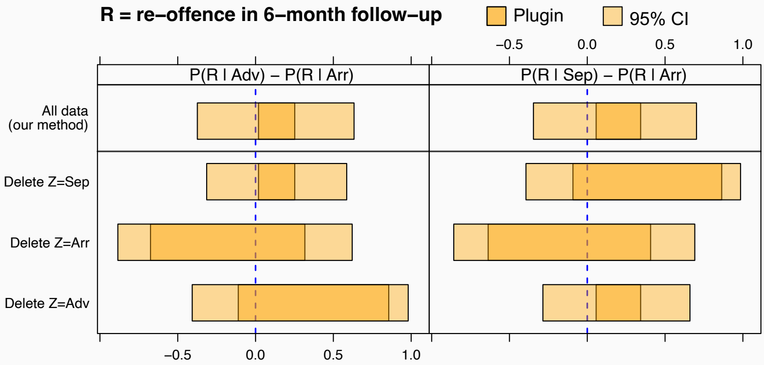


	$X = \text{Arr}$	$X = \text{Adv}$	$X = \text{Sep}$
$Z = \text{Arr}$	81/10	0/0	1/0
$Z = \text{Adv}$	15/3	69/15	3/3
$Z = \text{Sep}$	21/5	4/1	62/20

☞ Our method: Simultaneous coverage.

Number of IV inequalities $= 78 \ll 762$ Artstein's inequalities.

Revisiting Minneapolis Domestic Violence Experiment



	$X = \text{Arr}$	$X = \text{Adv}$	$X = \text{Sep}$
$Z = \text{Arr}$	81/10	0/0	1/0
$Z = \text{Adv}$	15/3	69/15	3/3
$Z = \text{Sep}$	21/5	4/1	62/20

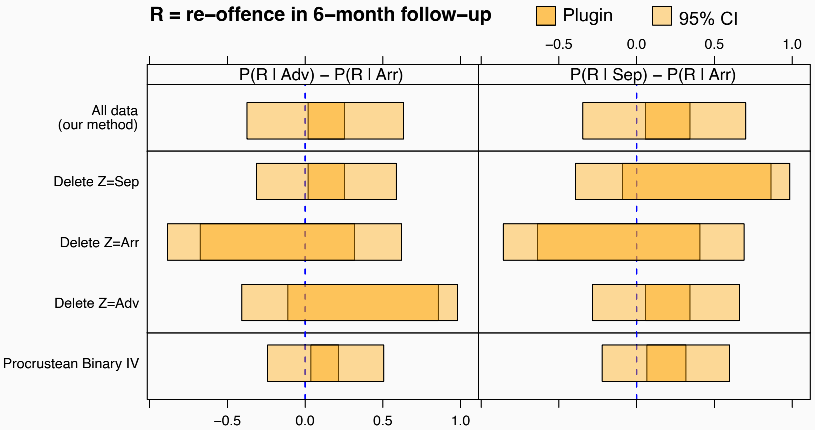
	$X = \text{Arr}$	$X = \text{Adv}$	$X = \text{Sep}$
$Z = \text{Arr}$	81/10	0/0	1/0
$Z = \text{Adv}$	15/3	69/15	3/3
$Z = \text{Sep}$	21/5	4/1	62/20

☞ Our method: Simultaneous coverage.

Number of IV inequalities $= 78 \ll 762$ Artstein's inequalities.

☞ Using more instrument arms can improve the efficiency of the results.

Revisiting Minneapolis Domestic Violence Experiment



	$X=$ Arr	$X=$ Adv	$X=$ Sep
$Z=$ Arr	81/10	0/0	1/0
$Z=$ Adv	15/3	69/15	3/3
$Z=$ Sep	21/5	4/1	62/20

	$X=$ Arr	$X=$ Adv	$X=$ Sep
$Z=$ Arr	81/10	0/0	1/0
$Z=$ Adv	15/3	69/15	3/3
$Z=$ Sep	21/5	4/1	62/20

	$X=$ Arr	$X=$ Adv	$X=$ Sep
$Z=$ Arr	81/10	0/0	1/0
$Z=$ Adv	15/3	69/15	3/3
$Z=$ Sep	21/5	4/1	62/20

- ☞ Our method: Simultaneous coverage.

Number of IV inequalities $= 78 \ll 762$ Artstein's inequalities.

- ☞ Using more instrument arms can improve the efficiency of the results.
- ☞ Procrustean binary IV analysis is biased and unnecessary.

THANKS

References

Beresteanu, A., Molchanov, I., & Molinari, F. (2012).**Partial identification using random set theory.** *Journal of Econometrics*, *166*(1), 17–32. https://doi.org/https://doi.org/10.1016/j.jeconom.2011.06.003

Bhadane, S., Mooij, J. M., Boeken, P., & Zoeter, O. (2025).**Revisiting the Berkeley admissions data: Statistical tests for causal hypotheses.** *Proceedings of the Forty-First Conference on Uncertainty in Artificial Intelligence*.

