

Lecture 6 - Identification: IV and RDD

Topics in Econometrics - M2 ENS Lyon

Vincent Bagilet

Introduction

Goal of the session

- Strengthen **intuition** on Instrumental Variables (**IV**) and Regression Discontinuity Designs (**RDD**)
- Review **key assumptions**
- Focus on some specific points:
 - Locality
 - Weak instruments
 - Shift-share instruments
- Learning through **applied examples** and explore some limits through **simulations**

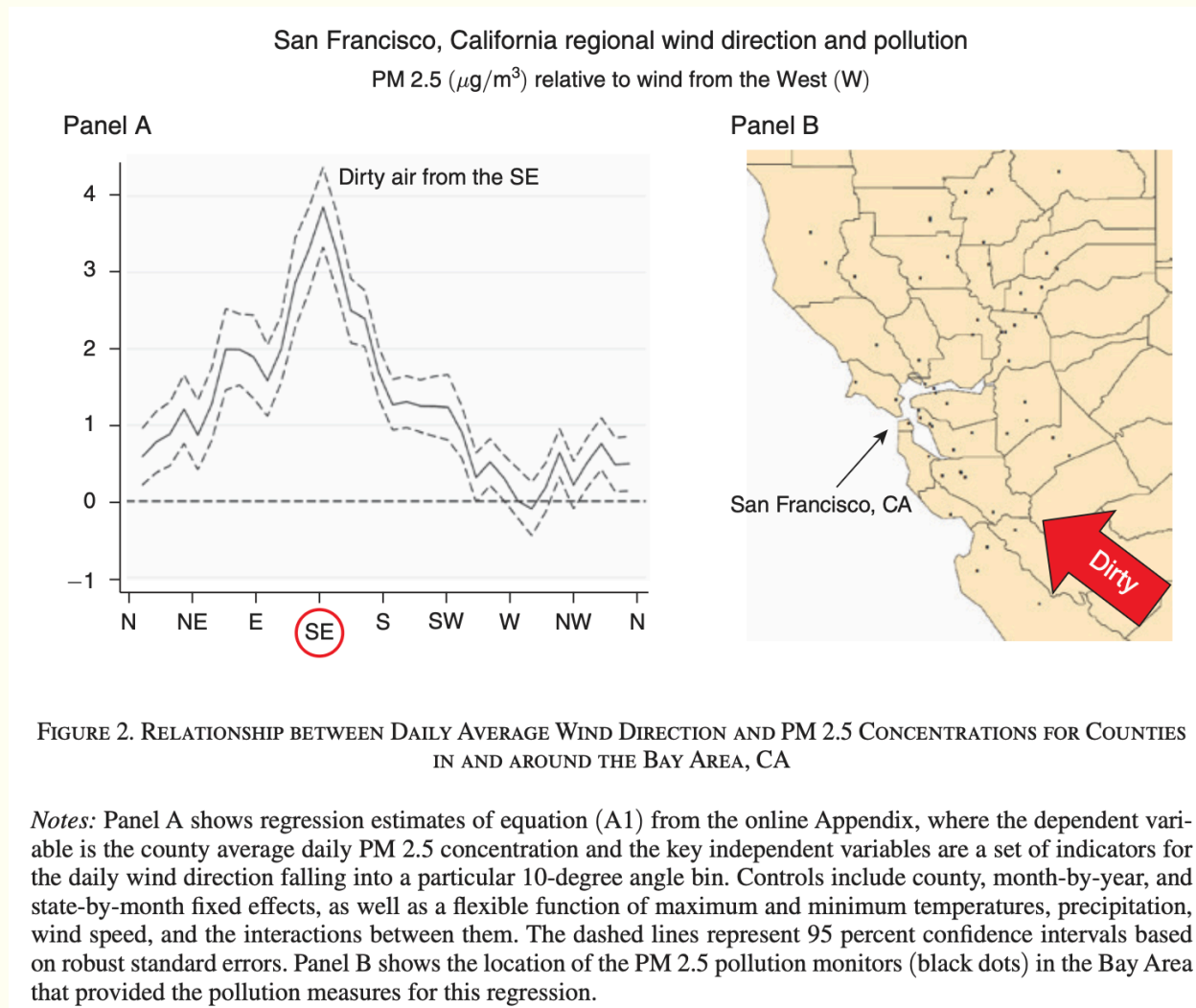
IV: Introductory example

Air pollution and health

- How would you measure health effects of air pollution?
- Why not regress an health outcome (mortality or hospital admissions for instance) on air pollution?
 - OVB: air pollution might be affected by unobserved factors (eg economic activity, income, healthcare, etc)
 - Measurement error (if correlated with treatment)
 - How to separate between pollutants?
- How would you proceed?
 - Need to find a source of **exogeneity**
 - Something that affects air pollution levels but do not otherwise affect health

An instrument for air pollution

- Find an exogenous shifter in air pollution
- Wind direction (Deryugina et al. 2019)



Why does that work?

