Data Science for Public Policy

Applied Micro Methods II

Dr. Sergio Galletta

ETHZ Zurich

13/03/2023

Synthetic control method

- With this method, it is possible to generate a "**synthetic control**" for a treated unit i, allowing comparisons free from confounding.
- By assigning weights to the Y values of untreated units, a "synthetic control" is produced

$$\hat{Y}_{t,post}(0) = \mu + \sum_{i \in c} \omega_i Y_{i,T}$$

- Typically, it is necessary to estimate the ω_i, which are formed by minimizing the distance between covariates in the pre-period.

$$\{\hat{\omega}\}_i = \arg\min_{\boldsymbol{W}} ||\boldsymbol{X}_{\text{treat}} - \boldsymbol{X}_{\text{control}} \boldsymbol{W}||$$

Synthetic control method

- Importantly, $\boldsymbol{X}$ can include both lagged outcomes and covariates.
- Reconnecting with the idea of what is observable and what is not:
  - Unobserved outcomes: $Y_{t,post}(0)$, $Y_{c,post}(1)$
  - Observed outcomes: $Y_{t,post}(1)$, $Y_{c,post}(0)$
  - Observed covariates / predictors: $Y_{t,pre}(0)$, $Y_{c,pre}(0)$, X_t, X_c
- Relevant method if one wants to study just one treated unit
- Recent papers have introduced Synthetic Difference-in-Differences (SDID)

Example

The Economic Costs of Organised Crime - Pinotti 2015 *EJ*

- Is organized crime good or bad for the economy?

Example

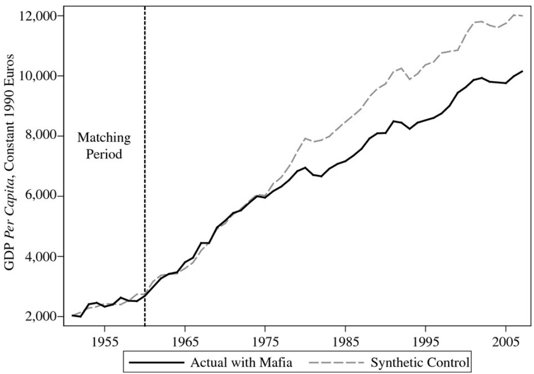
The Economic Costs of Organised Crime - Pinotti 2015 *EJ*

- Is organized crime good or bad for the economy?
- Expansion of Mafia to regions previously unaffected (Apulia and Basilicata)
- From the minimization problem, the control group is created by weights from Abruzzo (0.624) and Molise (0.376).

Example

The Economic Costs of Organised Crime - Pinotti 2015 *EJ*

- Is organized crime good or bad for the economy?
- Expansion of Mafia to regions previously unaffected (Apulia and Basilicata)
- From the minimization problem, the control group is created by weights from Abruzzo (0.624) and Molise (0.376).
- GDP growth slows down after mafia activities expanded to new regions



Example

Tax Reform and Foreign Inventors - Akcigit, Baslandze and Stantcheva 2016 *AER*

- Do people respond to changes in tax burden?

Example

Tax Reform and Foreign Inventors - Akcigit, Baslandze and Stantcheva 2016 *AER*

- Do people respond to changes in tax burden?
- In 1992, Denmark reduced taxes on foreign researchers
- From the minimization problem, the control group is created by weights from Switzerland, Canada, and Portugal

Example

Tax Reform and Foreign Inventors - Akcigit, Baslandze and Stantcheva 2016 *AER*

- Do people respond to changes in tax burden?
- In 1992, Denmark reduced taxes on foreign researchers
- From the minimization problem, the control group is created by weights from Switzerland, Canada, and Portugal
- Increased in the share of foreign inventors

