Causal design and urban policy evaluation: A very brief introduction

Nick Dorward

University of Bristol

December 7, 2022

nicholas.dorward@bristol.ac.uk

The case for causality

- To answer many urban policy questions we may need to adopt a set of tools that allow us to make **causal** rather than **statistical** (or predictive) inferences

The case for causality

- To answer many urban policy questions we may need to adopt a set of tools that allow us to make **causal** rather than **statistical** (or predictive) inferences
- For example:
  - Crime - Does intensive 'hot spot' policing work?
  - Health - Do "clean air zones" reduce asthma incidence?
  - Inequality - Does improving neighbourhood quality change health and economic outcomes?

The case for causality

- To answer many urban policy questions we may need to adopt a set of tools that allow us to make **causal** rather than **statistical** (or predictive) inferences

- For example:
  - Crime - Does intensive 'hot spot' policing work?
  - Health - Do "clean air zones" reduce asthma incidence?
  - Inequality - Does improving neighbourhood quality change health and economic outcomes?

- To answer many urban policy questions we may need to adopt a set of tools that allow us to make **causal** rather than **statistical** (or predictive) inferences

What is causality?

- “We may define a cause to be an object, followed by another, and where all the objects similar to the first are followed by objects similar to the second. **Or in other words where, if the first object had not been, the second never had existed**." (David Hume, 1748)

What is causality?

- “We may define a cause to be an object, followed by another, and where all the objects similar to the first are followed by objects similar to the second. **Or in other words where, if the first object had not been, the second never had existed**." (David Hume, 1748)

- “Causation is something that makes a difference, and the difference it makes must be a difference from what would have happened without it” (David Lewis, 1973)

What is causality?

- “We may define a cause to be an object, followed by another, and where all the objects similar to the first are followed by objects similar to the second. **Or in other words where, if the first object had not been, the second never had existed**." (David Hume, 1748)

- “Causation is something that makes a difference, and the difference it makes must be a difference from what would have happened without it” (David Lewis, 1973)

- **Key idea -** ***the counterfactual***. Alternative possibilities that we imagine in thought experiments to unpick causality.

Prediction vs causal inference

Prediction

- Detect patterns in data and fit functional relationships between variables with accuracy.

Prediction vs causal inference

Prediction

- Detect patterns in data and fit functional relationships between variables with accuracy.
- E.g., ‘the spread of an infectious disease?’, ‘how many people will use a service?’

Prediction vs causal inference

Prediction

- Detect patterns in data and fit functional relationships between variables with accuracy.
- E.g., ‘the spread of an infectious disease?’, ‘how many people will use a service?’
- Not a prediction of the effect that a **specific** choice or decision will have on an outcome

Prediction vs causal inference (cont'd)

Causal inference

- Prediction of a counterfactual associated with a particular decision or path taken

Prediction vs causal inference (cont'd)

Causal inference

- Prediction of a counterfactual associated with a particular decision or path taken

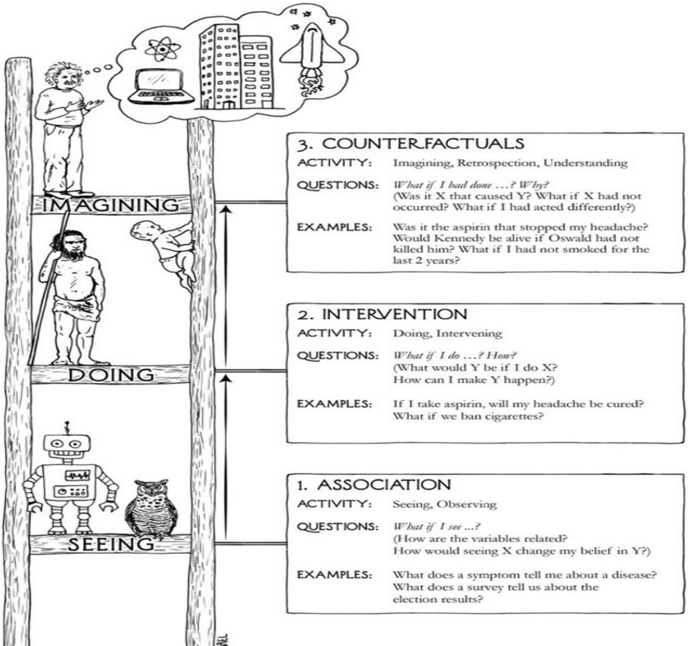
- Causal inference takes a predicted counterfactual and constructs a causal effect which, we hope, tells us something about the state of the future world in the event we make a specific choice.

Prediction vs causal inference (cont'd)

Causal inference

- Prediction of a counterfactual associated with a particular decision or path taken
- Causal inference takes a predicted counterfactual and constructs a causal effect which, we hope, tells us something about the state of the future world in the event we make a specific choice.
- Key for policy applications. We know not only the past, but the future

The ladder of causation



Heuristic vs Analytic

- **Heuristic:** Counterfactual thinking to formalise assumptions about processes. If this theory were true, what would prove me wrong?

Heuristic vs Analytic

- **Heuristic:** Counterfactual thinking to formalise assumptions about processes. If this theory were true, what would prove me wrong?

- **Analytic:** Counterfactuals to make causal inferences. Estimate the true *effect* of an intervention on an outcome/process.

Heuristic vs Analytic

- **Heuristic:** Counterfactual thinking to formalise assumptions about processes. If this theory were true, what would prove me wrong?
- **Analytic:** Counterfactuals to make causal inferences. Estimate the true *effect* of an intervention on an outcome/process.
- **Hybrid:** Counterfactuals as a system of thinking, design, and analysis to make *more credible* claims about causal relationships.

Potential outcomes

- The potential outcomes framework expresses causality in terms of counterfactuals.

Potential outcomes

- The potential outcomes framework expresses causality in terms of counterfactuals.

- A causal effect, let's call this δ_i, is defined as a comparison between two states of the world:
  - One in which a unit, i, receives an intervention - the actual state, let's call this Y_i^1.
  - One in which the unit does not receive the intervention - the counterfactual state, Y_i^0.

Potential outcomes

- The potential outcomes framework expresses causality in terms of counterfactuals.
- A causal effect, let's call this δ_i, is defined as a comparison between two states of the world:
  - One in which a unit, i, receives an intervention - the actual state, let's call this Y_i^1.
  - One in which the unit does not receive the intervention - the counterfactual state, Y_i^0.
- The individual causal (or treatment) effect of the intervention is the simple difference in outcomes (SDO) between the world in which the intervention occurs compared to the one where it does not:

$$\delta_i = Y_i^1 - Y_i^0 \tag{1}$$

The fundamental problem in causal inference

- However, we cannot observe both potential outcomes because counterfactuals are hypothetical and do not exist in reality - we only observe Y_i empirically.