Intuition

- Wind $\Rightarrow$ “as-good-as-random” variation in exposure to pollutants
 - Mimics randomized experiment
-
- Wind **affects pollution levels**
 - *ie* relevance
 - Wind direction **varies quasi-randomly over time**
 - *ie* independence (or exogeneity)
 - Wind itself **doesn't affect health** (after controlling for temperature and humidity)
 - *ie* exclusion restriction
 - Individuals **down-wind of pollution source receive more pollution** (no defiers)
 - *ie* monotonicity

Other ideas of instruments?

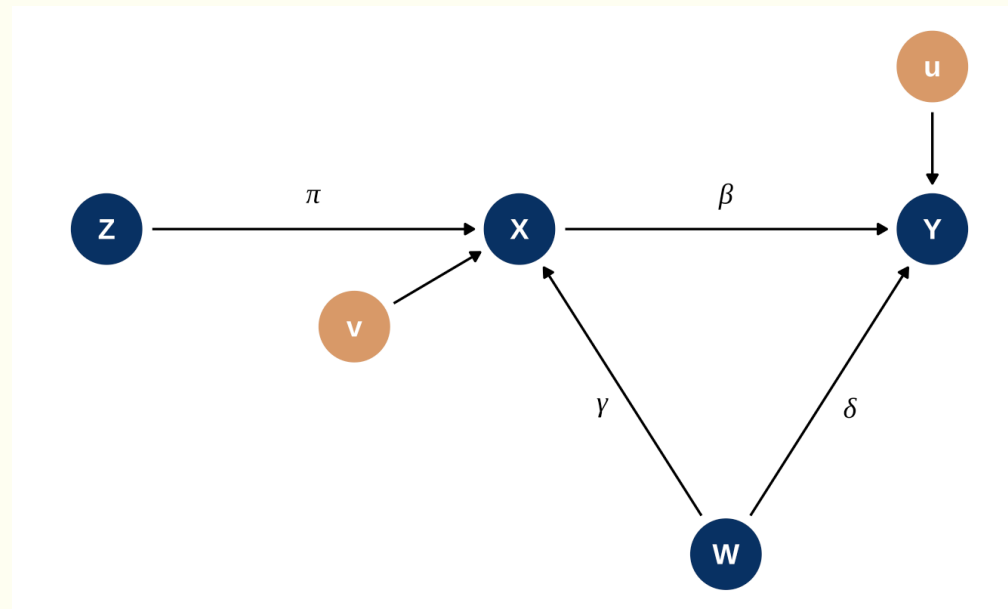
- Boat traffic in a port (Moretti and Neidell 2011)
- Thermal inversions (Arceo, Hanna, and Oliva 2016)
- Storms causing traffic jam in airports (Schlenker and Walker 2016)

- Why are they each valid instruments?

IV: Fundamentals

Formal intuition

- $y_i = \alpha + \beta x_i + u_i$ with $\text{cov}(x, u) \neq 0$
- Problem: x_i is **endogenous**
- **IV solution:**
 - Find a variable z that shifts x for reasons unrelated to unobservables in y
 - It isolates variation in x that is unrelated to u and recover β
 - Focuses on the variation that is explained by the instrument



Potential outcomes

Group	$d_i(z_i = 0)$	$d_i(z_i = 1)$	Effect of instrument
Compliers	0	1	Instr. $\nearrow$ proba of treatment
Defiers	1	0	Instr. $\searrow$ proba of treatment
Never-takers	0	0	Instr. has no effect
Always-takers	1	1	Instr. has no effect

- Effect of the instrument on the treatment **varies** across individuals
- We only observe z_i and d_i , not these groups

Warning

The IV estimator identifies the effect for **compliers only** (a *Local Average Treatment Effect*, LATE, under usual assumptions)