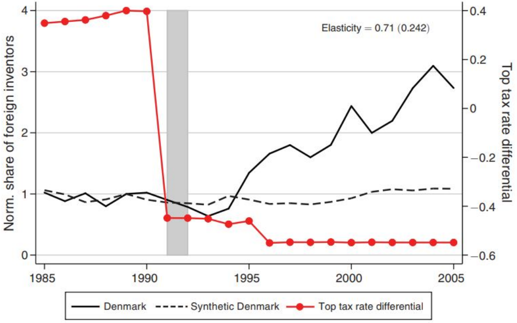


Figure 9. Denmark's 1992 Tax Reform and Foreign Inventors

Instrumental variables

- **Instrumental variables (IV)** is the most popular solution for dealing with endogenous treatments
- The 2021 Nobel prize for economics was assigned to researchers that linked the potential outcome framework with IV and introduced the concept of **local average treatment effect** (LATE), the actual type of estimands that IV delivers

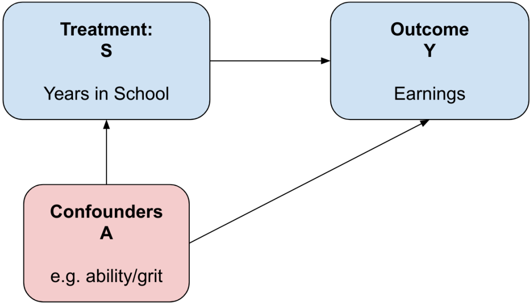
Instrumental variables

- **Instrumental variables (IV)** is the most popular solution for dealing with endogenous treatments
- The 2021 Nobel prize for economics was assigned to researchers that linked the potential outcome framework with IV and introduced the concept of **local average treatment effect** (LATE), the actual type of estimands that IV delivers
- Let's go back to the link between education and income

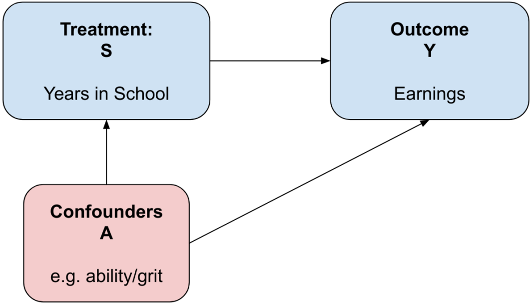
$$
\begin{aligned}
Y_i &= \alpha + \rho S_i + \epsilon_i \\
Y_i &= \alpha + \rho S_i + \phi \underbrace{A_i}_{\text{unobs}} + \eta_i
\end{aligned}
$$

- OLS estimates for $\hat{\rho}$ will be biased.