The fundamental problem in causal inference

- However, we cannot observe both potential outcomes because counterfactuals are hypothetical and do not exist in reality - we only observe Y_i empirically.

- If both potential outcomes are required to know the causal effect, then, since it is impossible to observe both Y_i^1 and Y_i^0 for the same individual, δ_i is *unknowable*.

The fundamental problem in causal inference

- However, we cannot observe both potential outcomes because counterfactuals are hypothetical and do not exist in reality - we only observe Y_i empirically.
- If both potential outcomes are required to know the causal effect, then, since it is impossible to observe both Y_i^1 and Y_i^0 for the same individual, δ_i is *unknowable*.
- Causal inference is a missing data problem where we need to make predictions, not of the present or future, but of a missing past.

The fundamental problem in causal inference (cont'd)

- Formalised by the switching equation which states that Yi is a function of its potential outcomes:

The fundamental problem in causal inference (cont'd)

- Formalised by the switching equation which states that Yi is a function of its potential outcomes:

$$Y_i = D_i Y_i^1 + (1 - D_i) Y_i^0 \tag{2}$$

The fundamental problem in causal inference (cont'd)

- Formalised by the switching equation which states that Yi is a function of its potential outcomes:

$$Y_i = D_i Y_i^1 + (1 - D_i) Y_i^0 \tag{2}$$

$$Y_i = \begin{cases} \mathrm{Y}_i^1 \text{ if } D_i = 1, \\ \mathrm{Y}_i^0 \text{ if } D_i = 0, \end{cases} \tag{3}$$

The fundamental problem in causal inference (cont'd)

- Formalised by the switching equation which states that Yi is a function of its potential outcomes:

$$Y_i = D_i Y_i^1 + (1 - D_i) Y_i^0 \tag{2}$$

$$Y_i = \begin{cases} \mathrm{Y}_i^1 \text{ if } D_i = 1, \\ \mathrm{Y}_i^0 \text{ if } D_i = 0, \end{cases} \tag{3}$$

- If $D_i = 1$, then $Y_i = Y_i^1$ because the second term in (2) zeroes out. And if $D_i = 0$, the first term zeroes and $Y_i = Y_i^0$.

Potential outcomes in action

- But we have a distribution of both y_i^1 and y_i^0 in the population. So, we can estimate 'average treatment effects' (ATE) across the population by comparing outcomes for 'treatment' (those with y_i^1) and 'control' (those with y_i^0) groups.

- Average treatment effects are *unknowable* because, according to the switching equation, we don't have both potential outcomes for each observation. But it can be *estimated* from samples of data.

- The simple difference in means between the treatment and control groups will give us the average treatment effect from across the population.

Potential outcomes in action (cont'd)

$$
\begin{aligned}
SDO &= \frac{1}{N_T}\sum_{i=1}^{N}(y_i|d_i=1) - \frac{1}{N_C}\sum_{i=1}^{N}(y_i|d_i=0) \\
&= E[Y_i|D_i=1] - E[Y_i|D_i=0]
\end{aligned}
$$

Potential outcomes in action (cont'd)

$$\begin{aligned} SDO &= \frac{1}{N_T}\sum_{i=1}^{N}(y_i|d_i=1) - \frac{1}{N_C}\sum_{i=1}^{N}(y_i|d_i=0) \\ &= E[Y_i|D_i=1] - E[Y_i|D_i=0] \end{aligned}$$

Which can be decomposed to:

Potential outcomes in action (cont'd)

$$
\begin{aligned}
SDO &= \frac{1}{N_T}\sum_{i=1}^{N}(y_i|d_i=1) - \frac{1}{N_C}\sum_{i=1}^{N}(y_i|d_i=0) \\
&= E[Y_i|D_i=1] - E[Y_i|D_i=0]
\end{aligned}
$$

Which can be decomposed to:

$$
\begin{aligned}
E[Y_i|D_i=1] - E[Y_i|D_i=0] = &\underbrace{E[Y_i^1|D_i=1] - E[Y_i^0|D_i=1]}_{\text{Average treatment effect}} \\
&+ \underbrace{E[Y_i^0|D_i=1] - E[Y_i^0|D_i=0]}_{\text{Selection bias}}.
\end{aligned}
$$

Potential outcomes in action (cont'd)

$$
\begin{aligned}
SDO &= \frac{1}{N_T}\sum_{i=1}^{N}(y_i|d_i=1) - \frac{1}{N_C}\sum_{i=1}^{N}(y_i|d_i=0)\\
&= E[Y_i|D_i=1] - E[Y_i|D_i=0]
\end{aligned}
$$

Which can be decomposed to:

$$
\begin{aligned}
E[Y_i|D_i=1] - E[Y_i|D_i=0] = &\underbrace{E[Y_i^1|D_i=1] - E[Y_i^0|D_i=1]}_{\text{Average treatment effect}}\\
&+ \underbrace{E[Y_i^0|D_i=1] - E[Y_i^0|D_i=0]}_{\text{Selection bias}}.
\end{aligned}
$$

Selection bias: the difference between treatment and control groups with no intervention.

Randomisation solves selection bias

- Randomisation of the assignment of D_i solves the selection problem because it ensures independence of potential outcomes:

Randomisation solves selection bias

- Randomisation of the assignment of D_i solves the selection problem because it ensures independence of potential outcomes:

$$
\begin{aligned}
E[Y_i|D_i = 1] - E[Y_i|D_i = 0] &= E[Y_i^1|D_i = 1] - E[Y_i^0|D_i = 0] \\
&= E[Y_i^1|D_i = 1] - E[Y_i^0|D_i = 1] \\
&= E[Y_i^1 - Y_i^0|D_i = 1] \\
&= E[Y_i^1 - Y_i^0].
\end{aligned}
$$

Randomisation solves selection bias

- Randomisation of the assignment of D_i solves the selection problem because it ensures independence of potential outcomes:

$$
\begin{aligned}
E[Y_i|D_i=1]-E[Y_i|D_i=0] &= E[Y_i^1|D_i=1]-E[Y_i^0|D_i=0] \\
&= E[Y_i^1|D_i=1]-E[Y_i^0|D_i=1] \\
&= E[Y_i^1-Y_i^0|D_i=1] \\
&= E[Y_i^1-Y_i^0].
\end{aligned}
$$

- Independence of Y_i^0 and D_i allows us to swap in $E[Y_i^1|D_i=1]$ in for $E[Y_0^1|D_i=1]$ in line 2 because because potential outcomes for Y_i^0 and Y_i^1 are the same.

Important assumptions

- Independence assumption - Assignment to treatment and control group is independent of potential outcomes: $E[Y_i^0|D=1] = E[Y_i^0|D=0]$.