Identifying assumptions

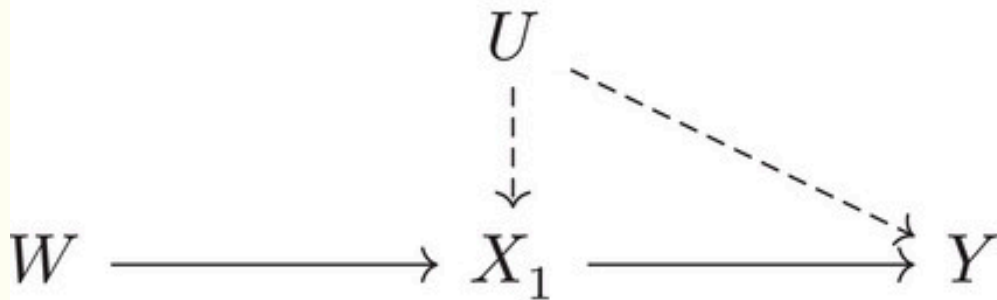
Assumption.	Formal expression	Intuition
Independence, exogeneity	$\text{Cov}(z, u) = 0$	No unobserved confounders affecting both z and x
Exclusion restriction	$\text{Cov}(z, u) = 0$	z affects y only through x
Relevance	$\text{Cov}(z, x) \neq 0$	z does affect x
Monotonicity	$x_i(z_i = 1) \geq x_i(z_i = 0)$ with z (no defiers)	z is an incentive, does not discourage treatment

- Exclusion restriction $\Rightarrow ITT_{\text{never-takers}} = ITT_{\text{always-takers}} = 0$.
- Monotonicity $\Rightarrow$ compliers are the only individuals whose treatment status could be altered by the instrument

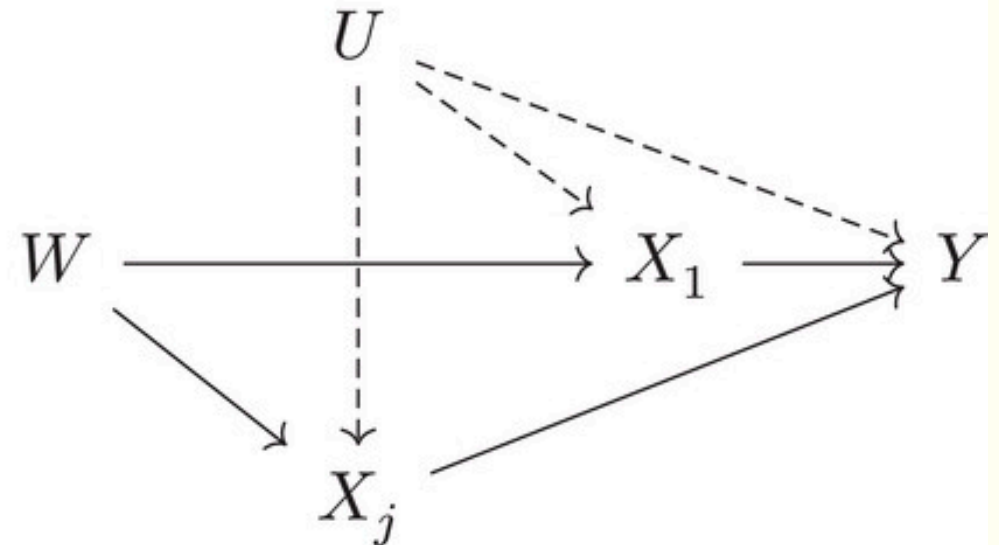
Exclusion restriction

- We **can never show** that the exclusion restriction holds
- Need to convince the audience that it does with **logic** arguments
- Need to find “clever” instruments
- But some instruments may turn out to be not so great

Instrumental variable



Exclusion-restriction violation



Exclusion restriction violations

- Sometimes an instrument affects many (many) other variables $\Rightarrow$ it may affect the outcome through another path (Mellon 2024)