Instrumental Variables: Main Intuition

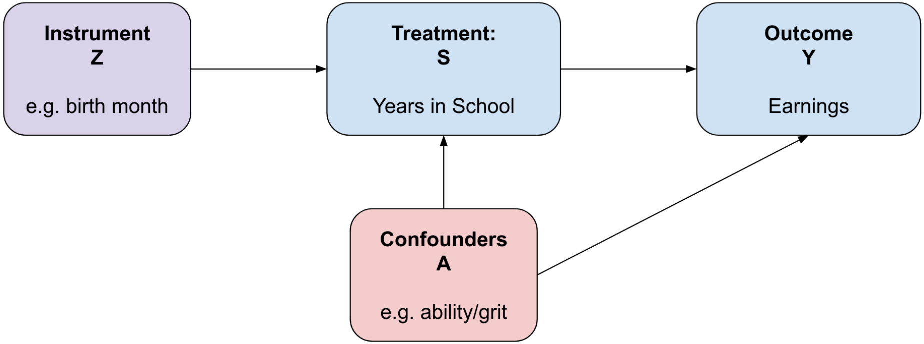


Instrumental Variables: Main Intuition



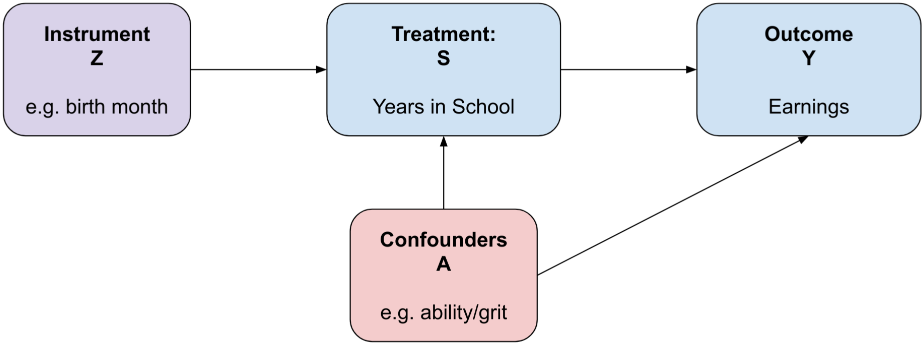
Instrumental Variable (IV): to identify a variable, that is correlated with S_i, but not correlated with anything else affecting Y_i.

Instrumental Variables: Main Intuition



- We identify a source of variation in treatment assignment that is as good as random – orthogonal to any relevant unobserved confounder.
- We compare individuals that, due to the instrument, are shifted between the control group and the treatment group.

What is a valid instrumental variable?



1. Correlated with the causal variable, e.g. S_i:

$$\text{Cov}[Z_i, S_i] \neq 0$$

2. Uncorrelated with any other determinants of outcome Y:

$$\text{Cov}[Z_i, \epsilon_i] = 0$$

What is a valid instrumental variable?

(1) Exogeneity: No unobserved factors affect both the outcome and the instrument:

$$\epsilon_i \nrightarrow Z_i$$

- **No "Z-confounders"**

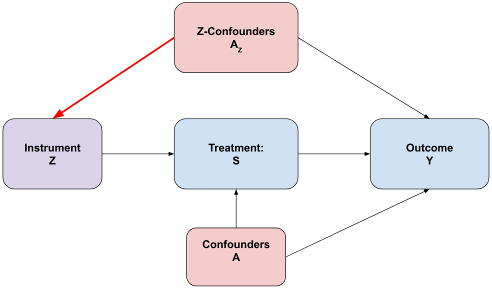
What is a valid instrumental variable?

(1) Exogeneity: No unobserved factors affect both the outcome and the instrument:

$$\epsilon_i \not\rightarrow Z_i$$

- **No "Z-confounders"**

Violation of exogeneity:



What is a valid instrumental variable?

(1) Exogeneity: No unobserved factors affect both the outcome and the instrument:

$$\epsilon_i \nrightarrow Z_i$$

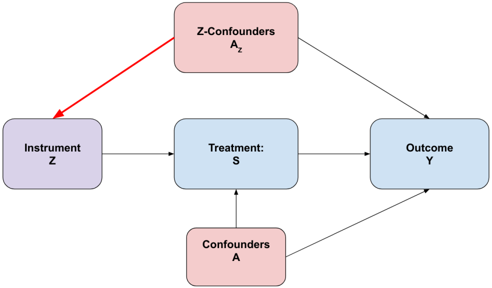
- **No "Z-confounders"**

(2) Exclusion: Instrument only affects outcome through treatment variable:

$$Z_i \nrightarrow \epsilon_i$$

- **"single mediator" condition**

Violation of exogeneity:



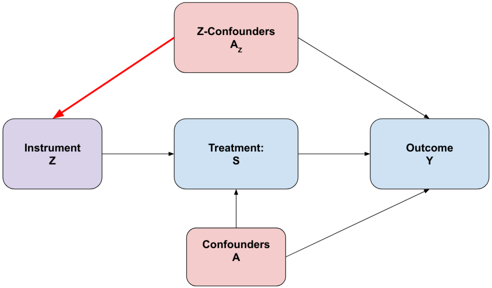
What is a valid instrumental variable?