Important assumptions

- Independence assumption - Assignment to treatment and control group is independent of potential outcomes: $E[Y_i^0|D=1] = E[Y_i^0|D=0]$.
- Stable Unit Treatment Value Assumption (SUTVA):

Important assumptions

- Independence assumption - Assignment to treatment and control group is independent of potential outcomes: $E[Y_i^0|D=1] = E[Y_i^0|D=0]$.
- Stable Unit Treatment Value Assumption (SUTVA):
  - Homogeneous treatment - the level (or dosage) of the treatment is homogeneous across groups.

Important assumptions

- Independence assumption - Assignment to treatment and control group is independent of potential outcomes: $E[Y_i^0|D=1] = E[Y_i^0|D=0]$.
- Stable Unit Treatment Value Assumption (SUTVA):
  - Homogeneous treatment - the level (or dosage) of the treatment is homogeneous across groups.
  - Non-interference - no externalities or spillover from treatment. Treatment status of unit i does not affect potential outcomes of unit j (e.g., (a)spatial networks).

Summary of potential outcomes

- Counterfactual thinking can help us imagine two potential outcomes for worlds in which a unit did or did not receive an intervention.

Summary of potential outcomes

- Counterfactual thinking can help us imagine two potential outcomes for worlds in which a unit did or did not receive an intervention.
- Fundamental problem in causal inference - we live in reality and cannot observe both potential outcomes.

Summary of potential outcomes

- Counterfactual thinking can help us imagine two potential outcomes for worlds in which a unit did or did not receive an intervention.
- Fundamental problem in causal inference - we live in reality and cannot observe both potential outcomes.
- We can make population wide comparisons across treatment and control groups to estimate average treatment effects.

Summary of potential outcomes

- Counterfactual thinking can help us imagine two potential outcomes for worlds in which a unit did or did not receive an intervention.
- Fundamental problem in causal inference - we live in reality and cannot observe both potential outcomes.
- We can make population wide comparisons across treatment and control groups to estimate average treatment effects.
- But selection bias reflects endogenous sorting into treatment and control.

Summary of potential outcomes

- Counterfactual thinking can help us imagine two potential outcomes for worlds in which a unit did or did not receive an intervention.
- Fundamental problem in causal inference - we live in reality and cannot observe both potential outcomes.
- We can make population wide comparisons across treatment and control groups to estimate average treatment effects.
- But selection bias reflects endogenous sorting into treatment and control.
- Randomisation solves the selection problem (under certain assumptions)

Summary of potential outcomes

- Counterfactual thinking can help us imagine two potential outcomes for worlds in which a unit did or did not receive an intervention.
- Fundamental problem in causal inference - we live in reality and cannot observe both potential outcomes.
- We can make population wide comparisons across treatment and control groups to estimate average treatment effects.
- But selection bias reflects endogenous sorting into treatment and control.
- Randomisation solves the selection problem (under certain assumptions)
- To make causal inferences we need random assignment of interventions or to be able to simulate randomness in some plausible way.

RCTs and (quasi)experiments

- **Experiments and RCTs** - explicitly randomise a policy intervention across treatment and control groups - 'balanced' on unobservables.

- **Natural experiments** - leverage arbitrary divergences in laws, policies, or practices to analyse the effects of an intervention on a population as is if they had been part of an experiment. Looks at differences across treatment and control groups 'as if' intervention was randomly assigned - regression discontinuity.

Moving to Opportunity (MTO) experiment:

- Offered low-income families from deprived urban neighbourhoods the opportunity to move to less ‘distressed’ areas.

Moving to Opportunity (MTO) experiment:

- Offered low-income families from deprived urban neighbourhoods the opportunity to move to less ‘distressed’ areas.
- Families randomly assigned to two treatment groups (where they got housing vouchers to move) and a control group.

Moving to Opportunity (MTO) experiment:

- Offered low-income families from deprived urban neighbourhoods the opportunity to move to less ‘distressed’ areas.
- Families randomly assigned to two treatment groups (where they got housing vouchers to move) and a control group.
- Families surveyed 10-15 years later to evaluate effects on adults and children.

Moving to Opportunity (MTO) experiment:

- Offered low-income families from deprived urban neighbourhoods the opportunity to move to less ‘distressed’ areas.
- Families randomly assigned to two treatment groups (where they got housing vouchers to move) and a control group.
- Families surveyed 10-15 years later to evaluate effects on adults and children.
- Treated adults had better physical and mental health outcomes but no effect on economic outcomes

Moving to Opportunity (MTO) experiment:

- Offered low-income families from deprived urban neighbourhoods the opportunity to move to less ‘distressed’ areas.
- Families randomly assigned to two treatment groups (where they got housing vouchers to move) and a control group.
- Families surveyed 10-15 years later to evaluate effects on adults and children.
- Treated adults had better physical and mental health outcomes but no effect on economic outcomes
- No detectable effect on education outcomes and no improvement to health outcomes expect improved mental health among girls.

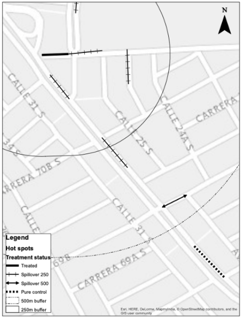
Moving to Opportunity (MTO) experiment:

- Offered low-income families from deprived urban neighbourhoods the opportunity to move to less 'distressed' areas.
- Families randomly assigned to two treatment groups (where they got housing vouchers to move) and a control group.
- Families surveyed 10-15 years later to evaluate effects on adults and children.
- Treated adults had better physical and mental health outcomes but no effect on economic outcomes
- No detectable effect on education outcomes and no improvement to health outcomes expect improved mental health among girls.
- Limitations include: non-compliance (50%), disruption of moving, non-random selection into destination neighbourhoods.

Policing in Bogata (Blattman et al., 2021)

- City randomly reallocated existing police and municipal resources across experimental street blocks.

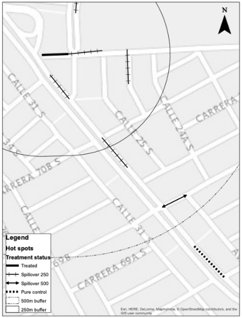
Figure 2: An example of assignment to the four treatment conditions



Policing in Bogata (Blattman et al., 2021)

- City randomly reallocated existing police and municipal resources across experimental street blocks.
- Streets randomly assigned to one of four treatment statuses: intensive policing, municipal services, both, or neither.