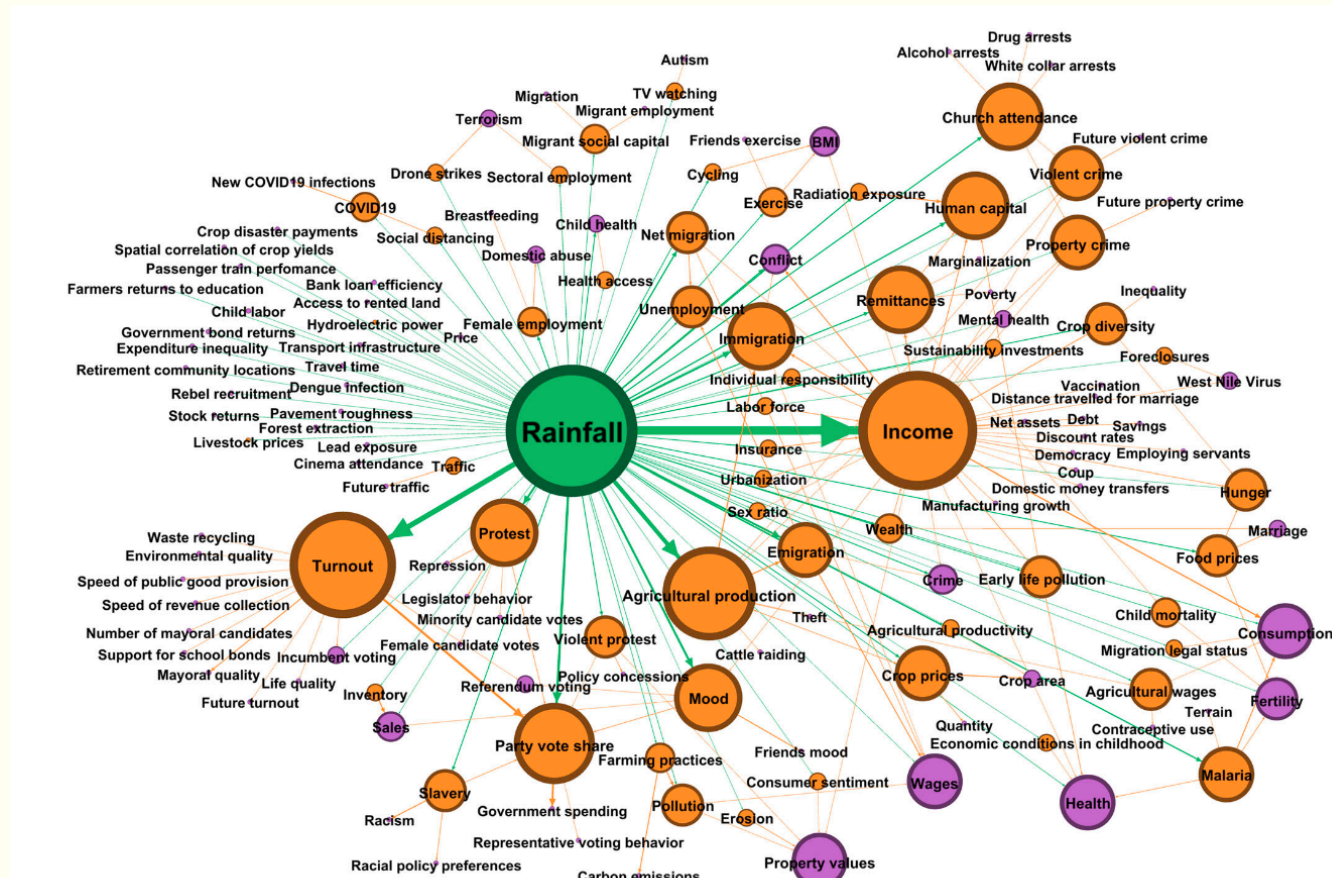


FIGURE 2 Causal web of rainfall from 192 papers. *Note:* This figure illustrates the plethora of variables causally linked to weather. These can represent exclusion-restriction violations. Node and tie sizes are proportional to appearances in literature. Colors: weather (green), instrumented variable (orange), and outcome (purple).

IV: Locality

What does the IV estimate?

$$\begin{aligned}\beta_{IV} &:= \frac{\text{Cov}(y_i, z_i)}{\text{Cov}(d_i, z_i)} = \dots \\ &= \frac{\mathbb{E}[y_i | z_i = 1] - \mathbb{E}[y_i | z_i = 0]}{\mathbb{E}[d_i | z_i = 1] - \mathbb{E}[d_i | z_i = 0]} = \dots \\ &= \underbrace{\mathbb{E}[y_i(1) - y_i(0) \mid d_i(1) = 1, d_i(0) = 0]}_{\text{LATE on the compliers}}\end{aligned}$$

IV $\neq$ ATE

- The IV estimator only captures the causal effect for those individuals whose treatment status changes because of the instrument (the compliers)
- ie the **ATE for compliers**

Intuition

- The instrument isolates the variation in treatment **driven by z**
- If treatment effects are heterogeneous, IV does **not** recover the overall average treatment effect (ATE)
- Instead, it identifies the effect **local** to those influenced by the instrument