(1) Exogeneity: No unobserved factors affect both the outcome and the instrument:

$$\epsilon_i \nrightarrow Z_i$$

- **No "Z-confounders"**

Violation of exogeneity:

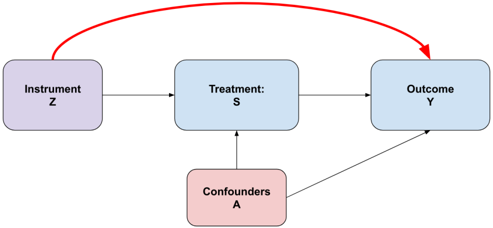


(2) Exclusion: Instrument only affects outcome through treatment variable:

$$Z_i \nrightarrow \epsilon_i$$

- **"single mediator" condition**

Violation of exclusion:



Good instruments are hard to find

- Good instruments come from a combination of three ingredients:
  - Good institutional knowledge
  - Theoretical modeling
  - Last but not least: Originality

Good instruments are hard to find

- Good instruments come from a combination of three ingredients:
  - Good institutional knowledge
  - Theoretical modeling
  - Last but not least: Originality
- Some usual sources of instruments:
  - Nature (e.g., genes, weather)
  - Assignment rules (e.g., random assignment of judges to cases)
  - ‘Natural’ experiments (e.g. lottery drafts, sudden policy changes, geographic or temporal cutoffs)

Two-Stage Least Squares (2SLS)

IV estimates are equivalent to running two separate OLS regressions:

1. Estimate"first stage", regressing treatment on instrument:

$$S_i = \gamma Z_i + \nu_i$$

Two-Stage Least Squares (2SLS)

IV estimates are equivalent to running two separate OLS regressions:

1. Estimate"first stage", regressing treatment on instrument:

$$S_i = \gamma Z_i + \nu_i$$

2. Form prediction $\hat{S}_i = \hat{\gamma} Z_i$ and estimate the "second stage", regressing outcome on first-stage-predicted treatment:

$$Y_i = \rho \hat{S}_i + \epsilon_i$$

Two-Stage Least Squares (2SLS)

IV estimates are equivalent to running two separate OLS regressions:

1. Estimate"first stage", regressing treatment on instrument:

$$S_i = \gamma Z_i + \nu_i$$

2. Form prediction $\hat{S}_i = \hat{\gamma} Z_i$ and estimate the "second stage", regressing outcome on first-stage-predicted treatment:

$$Y_i = \rho \hat{S}_i + \epsilon_i$$

- First stage is driven by "compliers" (responders to instrument).
- Standard 2SLS estimates give a "local average treatment effect" on the complier population.

Can we test validity of IV?

- Is Z_i correlated with causal variable of interest, S_i?
  - YES: check for the significance of the first stage (first-stage F-statistic)
  - The standard is F $>$ 10, but recent studies show you might need more
  - With weak instruments IV bias towards the OLS

Can we test validity of IV?

- Is Z_i correlated with causal variable of interest, S_i?
  - YES: check for the significance of the first stage (first-stage F-statistic)
  - The standard is F $>$ 10, but recent studies show you might need more
  - With weak instruments IV bias towards the OLS
- Is Z_i uncorrelated with any other determinants of Y_i?
  - Untestable, use logic and theory to argue in favor the assumption
  - But often indirect ways to probe exogeneity and exclusion

Can we test validity of IV?

- Is Z_i correlated with causal variable of interest, S_i?
  - YES: check for the significance of the first stage (first-stage F-statistic)
  - The standard is F $>$ 10, but recent studies show you might need more
  - With weak instruments IV bias towards the OLS
- Is Z_i uncorrelated with any other determinants of Y_i?
  - Untestable, use logic and theory to argue in favor the assumption
  - But often indirect ways to probe exogeneity and exclusion
- Additional assumption that is important (and untestable)
  - Monotonicity: the instrument(s) should have a monotonic relationship with the endogenous explanatory variable(s)

Reduced Form

"Reduced Form" (RF) means regressing the outcome directly on the instrument:

$$Y_i = \alpha + \phi Z_i + \epsilon_i$$

- Papers will normally report this along with 2SLS estimates.
- For causal interpretation, RF requires exogeneity but not exclusion.

Example

Media and local finance - Ash and Galletta, 2023 *AEJPOL*

- How national cable news affects local public policy?