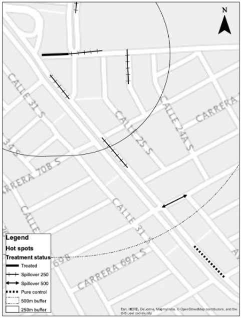
Figure 2: An example of assignment to the four treatment conditions



Policing in Bogata (Blattman et al., 2021)

- City randomly reallocated existing police and municipal resources across experimental street blocks.
- Streets randomly assigned to one of four treatment statuses: intensive policing, municipal services, both, or neither.
- Measure spatial spillover (interference) into non-experimental streets

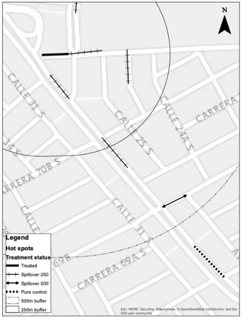
Figure 2: An example of assignment to the four treatment conditions



Policing in Bogata (Blattman et al., 2021)

- City randomly reallocated existing police and municipal resources across experimental street blocks.
- Streets randomly assigned to one of four treatment statuses: intensive policing, municipal services, both, or neither.
- Measure spatial spillover (interference) into non-experimental streets
- No overall decline in crime from policing or services, differences across violent and property crimes, policing treatment displaced crime into non-treatment streets.

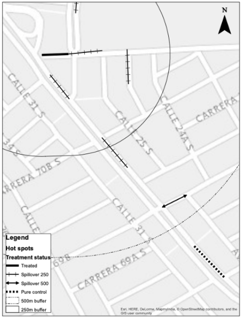
Figure 2: An example of assignment to the four treatment conditions



Policing in Bogata (Blattman et al., 2021)

- City randomly reallocated existing police and municipal resources across experimental street blocks.
- Streets randomly assigned to one of four treatment statuses: intensive policing, municipal services, both, or neither.
- Measure spatial spillover (interference) into non-experimental streets
- No overall decline in crime from policing or services, differences across violent and property crimes, policing treatment displaced crime into non-treatment streets.
- Many results not statistically significant, but may be substantively meaningful to policy makers.

Figure 2: An example of assignment to the four treatment conditions



Part 2. The causal inference 'tool box'

- Causal inference without explicit randomisation.

Part 2. The causal inference 'tool box'

- Causal inference without explicit randomisation.

- Toolbox of post-assignment corrections that leverage 'as if' random variation in interventions to recover causal parameters:
  - Controls, matching, & fixed-effects
  - Difference-in-differences
  - Regression discontinuity
  - Instrumental variables

Backdoor criterion

$$\begin{tikzcd}
& U \arrow[dl, dashed] \arrow[dr, dashed] & \\
T \arrow[rr] & & Y
\end{tikzcd}$$

Figure: Unobservable U

$$\begin{tikzcd}
& X \arrow[dl] \arrow[dr] & \\
T \arrow[rr] & & Y
\end{tikzcd}$$

Figure: Observable X

- Consider the standard regression equation:

Backdoor criterion

$$\begin{tikzcd}
& U \arrow[dl, dashed] \arrow[dr, dashed] & \\
T \arrow[rr] & & Y
\end{tikzcd}$$

Figure: Unobservable U

$$\begin{tikzcd}
& X \arrow[dl] \arrow[dr] & \\
T \arrow[rr] & & Y
\end{tikzcd}$$

Figure: Observable X

- Consider the standard regression equation:

$$y_i = T_i \beta_i + \color{red}{X_i \delta_i} + U_i + \epsilon_i$$

Backdoor criterion

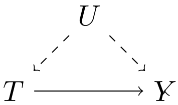


Figure: Unobservable U

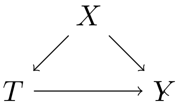


Figure: Observable X

- Consider the standard regression equation:

$$y_i = T_i\beta_i + X_i\delta_i + U_i + \epsilon_i$$

- where $T_i\beta_i$ is the treatment, $X_i\delta_i$ are observed controls, and U_i are unobserved confounders.

Selection on observables: Controls

- Two primary reasons for including controls, $X_i\delta_i$:

Selection on observables: Controls

- Two primary reasons for including controls, $X_i\delta_i$:
  - Conditional independence assumption - assignment of an intervention is random conditional on some observable(s) (e.g., gender, school, neighbourhood).

Selection on observables: Controls

- Two primary reasons for including controls, $X_i\delta_i$:
  - Conditional independence assumption - assignment of an intervention is random conditional on some observable(s) (e.g., gender, school, neighbourhood).
    - In reality T is likely to be correlated with unobservables, U

Selection on observables: Controls

- Two primary reasons for including controls, $X_i \delta_i$:
  - Conditional independence assumption - assignment of an intervention is random conditional on some observable(s) (e.g., gender, school, neighbourhood).
    - In reality T is likely to be correlated with unobservables, U
    - Can't be sure that all relevant X are accounted for - unobservables

Selection on observables: Controls

- Two primary reasons for including controls, $X_i\delta_i$:
  - Conditional independence assumption - assignment of an intervention is random conditional on some observable(s) (e.g., gender, school, neighbourhood).
    - In reality T is likely to be correlated with unobservables, U
    - Can't be sure that all relevant X are accounted for - unobservables
  - Precision - Even if not related to the assignment probability, including controls that are related to the outcome will reduce residual variance increasing precision of estimates.

Selection on observables: Matching

- Alternative to regression that closes back doors between intervention and outcome

Selection on observables: Matching

- Alternative to regression that closes back doors between intervention and outcome
- Constructs comparison groups that are similar along a set of *observed* matching variables using weights.

Selection on observables: Matching

- Alternative to regression that closes back doors between intervention and outcome
- Constructs comparison groups that are similar along a set of *observed* matching variables using weights.
- In many ways similar to regression.

Selection on observables: Matching

- Alternative to regression that closes back doors between intervention and outcome
- Constructs comparison groups that are similar along a set of *observed* matching variables using weights.
- In many ways similar to regression.
- Uses a different set of assumptions and is less model dependent than regression.

Selection on observables: Matching

- Alternative to regression that closes back doors between intervention and outcome
- Constructs comparison groups that are similar along a set of *observed* matching variables using weights.
- In many ways similar to regression.
- Uses a different set of assumptions and is less model dependent than regression.
- Suffers from the same fatal flaw - at least when it comes to estimating causal effects - of assuming that our set of observed variables are enough to close all back doors.

Selection on observables: Fixed effects

- An approach to controlling for all unobserved confounders that are fixed for some category or context.

Selection on observables: Fixed effects

- An approach to controlling for all unobserved confounders that are fixed for some category or context.
- Controls for the individual unit (people, firm, neighbourhood, city, country ...)
  - Include a dummy variable for that 'higher-level' unit.