Take-away message

- Always ask yourself: **who are the compliers?**
- The LATE is causal, but is local: it **depends on the instrument used**
- Using a different instrument identifies a different LATE

Example

- Want to estimate the causal effect of **education** (x) on **earnings** (y)
- Education not randomly assigned
- Studied in Card (1993)
- Solution:
 - Use an instrument z , such as **proximity to a college**
 - Is it a valid instrument? Yes, arguably. It affects the likelihood of getting more education but does not affect earnings directly
- It estimates:
 - The ATE of education on earnings **for those whose education decisions are changed by the instrument**
 - *ie* students who attend college because they live nearby

Differences in estimands

Estimand	Definition	Estimates the effect for?	Identified by
ATE	$\mathbb{E}[y_i(1) - y_i(0)]$	The entire population	Randomized experiment
ATT	$\mathbb{E}[y_i(1) - y_i(0) \mid d_i = 1]$	Those who actually receive treatment	Selection on observables
LATE	$\mathbb{E}[y_i(1) - y_i(0) \mid d_i(1) = 1, d_i(0) = 0]$	Those whose treatment status is affected by the instrument (the compliers)	Instrumental Variables (IV)

Why instrument a *randomised* treatment?

- **Oregon**, 2008: **Medicaid** expansion, but cannot cover everyone
- 85,000 people sign up → 30,000 drawn at random → only 10,000 actually enrol
- The lottery randomises the offer, not the treatment
- Which comparison should we use?
 - Enrolled vs not enrolled?
 - Winners vs losers?
 - Finkelstein et al. (2012) instrument **being insured** with **winning the lottery**

Different estimands

- Find much larger effects with the IV than its reduced form
- Estimate different estimands

Two different questions

- **ITT**: what does *offering* insurance do? The policy-relevant number if take-up stays the same
- **LATE**: what does *having* insurance do, for those who enrolled because they won

2SLS actually often not a LATE

- LATE theorem assumes **no covariates** (or a fully saturated specification)
- But virtually all applied studies uses controls
- With controls, 2SLS = a weighted average of treatment effects not only for compliers (Blandhol et al. 2026)
 - Some **weights are negative**
 - Show that the issue is widespread

IV: Estimation

IV Estimator and 2SLS

- **Wald estimator**: sample analog of the estimand

$$\widehat{\beta}_W = \frac{\widehat{\text{Cov}}(y, z)}{\widehat{\text{Cov}}(d, z)}$$

- Numerically equivalent to **Two-Stage Least Squares (2SLS) estimator** $\widehat{\beta}_{2SLS}$ obtained through:

1. Regress d on instrument(s) z and controls and retrieve predicted values

$$d_i = \pi_0 + \pi_1 z_i + v_i \quad \Rightarrow \quad \widehat{d}_i = \widehat{\mathbb{E}}[d_i | z_i]$$

1. Regress y on predicted treatment $\widehat{d}$

$$y_i = \alpha + \beta \widehat{d}_i + e_i$$

Warning

- SEs of the 2nd stage do not give correct SEs (they do not account for the first-stage estimation)
- Need to adjust for the two stages
- 2SLS packages (eg `ivreg`, `fixest`) automatically adjust SEs
- Alternatively, use bootstrapping to get correct standard errors

Control function approach

- Alternative approach for the IV
- Steps:
 1. Regress d on z and controls
 2. Retrieve the residuals (the part of d unexplained by z)
 3. Regress y on d and these residuals
- Adjusting for these residuals **partials this unexplained variation out**
- Only leaves out the explained variation

Warning

- The point estimate is exactly $\hat{\beta}_{2SLS}$, but the OLS standard error is **too small**
- Fix: **bootstrap** both steps, or just read the SE off `feols/ivreg`

Practical advice

- Think about **what are your compliers**
 - Different instruments $\Rightarrow$ different estimands
- Make sure instrument is **relevant** w.r.t. the policy of interest
- If weak IV (and you don't want to give up)
 - Use approaches more **robust to weak instruments** (Andrews, Stock, and Sun 2019)
- Adjust for all relevant pre-treatment variables (predictors of y not affected by d)
- For models non-linear in d , properties of 2SLS do not necessarily hold
 - Consider **alternative estimation strategies** (e.g., control function method)