Example

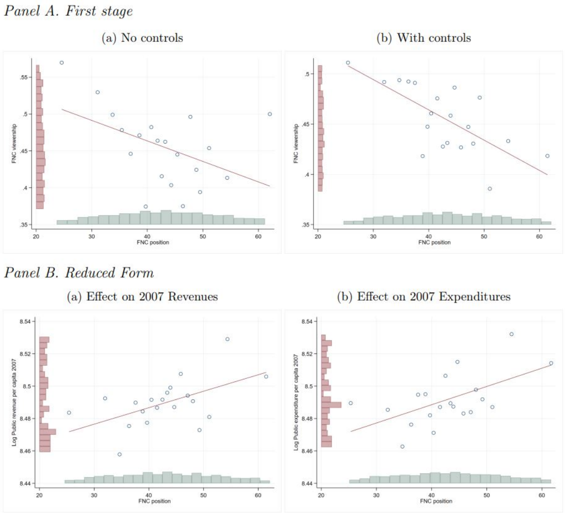
Media and local finance - Ash and Galletta, 2023 *AEJPOL*

- How national cable news affects local public policy?
- Variation in channel position depending on the area of residence
- Higher channel position lower viewership

Example

Media and local finance - Ash and Galletta, 2023 *AEJPOL*

- How national cable news affects local public policy?
- Variation in channel position depending on the area of residence
- Higher channel position lower viewership
- Fox News did decrease the size of local budgets



Example

War and Tax Evasion - Galletta and Giommoni 2023 WP

- Does exposure to war violence affect tax evasion?

Example

War and Tax Evasion - Galletta and Giommoni 2023 WP

- Does exposure to war violence affect tax evasion?
- Death of a relative in WWI as treatment
- Exogenous allocation of soldiers to more/less risky military units

Example

War and Tax Evasion - Galletta and Giommoni 2023 WP

- Does exposure to war violence affect tax evasion?
- Death of a relative in WWI as treatment
- Exogenous allocation of soldiers to more/less risky military units
- Yes, higher tax evasion when a relative died during the war

Table 3: Effect of War on Tax noncompliance - IV

	(1)	(2)	(3)	(4)
***Panel A*: OLS**				
Death of a relative in the battlefield	0.028***	0.035***	0.011*	0.011*
	(0.007)	(0.006)	(0.006)	(0.006)
***Panel B*: First stage**				
Risk of military unit	0.101***	0.105***	0.098***	0.104***
	(0.004)	(0.004)	(0.005)	(0.007)
F-stat	634.1335	835.1041	332.6722	219.2009
***Panel C*: Reduced form**				
Risk of military unit	0.019***	0.021***	0.004*	0.007**
	(0.003)	(0.003)	(0.003)	(0.003)
***Panel D*: IV**				
Death of a relative in the battlefield	0.186***	0.200***	0.044*	0.064**
	(0.031)	(0.028)	(0.026)	(0.030)
N Observations	54,990	54,771	51,486	49,472
Baseline controls		✓	✓	✓
Surname FE			✓	
Province FE			✓	
Municipality FE				✓
Surname FE × Province FE				✓

Regression Discontinuity Design (RDD)

- **Regression Discontinuity Design (RDD)** is a quasi-experimental design that has gained popularity among researchers because it can provide more credible causal estimates than other designs.
- It exploits that individuals are assigned to treatment or control groups based on a running variable (e.g., test score, distance, or class size) with a discontinuity at a certain threshold or **cutoff point**.

Sharp vs Fuzzy Regression Discontinuity Design

- In the Sharp RDD, individuals who score above the cutoff receive the treatment, and those who score below the cutoff do not.
- In the Fuzzy RDD, the probability of receiving the treatment changes discontinuously at the cutoff, but not all individuals who score above the cutoff receive the treatment.
- Fuzzy RDD is essentially equivalent to an Instrumental Variables (IV) design, where the running variable serves as the instrument for the treatment.