Selection on observables: Fixed effects

- An approach to controlling for all unobserved confounders that are fixed for some category or context.
- Controls for the individual unit (people, firm, neighbourhood, city, country ...)
  - Include a dummy variable for that 'higher-level' unit.
- Typically used in context of cross-sectional time-series data.
  - Removing variation between units focusing upon within unit variation over time.

Selection on observables: Fixed effects

- An approach to controlling for all unobserved confounders that are fixed for some category or context.
- Controls for the individual unit (people, firm, neighbourhood, city, country ...)
  - Include a dummy variable for that 'higher-level' unit.
- Typically used in context of cross-sectional time-series data.
  - Removing variation between units focusing upon within unit variation over time.
- Can be extended to multiple fixed effects and time as well as geography (TWFE).

Difference-in-Differences Design

- Extension of TWFE logic - typically applied when we have treatment and control groups measured across at least two time periods.

Difference-in-Differences Design

- Extension of TWFE logic - typically applied when we have treatment and control groups measured across at least two time periods.
- Removes time-invariant components of unobservables, U, that are common to treatment and control groups.

Difference-in-Differences Design

- Extension of TWFE logic - typically applied when we have treatment and control groups measured across at least two time periods.
- Removes time-invariant components of unobservables, U, that are common to treatment and control groups.
- Primary identifying assumption - parallel trends.
  - No time-varying differences in unobservables between treatment and control groups.

Difference-in-Differences Design

- Extension of TWFE logic - typically applied when we have treatment and control groups measured across at least two time periods.
- Removes time-invariant components of unobservables, U, that are common to treatment and control groups.
- Primary identifying assumption - parallel trends.
  - No time-varying differences in unobservables between treatment and control groups.
- Commonly applied to natural experiments where some areas receive intervention by chance.

John Snow's cholera study (1855)

John Snow's cholera study (cont'd)

- Years of studying cholera epidemics in the 19th C led Snow to question the prevailing miasma hypothesis

John Snow's cholera study (cont'd)

- Years of studying cholera epidemics in the 19th C led Snow to question the prevailing miasma hypothesis
- Developed alternative hypothesis that disease caused by contaminated drinking water

John Snow's cholera study (cont'd)

- Years of studying cholera epidemics in the 19th C led Snow to question the prevailing miasma hypothesis
- Developed alternative hypothesis that disease caused by contaminated drinking water
- Snow leveraged a 'natural experiment' that allocated clean drinking water "as if" by random

John Snow's cholera study (cont'd)

- Years of studying cholera epidemics in the 19th C led Snow to question the prevailing miasma hypothesis
- Developed alternative hypothesis that disease caused by contaminated drinking water
- Snow leveraged a 'natural experiment' that allocated clean drinking water "as if" by random
- Lambeth water company moved its intake pipes upstream beyond main sewage discharge point - uncontaminated water.

John Snow's cholera study (cont'd)

- Years of studying cholera epidemics in the 19th C led Snow to question the prevailing miasma hypothesis
- Developed alternative hypothesis that disease caused by contaminated drinking water
- Snow leveraged a 'natural experiment' that allocated clean drinking water "as if" by random
- Lambeth water company moved its intake pipes upstream beyond main sewage discharge point - uncontaminated water.
- Southwark and Vauxhall water company did not - contaminated water.

John Snow's cholera study (cont'd)

- Years of studying cholera epidemics in the 19th C led Snow to question the prevailing miasma hypothesis
- Developed alternative hypothesis that disease caused by contaminated drinking water
- Snow leveraged a 'natural experiment' that allocated clean drinking water "as if" by random
- Lambeth water company moved its intake pipes upstream beyond main sewage discharge point - uncontaminated water.
- Southwark and Vauxhall water company did not - contaminated water.
- Proved that both companies served similar households within the same neighbourhoods (i.e., balance on covariates).

John Snow's cholera study (cont'd)

- Years of studying cholera epidemics in the 19th C led Snow to question the prevailing miasma hypothesis
- Developed alternative hypothesis that disease caused by contaminated drinking water
- Snow leveraged a 'natural experiment' that allocated clean drinking water "as if" by random
- Lambeth water company moved its intake pipes upstream beyond main sewage discharge point - uncontaminated water.
- Southwark and Vauxhall water company did not - contaminated water.
- Proved that both companies served similar households within the same neighbourhoods (i.e., balance on covariates).
- Interpret effect of clean water while holding confounders - hygiene, poverty, neighbourhood - constant.

How diff-in-diff works

Table: Snow's data

Company name	**1849**	**1854**
Southwark and Vauxhall	135	147
Lambeth	85	19

1) First difference - difference in Lambeth and S&V outcomes in 1854.

- Selection bias

How diff-in-diff works

Table: Snow's data

Company name	**1849**	**1854**
Southwark and Vauxhall	135	147
Lambeth	85	19

1) First difference - difference in Lambeth and S&V outcomes in 1854.

- Selection bias

2) Second difference - compare Lambeth before and after intervention.

- Time trends.

How diff-in-diff works

Table: Snow's data

Company name	**1849**	**1854**
Southwark and Vauxhall	135	147
Lambeth	85	19

1) First difference - difference in Lambeth and S&V outcomes in 1854.
   - Selection bias

2) Second difference - compare Lambeth before and after intervention.
   - Time trends.

3) Diff-in-diff - combine differences to eliminate selection bias and time trend
   - Parallel trends - difference between treated and untreated units the same pre- and post-treatment without intervention.

How diff-in-diff works (cont'd)

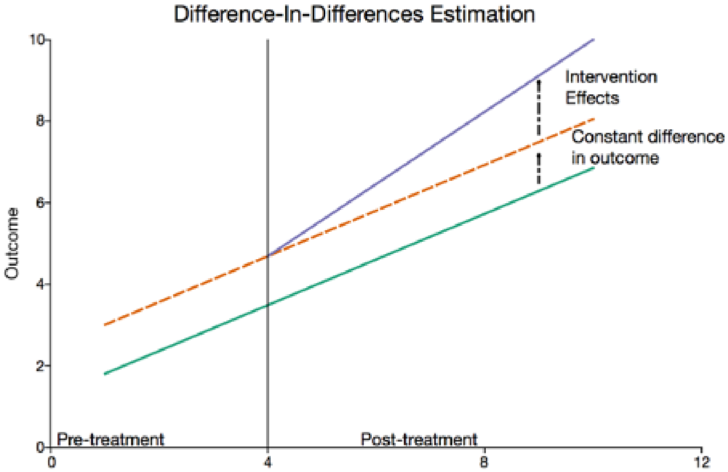


Figure: difference-in-differences estimator

How diff-in-diff works (cont'd)

- Diff-in-diff estimator:

  $$ATT = (E[Y_k|Post] - E[Y_k|Pre]) - (E[Y_U|Post] - E[Y_U|Pre])$$

How diff-in-diff works (cont'd)

- Diff-in-diff estimator:

  $$ATT = (E[Y_k|Post] - E[Y_k|Pre]) - (E[Y_U|Post] - E[Y_U|Pre])$$

- If we plug in Snow's data:

$$ATT = (19 - 85) - (147 - 135)$$

How diff-in-diff works (cont'd)

- Diff-in-diff estimator:

$$ATT = (E[Y_k|Post] - E[Y_k|Pre]) - (E[Y_U|Post] - E[Y_U|Pre])$$

- If we plug in Snow's data:

$$ATT = (19 - 85) - (147 - 135)$$

- This generalises to:
  - Multiple cross-sectional units
  - Multiple temporal units
  - Treatment in multiple periods

Diff-in-diff: summary

- Compare changes in outcomes pre- and post-intervention using a naturally occurring control and treatment group.