Exercise

1. Run two regressions on the `wooldridge::card` data using `fixest::feols` to compare the 2SLS and the **control function** approaches
 - Key variables:
 - Outcome: `lwage` (log wage)
 - Treatment: `educ` (years of schooling)
 - Instrument: `nearc4` (near 4-year college)
 - Controls: `exper`, `black`, `south`, `smsa`, `fatheduc`, `motheduc`
2. Build a simple and naive simulation to explore the impact of the use of instruments on **measurement error** (depending on the correlation between measurement error and the treatment)

 More precise instructions on the website

IV: Weak instruments

Intuition

- In IV, use the variation that is **explained by the instrument**
- Throws out variation not explained by the instrument
- Throwing out variation increases variance $\Rightarrow \sigma_{\beta_{IV}}^2 \geq \sigma_{\beta_{OLS}}^2$
- May end up with too little variation to estimate the effect of the treatment

Definition

- What is a weak instrument?
 - An instrument that **does not strongly predict treatment** d_i
 - In the first stage ($d_i = \pi_0 + \pi_1 z_i + v_i$) a weak instrument has $\pi_1 \approx 0$.
- Why problematic?
 - IV uses the variation in d explained by z
 - If z barely changes d , the IV estimator is **noisy** and can be biased
 - Small changes in sample or error can lead to large changes in the estimate

Consequences

1. **Large variance:** Confidence intervals can be very wide
2. **Finite-sample bias:** IV can be biased towards OLS (especially in small samples) because the IV leverages endogenous variation
3. **Distorted inference:** Standard t-tests and F-tests become unreliable (non-normal errors)



Take-away message

Even if the instrument is valid, if it is weak, IV estimates are not trustworthy

- Testing for relevance: compute the **first-stage F-stat** but use the Olea and Pflueger (2013) effective F-stat
- Report identification-robust Anderson-Rubin confidence intervals
- But basically look for a stronger instrument

Behind the F-stat

- With one instrument, the first-stage F is (Keane and Neal 2024)

$$F = N \times \frac{R_{\text{partial}}^2}{1 - R_{\text{partial}}^2}$$

- $\Rightarrow F$ grows with the **sample size**. Instrument strength does not
- A negligible R_{partial}^2 passes $F > 10$ if N is large enough (Bound, Jaeger, and Baker 1995)
- And 10 is the wrong number: valid 5% t -inference needs $F > 100$ (Lee et al. 2022)

Exercise

- Build a simple **simulation** to explore the impact of weak instruments on your estimator (point estimate and standard errors)
- Do not forget to add some endogeneity
- Is the first-stage F-stat a good metric?

IV: Shift-share Instruments

Bartik's setting

- Shift-share instruments also named **Bartik instruments** after Bartik (1991)
- Explores the effect of immigration (I) on local wages (y)

$$y_{l,t} = \alpha + \delta I_{l,t} + \rho X_{l,t} + \varepsilon_{l,t}$$

- Problem: endogeneity of I , eg migrants locate where the economy is doing well
- Bartik instrument interacts:
 - **Shares**: where migrants from each origin k historically settled
 - **Shifts**: the national inflow from each origin

$$B_{l,t} = \sum_{k=1}^K \underbrace{z_{l,k,t_0}}_{\text{share}} \times \underbrace{m_{k,t}}_{\text{shift}}$$

- A city's predicted inflow is a weighted average of **national** flows, weighted by its **initial** composition

Intuition

- Exogeneity: the instrument uses **no local information about today**
 - National flows are not driven by one city
 - Shares are measured *before* the period studied
- The exogeneity and identification come from either:
 - The **initial shares** → argue that they are exogenous, conditional on controls. Shifts only give first-stage strength
 - The **many national shocks** → argue that shocks are numerous and quasi-randomly assigned. Shares need not be exogenous

Identifying assumptions

- Exogeneity comes from either **one** of, never both at once:
 - **Shares**: using the K shares as separate instruments
 - **Shocks**: the many shocks are as-good-as-randomly assigned
- The exclusion restriction...
 - is likely **not violated because** national inflows are driven by push factors in the origin countries, and the shares are predetermined
 - **would be violated if** the initial composition predicts wage *growth* directly
- Relevance is ensured because settlement is **persistent**: migrants go where their community already is $\Rightarrow$ usually a very strong first stage
- Monotonicity is not guaranteed here

To check

- Depending on the approach, check:
 - Shares approach: which shares actually drive the estimate (**Rotemberg weights**)?
 - Shocks approach: are there **enough** shocks, are they plausibly as-good-as-random?
- Shares often correlate with a lot of things
- With hundreds of industries and many periods, “the exclusion restriction holds” is a **big** claim

Take-away message

A shift-share instrument is not exogenous because it is a shift-share instrument. Need to argue **which part** provides the identification and how

RDD: An example

Air quality alerts in Mexico City

- **Air quality alerts** in Mexico City on highly polluted days
 - Vehicle restrictions, industrial curbs, public messaging
- Which impact on pollution levels and health?
- Problem: cannot compare alert days (highly polluted) to “normal” days
- But alert triggered by a **rule**: ozone above **155 ppb** in any station (Aguilar-Gomez 2025)
 - So compare days just **above** 155 ppb to days just **below**

Why does that work?