Notation and Key Assumptions

- Let $Y_i(0)$ and $Y_i(1)$ be the potential outcomes of individual i when they do not receive the treatment and when they do, respectively.
- Let D_i be the treatment indicator such that $D_i = 1$ if individual i receives the treatment, and $D_i = 0$ otherwise.
- Let Z_i be the running variable that assigns individuals to treatment or control.
- We assume that Z_i **has a discontinuity at a threshold** value z_0, where the treatment is assigned to individuals with $Z_i \geq z_0$.
- We assume that $Y_i(0)$ **and** $Y_i(1)$ **are continuous in** Z_i **around the threshold value**, which allows us to estimate the LATE at the threshold.

Regression Discontinuity Design (RDD)

- If the previous assumptions hold

$$\tau_{ATE} = E(Y_i(1) - Y_i(0)|Z_i = 0) = \lim_{z \downarrow 0} E(Y_i|Z_i = z) - \lim_{z \uparrow 0} E(Y_i|Z_i = z)$$

- But, this is a very particular subgroup of individuals right at the cutoff

Regression Discontinuity Design (RDD)

- The basic RDD model can be expressed as:

$$Y_i = \alpha + \beta T_i + \gamma(Z_i - c) + \epsilon_i$$

- where Y_i is the outcome variable
- T_i is a binary treatment
- Z_i is the continuous variable used for assignment
- c is the cutoff point, and ϵ_i is the error term
- Local linear regression method by selecting a small neighborhood around the cutoff point.

RDD Check list

- A graphical representation and test of “balance” and first stage (if fuzzy)
- Permutation test of characteristic at cutoff
- The density of the forcing variable (Mcrary test)
- Placebo checks
- A graphical representation of the outcomes
- Estimates based on optimal bandwidth choice and robust inference, using local linear analysis
  - These decisions vary depending on running variable. If discrete running variable, need to account for discreteness (Kolesar and Rothe (2018))
  - Should use local linear regression, and not global polynomials (Gelman and Imbens)
- Robustness analysis along bandwidth choice (and other tuning parameters)
  - Present this graphically

Example

Female Mayor and Violence against Women - Bochenkova, Buonanno and Galletta 2023 *JDevE*

- Does female mayor influence violence against women?

Example

Female Mayor and Violence against Women - Bochenkova, Buonanno and Galletta 2023 *JDevE*

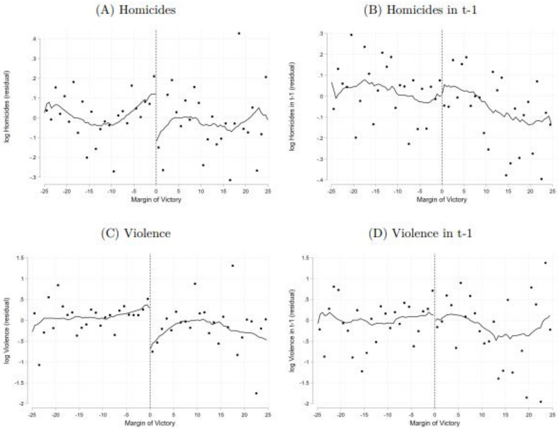
- Does female mayor influence violence against women?
- Compare Brazilian municipalities where a female candidate barely won to those where a female candidate barely lost mayoral elections
- Winning is random for those close to 50%

Example

Female Mayor and Violence against Women - Bochenkova, Buonanno and Galletta 2023 *JDevE*

- Does female mayor influence violence against women?
- Compare Brazilian municipalities where a female candidate barely won to those where a female candidate barely lost mayoral elections
- Winning is random for those close to 50%
- Yes, having a female mayor reduces crime against women

Figure 1: Female Mayor and Violence against Women



Example

Direct democracy and social preferences - Galletta 2021 *JEBO*

- Does political institutions affect preferences for redistribution?