Diff-in-diff: summary

- Compare changes in outcomes pre- and post-intervention using a naturally occurring control and treatment group.
- Difference out unobervables across units and time.

Diff-in-diff: summary

- Compare changes in outcomes pre- and post-intervention using a naturally occurring control and treatment group.
- Difference out unobervables across units and time.
- Key assumption - parallel trends. Based on a counterfactual that cannot be empirically validated.

Diff-in-diff: summary

- Compare changes in outcomes pre- and post-intervention using a naturally occurring control and treatment group.
- Difference out unobervables across units and time.
- Key assumption - parallel trends. Based on a counterfactual that cannot be empirically validated.
- John Snow example shows these assumptions rest on deep empirical and contextual knowledge of the problem.

Regression Discontinuity Design

- Regression discontinuity designs (RDD) leverage interventions that are assigned at a cutoff or threshold.

Regression Discontinuity Design

- Regression discontinuity designs (RDD) leverage interventions that are assigned at a cutoff or threshold.
- Units on one side of the threshold get the intervention while those on the other do not.

Regression Discontinuity Design

- Regression discontinuity designs (RDD) leverage interventions that are assigned at a cutoff or threshold.
- Units on one side of the threshold get the intervention while those on the other do not.
- The idea behind RDD is that units either side of the threshold should be very similar in terms of their observables and, within this subpopulation, treatments is assigned as if by random.

Regression Discontinuity Design

- Regression discontinuity designs (RDD) leverage interventions that are assigned at a cutoff or threshold.
- Units on one side of the threshold get the intervention while those on the other do not.
- The idea behind RDD is that units either side of the threshold should be very similar in terms of their observables and, within this subpopulation, treatments is assigned as if by random.
- We can estimate the causal effect of the intervention by comparing the sub-population of units around the threshold.

Regression Discontinuity Design

- Regression discontinuity designs (RDD) leverage interventions that are assigned at a cutoff or threshold.
- Units on one side of the threshold get the intervention while those on the other do not.
- The idea behind RDD is that units either side of the threshold should be very similar in terms of their observables and, within this subpopulation, treatments is assigned as if by random.
- We can estimate the causal effect of the intervention by comparing the sub-population of units around the threshold.
- Seen by many as the 'gold standard' in causal inference with observational data.

Key terminology

- **Running variable** - determines treatment status of unit.

Key terminology

- **Running variable** - determines treatment status of unit.
- **Cutoff/threshold** - the specific value long the running variable at which treatment is assigned.

Key terminology

- **Running variable** - determines treatment status of unit.
- **Cutoff/threshold** - the specific value long the running variable at which treatment is assigned.
- **Bandwidth** - Everything is related to everything else but those things closer to the cutoff are more similar than things farther from the cutoff. The bandwidth determines how close to the cutoff we look to make our comparison.

Doing RDD

- Account for how the running variable normally affects the outcome.

Doing RDD

- Account for how the running variable normally affects the outcome.
- Choose a method for estimating the outside either side of the cutoff (i.e., OLS, LOESS).

Doing RDD

- Account for how the running variable normally affects the outcome.
- Choose a method for estimating the outside either side of the cutoff (i.e., OLS, LOESS).
- Select a bandwidth (i.e., Gaussian kernal, IDW).

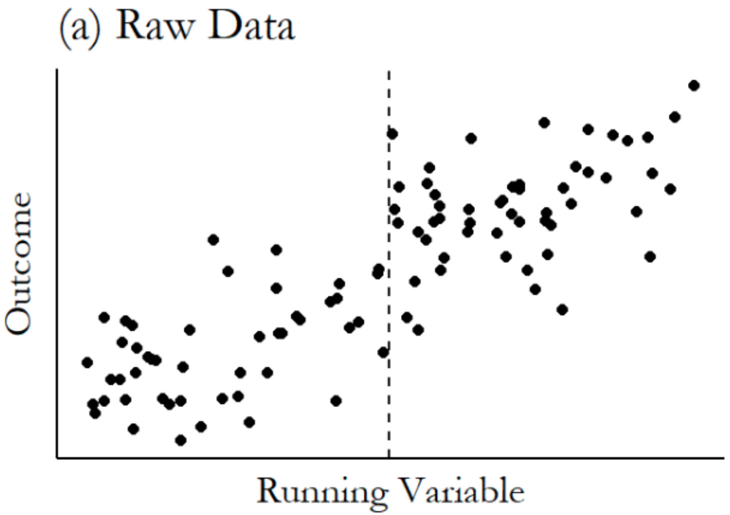
Doing RDD

- Account for how the running variable normally affects the outcome.
- Choose a method for estimating the outside either side of the cutoff (i.e., OLS, LOESS).
- Select a bandwidth (i.e., Gaussian kernal, IDW).
- Focus upon observations around the cutoff within the bandwidth.

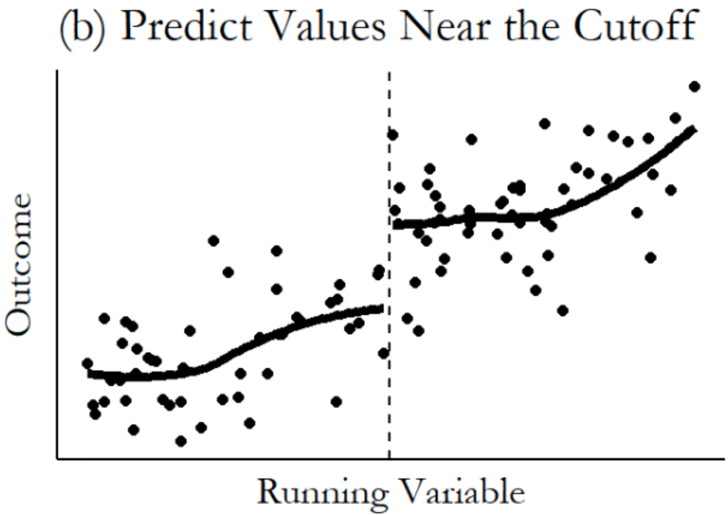
Doing RDD

- Account for how the running variable normally affects the outcome.
- Choose a method for estimating the outside either side of the cutoff (i.e., OLS, LOESS).
- Select a bandwidth (i.e., Gaussian kernal, IDW).
- Focus upon observations around the cutoff within the bandwidth.
- Compare the just-barely treated units against the just-barely untreated units.