Intuition

- The threshold is **arbitrary**: nothing happens to the atmosphere at 154.9 ppb
 - A 154 ppb day and a 156 ppb day are the **same kind of day**, except one triggers an alert
-
- Nothing else about a day changes at 155 ppb
 - *ie* **local continuity**
 - There are days just above and just below
 - *ie* **local support**
 - Crossing 155 ppb does raise the chance of an alert, by 74 pp
 - *ie* **relevance**. But here **not** mechanical $\Rightarrow$ fuzzy design
 - Crossing it only matters because it triggers the alert
 - *ie* **exclusion restriction**

Other examples of discontinuities?

- **Eligibility thresholds** to a program (based on income, age, test score)
- **Close elections**: winning with 50.1% rather than 49.9% of the votes
- **Spatial RDD**: being just on one side of a geographic administrative boundary
- **Class-size rules**: a class is split once it passes a given number of pupils

RDD: Fundamentals

In a nutshell

- Treatment is not randomly assigned, it is assigned by a **rule**:
 - A continuous **running variable** x_i crosses a **cutoff** c
 - $d_i = \mathbb{1}\{x_i \geq c\}$
 - eg a county is regulated if its pollution exceeds a federal threshold
- Estimand:

$$\beta_{RD} = \lim_{x \rightarrow c^+} \mathbb{E}[y_i | x_i = x] - \lim_{x \rightarrow c^-} \mathbb{E}[y_i | x_i = x] = \underbrace{\mathbb{E}[y_i(1) - y_i(0) | x_i = c]}_{\text{ATE at the cutoff}}$$

- Econometric model:

$$y_i = \alpha + \beta d_i + f_l(x_i) + f_r(x_i)d_i + e_i$$

Identifying assumptions

- **Local continuity**

- $\mathbb{E}[y_i^1|x_i]$ and $\mathbb{E}[y_i^0|x_i]$ are continuous at c
- Everything else varies smoothly at c
- Units just below c are a valid counterfactual for units just above

- **Relevance and local support**

- Treatment does jump at c , and there are units on both sides of it

Threats to local continuity

- Local continuity can break when:
 - Individuals can **sort** across the cutoff. That happens if:
 1. The rule is **known in advance**
 2. Individuals **want** to be on the other side
 3. Individuals **have time** to adjust
 - The cutoff is **endogenous**: several factors affecting the outcome vary at the threshold
 - **Heaping**: excess mass at particular values of the running variable (eg round numbers, that may coincide with the threshold)

No overlapping support

- A RDD **never** compares a treated and an untreated unit with the same x_i
 - By construction, they do not exist
- $\neq$ from other strategies: no CIA
- Counterfactual built by **extrapolating** to the other side of the cutoff
- Narrower bandwidth size $\Rightarrow$ less bias, more imprecise
- The **functional form** of f matters $\rightarrow$ global or high-order polynomials **overfit** (Gelman and Imbens 2019)

Locality

- An RDD identifies an effect **at the cutoff** $\Rightarrow$ local
- Says nothing about units far from the threshold
- Extrapolating away from c requires **extra assumptions**, not more data
- In a fuzzy RD two forms of locality combined: effect on compliers at the cutoff

Warning

The identification can be credible but the estimand can remain uninformative of what we care about

RDD: Fuzzy-RD

Sharp and fuzzy

	Sharp	Fuzzy.
What jumps at c	Treatment itself: $d_i = \mathbb{1}\{x_i \geq c\}$	The probability of treatment: $\Pr(d_i = 1 \mid x_i)$
Assumptions	Local continuity, relevance	+ exclusion restriction and local monotonicity
Estimand	$\mathbb{E}[y_i(1) - y_i(0) \mid x_i = c]$	$\mathbb{E}[y_i(1) - y_i(0) \mid x_i = c, d_i(1) > d_i(0)]$
	ATE at the cutoff	ATE on compliers at the cutoff
Estimator	The jump in y at c	Jump in y divided by the jump in d , ie 2SLS

- Crossing only makes the treatment **more likely**
 - eg individuals above the threshold are designated, but not all comply

Fuzzy RD as an IV

- The cutoff indicator $z_i := \mathbb{1}\{x_i \geq c\}$ is an **instrument** for d_i
- Link to what we discussed for IV:
 - Small jump in the treatment probability = **weak instrument**
 - Exclusion restriction must hold **at the cutoff**: does crossing c trigger anything else?