Bonet, B. (2001).**Instrumentality tests revisited.** *Proceedings of the 17th Conference in Uncertainty in Artificial Intelligence*, 48–55. https://arxiv.org/abs/1301.2258

Brown, L. D., Cai, T., & DasGupta, A. (2001).**Interval estimation for a binomial proportion.** *Statistical science*, 101–117.

Clopper, C. J., & Pearson, E. S. (1934).**The use of confidence or fiducial limits illustrated in the case of the binomial.** *Biometrika*, 404–413.

Guo, F. R., & Richardson, T. S. (2021).**Chernoff-type concentration of empirical probabilities in relative entropy.** *IEEE Transactions on Information Theory*, *67*, 549–558. https://api.semanticscholar.org/CorpusID:213004900

Imbens, G. W., & Angrist, J. D. (1994).**Identification and estimation of local average treatment effects.** *Econometrica*, *62*(2), 467–475. Retrieved May 14, 2025, from http://www.jstor.org/stable/2951620

Kédagni, D., & Mourifié, I. (2020).**Generalized instrumental inequalities: testing the instrumental variable independence assumption.** *Biometrika*, *107*(3), 661–675. https://doi.org/10.1093/biomet/asaa003

Kennedy, E. H., Balakrishnan, S., & G'Sell, M. (2020).**Sharp instruments for classifying compliers and generalizing causal effects.** *The Annals of Statistics*, *48*(4), 2008–2030. https://doi.org/10.1214/19-AOS1874

Koperberg, T. (2024).**Couplings and matchings: Combinatorial notes on Strassen's theorem.** *Statistics & Probability Letters*, *209*, 110089. https://doi.org/https://doi.org/10.1016/j.spl.2024.110089

Luo, Y., & Wang, H. (2017).**Core determining class and inequality selection.** *The American Economic Review*, *107*(5), 274–77. https://doi.org/10.1257/aer.p20171041

Pearl, J. (1995).**On the testability of causal models with latent and instrumental variables.** *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, 435–443.

References iv

Richardson, T. S., & Robins, J. M. (2014).**ACE bounds; SEMs with equilibrium conditions.** *Statistical Science*, *29*(3), 363–366. Retrieved May 3, 2024, from http://www.jstor.org/stable/43288513

Russell, T. M. (2021).**Sharp bounds on functionals of the joint distribution in the analysis of treatment effects.** *Journal of Business & Economic Statistics*, *39*(2), 532–546. https://doi.org/10.1080/07350015.2019.1684300

Sherman, L. W., & Berk, R. A. (1984). **The Minneapolis domestic violence experiment (tech. rep.).** (Report). National Policing Institute. Washington, DC. https://www.policinginstitute.org/publication/the-minneapolis-domestic-violence-experiment/

Strassen, V. (1965).**The existence of probability measures with given marginals.** *The Annals of Mathematical Statistics*, *36*(2), 423–439. https://doi.org/10.1214/aoms/1177700153

References v

Swanson, S. A., Hernán, M. A., Miller, M., Robins, J. M., & Richardson, T. S. (2018).**Partial identification of the average treatment effect using instrumental variables: Review of methods for binary instruments, treatments, and outcomes [PMID: 31537952].** *Journal of the American Statistical Association*, *113*(522), 933–947. https://doi.org/10.1080/01621459.2018.1434530

Wang, L., Robins, J. M., & Richardson, T. S. (2017).**On falsification of the binary instrumental variable model.** *Biometrika*, *104*(1), 229–236.

Exclusion restriction, versions of

(V1) Individual-level Exclusion

$$Y(x_i, z) = Y(x_i, \widetilde{z}) \text{ for all } z, \widetilde{z} \in [Q],\ i \in [K], \text{ and } q \in [Q]$$

(V2) Joint Stochastic Exclusion

$$P(Y(x_1, z) = y^1, \ldots, Y(x_K, z) = y^K) = P(Y(x_1, \widetilde{z}) = y^1, \ldots, Y(x_K, \widetilde{z}) = y^K)$$
$$\text{for all } z, \widetilde{z} \in [Q] \text{ and } y^1, \ldots, y^K \in [M]$$

(V3) Latent Exclusion

$$P(Y(x, z) = y \mid U = u) = P(Y(x, \widetilde{z}) = y \mid U = u) \text{ for all } z, \widetilde{z} \in [Q],$$
$$x \in [K] \text{ and } y \in [M] \text{ and latent state } u.$$

Independence assumption, versions of

(V1) Random assignment

$$Z \perp\!\!\!\perp (Y(x,z), X(z) : x \in [K], z \in [Q])$$

(V2) Joint independence

$$Z \perp\!\!\!\perp (Y(x,z) : x \in [K], z \in [Q])$$

(V3) Single-world independence

$$Z \perp\!\!\!\perp X(z),\ Y(x,z), \quad \text{for all } z \in [Q],\ x \in [K]$$

(V4) Latent-variable exogeneity There exists U such that $U \perp\!\!\!\perp Z$, and

$$Y(x,z) \perp\!\!\!\perp X, Z \mid U, \quad \text{for all } z \in [Q],\ x \in [K]$$