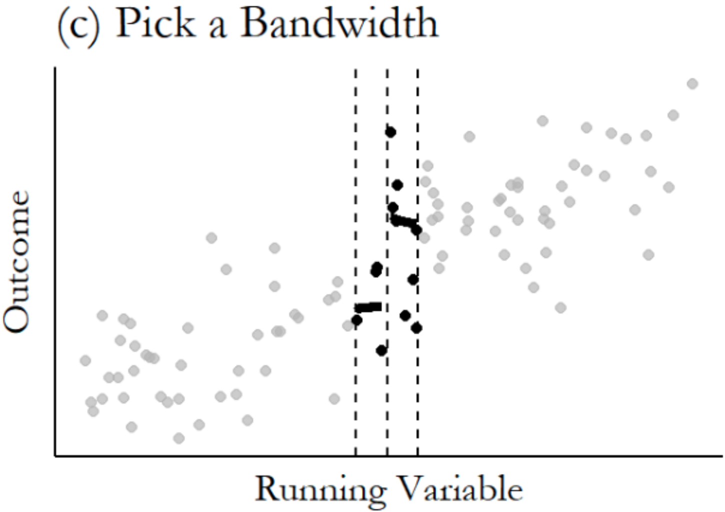
How RDD works



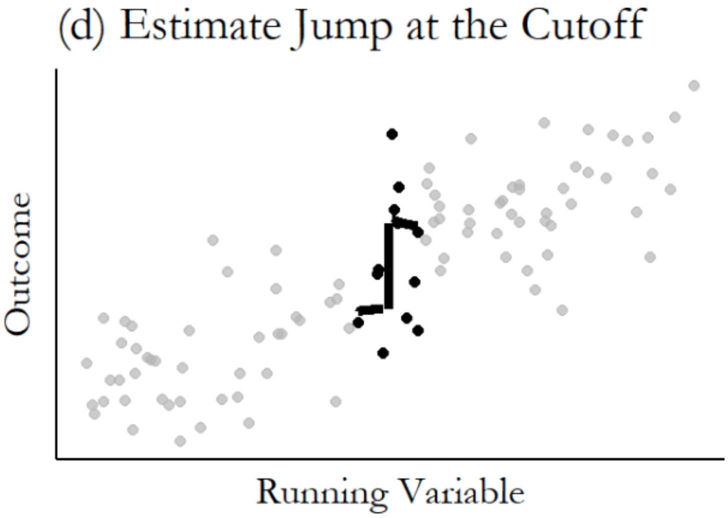
How RDD works (cont'd)



How RDD works (cont'd)

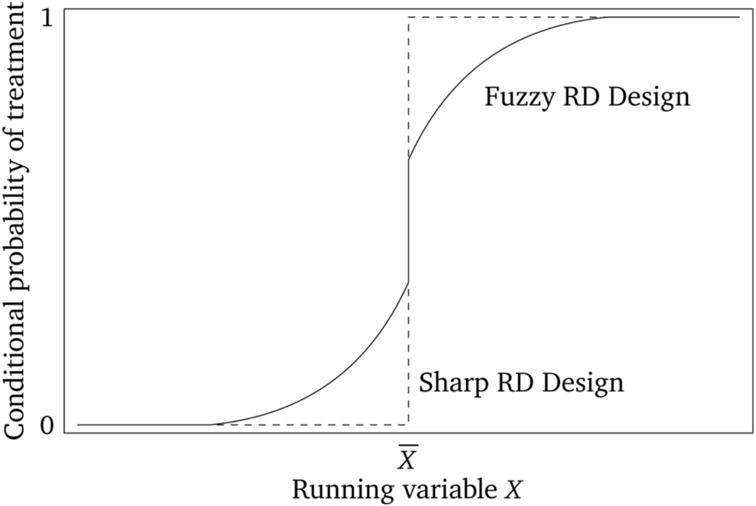


How RDD works (cont'd)



Sharp vs. fuzzy RDD

- In fuzzy RDD the threshold is not discrete and only changes the probability of being assigned an intervention.



Key RDD assumptions

- Continuity assumption: In the absence of the intervention, potential outcomes across treated and untreated groups would not change.

Key RDD assumptions

- Continuity assumption: In the absence of the intervention, potential outcomes across treated and untreated groups would not change.
- In other words, there is no omitted variable at the cutoff. There is nothing else causing the discontinuity.

Key RDD assumptions

- Continuity assumption: In the absence of the intervention, potential outcomes across treated and untreated groups would not change.
- In other words, there is no omitted variable at the cutoff. There is nothing else causing the discontinuity.
- However, this can be violated when:
  - Units can sort their treatment status.
  - Cutoff is endogenous to unobservables that influence the outcome.

Key RDD assumptions

- Continuity assumption: In the absence of the intervention, potential outcomes across treated and untreated groups would not change.
- In other words, there is no omitted variable at the cutoff. There is nothing else causing the discontinuity.
- However, this can be violated when:
  - Units can sort their treatment status.
  - Cutoff is endogenous to unobservables that influence the outcome.
- Analyst must know the assignment rule!

RDD with geographic boundaries

- Geographic borders can act as a discontinuity.
- When one a policy is arbitrarily implemented on one side of a boundary and not the other.
- Places or observational either side are more likely to be similar - comparable on unobservables and potential outcomes.
- Challenges:
  - Sorting - individual can sort across geographic boundaries
  - Interference - aka spatial diffusion/spillover
  - Context - borders not randomly assigned and thus endogenous to outcome and potential outcomes - think gerrymandered districts and political outcomes.