Example

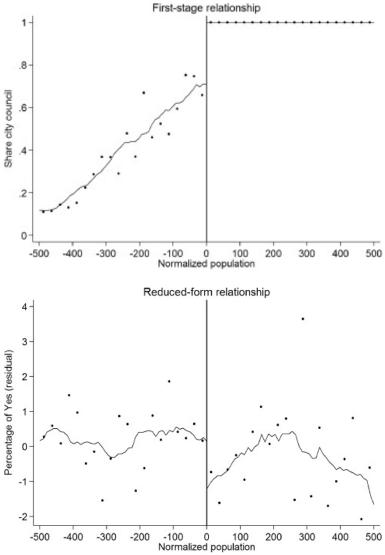
Direct democracy and social preferences - Galletta 2021 *JEBO*

- Does political institutions affect preferences for redistribution?
- Exploits a discrete change in the probability that a municipality has representative democracy based on a population threshold
- $\geq$ 800 inhabitants adopt a city council

Example

Direct democracy and social preferences - Galletta 2021 *JEBO*

- Does political institutions affect preferences for redistribution?
- Exploits a discrete change in the probability that a municipality has representative democracy based on a population threshold
- $\geq$ 800 inhabitants adopt a city council
- Yes, representative democracy reduces the share of votes in favor of public spending



Regression Discontinuity Aggregation (RDA)

- In a classic RDD, treatment is assigned by whether an observation's running variable crosses a fixed cutoff.
- Borusyak and Kolerman-Shemer 2024 extend the RDD framework to cases where a unit is exposed to multiple discontinuity events
- In such settings, a unit's "treatment" is no longer a simple binary indicator at a single cutoff, but rather an aggregate measure of several RD shocks across events
- Important finding: RDA solutions provide a causal estimate as far as the standard RDD assumptions are satisfied

Upper-Level IV Estimator (RDA)

- Estimate causal effects at an aggregate level by combining multiple regression discontinuity events.
- Aggregate treatment indicators from multiple local RD events into a single measure per aggregate unit (e.g., multiple elections results).
- Construct an instrumental variable from aggregated near-threshold discontinuities (e.g., elections results that happen "by chance").
- Use this IV approach to estimate how aggregated treatments affect the aggregated outcome (e.g., economic outcomes).
- Include aggregated controls to mimic local comparisons at each RD threshold.

Lower-Level (Stacking) Estimator (RDA)

- Stack multiple individual RD events into one dataset and performing a fuzzy RD analysis.
- Create a dataset with each RD event as a separate observation.
- Assign identical aggregated outcomes and treatments to all events from the same aggregate unit.
- Perform regression discontinuity analysis on the stacked data, controlling for individual running variables.

Example

Political Influence during Childhood - Daniele, Galletta, Le Moglie and Masera 2025 *almost a WP*

- Does early-life exposure to political ideology shape voters' political preferences?

Example

Political Influence during Childhood - Daniele, Galletta, Le Moglie and Masera 2025 *almost a WP*

- Does early-life exposure to political ideology shape voters' political preferences?
- Study context: Ideological exposure to U.S. state governors during childhood

Example

Political Influence during Childhood - Daniele, Galletta, Le Moglie and Masera 2025 *almost a WP*

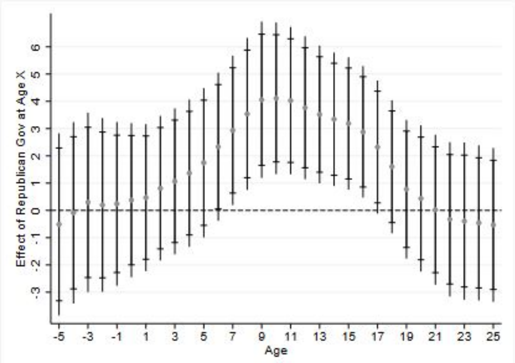
- Does early-life exposure to political ideology shape voters' political preferences?
- Study context: Ideological exposure to U.S. state governors during childhood
- Data: Party preferences from over 200 million U.S. voters

Example

Political Influence during Childhood - Daniele, Galletta, Le Moglie and Masera 2025 *almost a WP*

- Does early-life exposure to political ideology shape voters' political preferences?
- Study context: Ideological exposure to U.S. state governors during childhood
- Data: Party preferences from over 200 million U.S. voters
- Main finding: Significant ideological imprinting, strongest effect at age 10