Important ideas not discussed

- Estimands/treatment effects
- Heterogeneous treatment effects
- Instrumental variables (see back of slides)
- Synthetic controls
- Spatial causal inference
- Causal machine learning

Summary (cont'd)

Table 1 Summary of the key analytical methods used to assess health interventions and their relative trade-offs

Analytical method	Description	Advantages	Disadvantages	Trade-offs relative to other methods
Interrupted Time Series (ITS)	A before-after comparison in the level and trend of outcomes pre and post intervention [17, 21, 22]	Straightforward methodological approach without reliance on simplifying assumptions [17, 21, 22]	Influenced by simultaneous events occurring at the time of intervention [17, 21, 22]	No control group to compare intervention effects against a group exposed to the intervention which can bias estimated intervention effects [23]
Difference-in-differences (DiD)	A contrast of outcome changes pre and post intervention using a naturally occurring control group and treatment group subject to the intervention change [18, 24]	Using the intervention itself as a naturally occurring experiment, allows to difference out any exogenous effects from events occurring simultaneously [18, 24]	The parallel trends assumption is based on counter-factual intervention trends which cannot be tested [18, 24]	Use of a naturally occurring control group to compare intervention effects naturally isolates group differences from intervention effects. No statistical test to verify the parallel trends assumption can bias estimated effects [18, 24]
Synthetic Control (SC)	Comparison of treatment effects between a treatment group and a constructed control i.e. a synthetic control using weights similar to treatment outcomes pre-intervention [25, 26]	Can complement other analytical methods particularly when a naturally occurring control group cannot be established and/or when simplification assumptions do not hold e.g. the parallel trends assumption in DiD [25, 26]	Requirement of sufficient data pre and post intervention containing sufficient detail of control weights similar to the treatment group [19]	Can overcome parallel trends assumption required for DiD. Cannot test for similarity of controls used to construct the synthetic control which may bias estimated intervention effects. Heavy data requirement pre and post intervention [19, 25]
Matching	A comparison of outcomes between treatment and control groups pre and post intervention post matching groups with similar observable factors [18, 27]	Reduction of biases within groups is eliminated due to matching [18, 27]	Requirement of sufficient data pre and post intervention for matching similar observable characteristics between treatment and control groups. No statistical means to testing 'similarity' [27]	Heavy data requirement to match similar characteristics. Matching is limited to observable factors and does not account for non-observable factors. 'Similarity' determined using subjective judgment and cannot be statistically measured and can bias estimates [27].
Instrumental Variables (IV)	An observable variable i.e. the instrument is selected to randomise the estimation of treatment effects [18, 20, 28]	Introduction of randomness when estimating treatment effects to reflect similarity to a RCT [18]	Dependence on choosing the most appropriate instrument to satisfy the assumption of no relationship between the outcome and assuming outcome is affected only via intervention exposure [18, 29]	Imposed randomisation using an instrument useful for estimating intervention effects. Randomisation is imposed and not naturally occurring like with DiD and can bias estimated effects [18, 20, 28, 29]

Instrumental variables

A way of identifying causal effect of an intervention by identifying a source of random variation in treatment assignment that is not affected by unobservables.

The instrument, Z, mimics the explicit random assignment of T in RCTs with something that has already randomised T in the real-world.

Use Z to statistically isolate variation in T driven by Z and identify causal effect of T on Y:

- 1) Use Z to explain T
- 2) Remove any part of the T that is not explained by Z
- 3) Use Z to explain tY removing any Y not explained by Z
- 4) Assess relationship between Z-explained part of T and Z-explained part of Y

How IV works

$$\begin{tikzcd}
& U \arrow[dl, dashed] \arrow[dr, dashed] & \\
T \arrow[rr] & & Y
\end{tikzcd}$$

Figure: Endogeneity

$$\begin{tikzcd}
& & U \arrow[dl, dashed] \arrow[dr, dashed] & \\
& T \arrow[rr] & & Y \\
Z \arrow[ur] & & &
\end{tikzcd}$$

Figure: Instrumental variable

$$\begin{tikzcd}
& & U \arrow[dl, dashed] \arrow[dr, dashed] & \\
& T \arrow[rr] & & Y \\
Z \arrow[ur] \arrow[urrr, bend right] & & &
\end{tikzcd}$$

Figure: Exclusion restriction violation

How IV works (cont'd)

Instrument variables estimation - For each Z-explained movement in T, how much Z-explained movement in Y was there?

Actual estimation is comparatively simple.

Most commonly performed via Two-stage least squares (2SLS):

$$T = \gamma_0 + \gamma_1 Z + \gamma_2 W + \upsilon \tag{4}$$

$$Y = \beta_0 + \beta_1 \hat{T} + \beta_2 W + \epsilon \tag{5}$$

Where W are controls, γ are first stage regression coefficients, and $\hat{T}$ are predicted values of T.

Choosing instruments

Credible inference in IV depends upon the choice of IV

A valid instrumental variable must satisfy three key criteria:

- Relevancy: $Cov(Z, Y) \neq 0$. Statistical vs substantive relevancy. Does Z theoretically *cause* Y?
- Exogeneiety: Z is assigned randomly or conditionally on controlled covariance, $\gamma_2 W$ in first-stage equation.
- Exclusion restriction: Z affects Y only through its influence on T. No "backdoor" between Z Y.

Choosing instruments (cont'd)

Selecting an instrument:

- Theoretically identify all possible source of variation in T
- Select ones that are least likely to be correlated with U. Exclusion resctrivtion.
  - DAGs are especially helpful here
- Estimate first stage equation to see if Z is a sufficiently strong (relevant) predictor of T.

Bad instruments

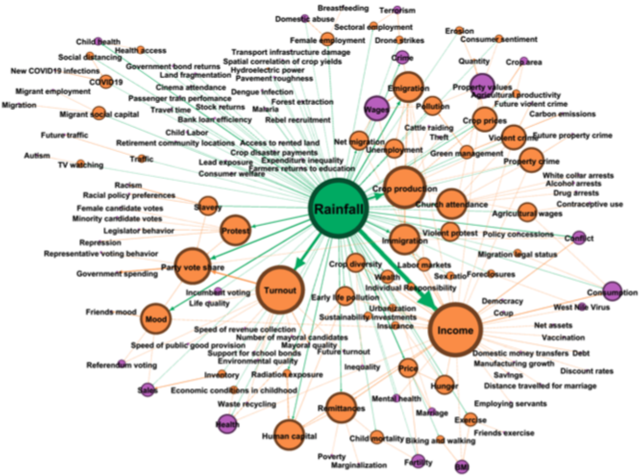


Figure: Rainfall IV

Good instruments?

Do highways cause suburbanisation?

Baum-Snow (2007) - did contruction of radial highways cause population decentralisation in US cities?

Baum-Snow et al., (2014) - did contruction of radial highways cause population decentralisation in Chinese cities?

IV summary

- Well identified ID can recover causal effects of urban policy interventions.
- However, credible inference from IV is not mechanistic.
- Requires strong theoretical consideration of the instrument and variation in T.
- Strong theory must be used to justify the two main identifying assumptions:
  - Relevance: Z is relevant predictor of T.
  - Exogenous: Z is assigned randomly or "as if" by random
  - Exclusion restriction: Z is uncorrelated with Y.
    $Z \rightarrow T \rightarrow Y$