Figure 6: Growing up under a Republican governor: Different ages



Example

Political Influence during Childhood - Daniele, Galletta, Le Moglie and Masera 2025 *WP*

Upper-Level Estimator

$$ShRep_{i,s}^{a,b} = \gamma_s + \gamma_0 ShRepNE_{i,s}^{a,b} + \gamma_1 Q_{i,s} + \gamma_2 BY_{i,s} + \epsilon_{i,s}, \tag{2 FS}$$

$$Y_{i,s} = \beta_s + \beta_0 \widehat{ShRep_{i,s}^{a,b}} + \beta_1 Q_{i,s} + \beta_2 BY_{i,s} + \rho_{i,s}, \tag{2 SS}$$

- $ShRep_{i,s}^{a,b} = \sum_{\tau=a}^{b} \frac{RepGov_{i,s}^{\tau}}{(b-a+1)}$: Share of years under a Republican governor.
- $ShRepNE_{i,s}^{a,b}$: Same as $ShRep_{i,s}^{a,b}$, but only using elections with $MV_{i,s}^{\tau} < c$.
- $Q_{i,s}$: Aggregated controls (e.g., share of narrow elections, average margin of victory).
- $BY_{i,s}$: Birth year fixed effects.

Example

Political Influence during Childhood - Daniele, Galletta, Le Moglie and Masera 2025 *WP*

Lower Level Estimator

$$ShRep_{i,s}^{a,b} = \gamma_s + \gamma_0 RepGov_{i,s}^{\tau} + \gamma_1 MV_{i,s}^{\tau} + \gamma_2 (MV_{i,s}^{\tau} \times RepGov_{i,s}^{\tau}) + \gamma_3 BY_{i,s} + \epsilon_{i,s}, \tag{3 FS}$$

$$Y_{i,s} = \beta_s + \beta_0 \widehat{ShRep_{i,s}^{a,b}} + \beta_1 MV_{i,s}^{\tau} + \beta_2 (MV_{i,s}^{\tau} \times RepGov_{i,s}^{\tau}) + \beta_3 BY_{i,s} + \rho_{i,s}, \tag{3 SS}$$

- $RepGov_{i,s}^{\tau}$: Dummy for a Republican governor at age τ.
- $MV_{i,s}^{\tau}$: Margin of victory in the election at age τ.
- $ShRep_{i,s}^{a,b}$: Defined as in the upper-level estimator.

Example

Political Influence during Childhood - Daniele, Galletta, Le Moglie and Masera 2025 *WP*

	Share Years Rep Governor (7-16) (1)	Republican (2)	Republican (3)	Democrat (4)	Independent (5)	Republican (6)
			Panel A: Upper-Level Estimation			
Share Years Rep Governor (7-16)		6.76***	8.65***	-4.49**	-2.48	
		(1.44)	(1.86)	(1.75)	(1.80)	
Upper Level Instrument	0.33***					2.82***
	(0.03)					(0.52)
Observations	141.82M	141.82M	98.67M	141.82M	141.82M	98.67M
Sample	Full Sample	Full Sample	2Party-Registered	Full Sample	Full Sample	2Party-Registered
F-stat		124.8	121.7	124.8	124.8	
			Panel B: Lower-Level Estimation			
Share Years Rep Governor (7-16)		8.78***	11.99***	-8.33***	-1.11	
		(1.11)	(1.58)	(1.45)	(1.29)	
Lower Level Instrument	0.32***					3.58***
	(0.03)					(0.38)
Observations	724.73M	724.73M	509.38M	724.73M	724.73M	509.38M
Sample	Full Sample	Full Sample	2Party-Registered	Full Sample	Full Sample	2Party-Registered
F-stat		154.8	131.0	154.8	154.8	

To recap

- We care about causality
- Potential outcome framework
- Quasi-experimental methods
  - Diff-in-Diff
  - Synthetic controls
  - IV
  - RDD