RDD: Challenges

Functional form near the cutoff

Raw data

Linear

Quadratic

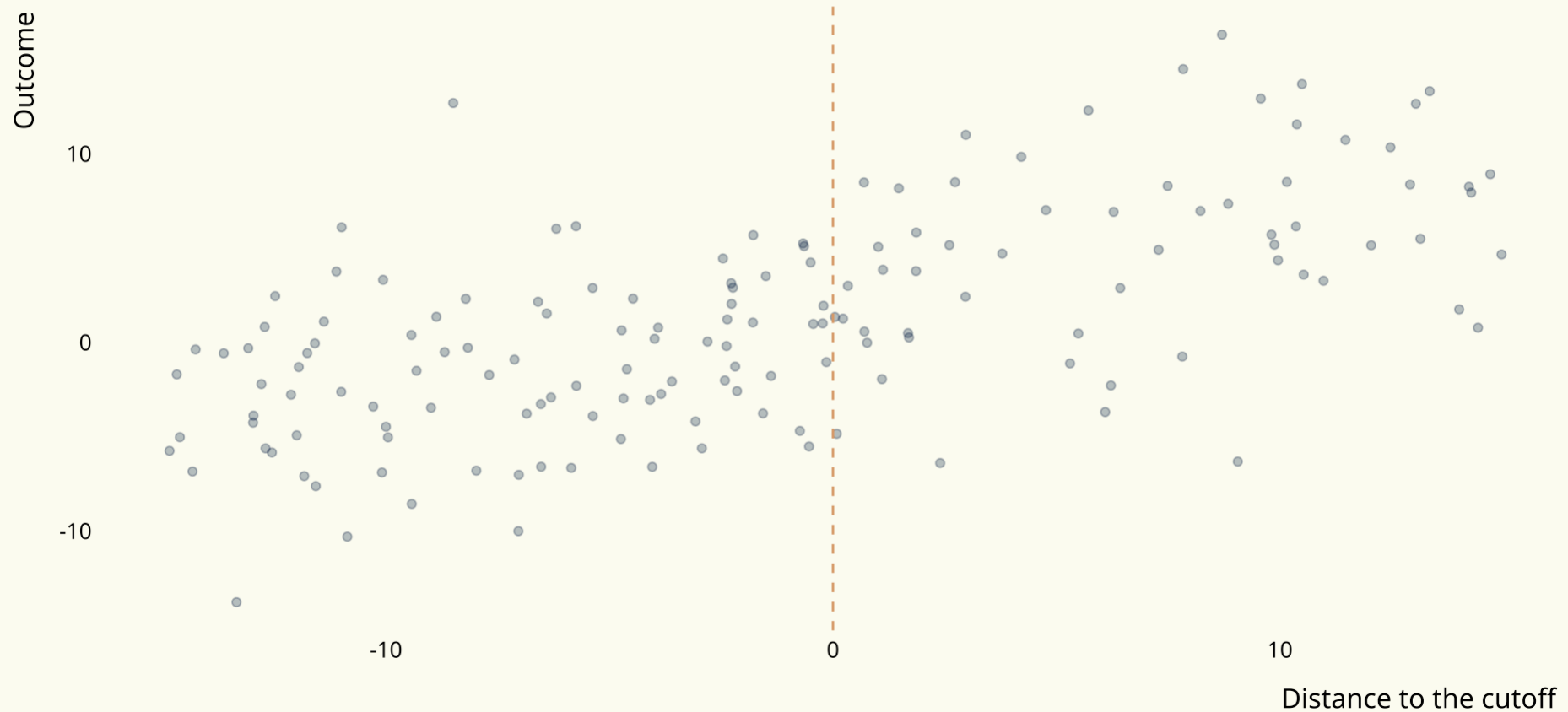
Flexible

Repeat

Code

Raw data

One draw, within a bandwidth of 15. The true jump is 3



Bandwidth size

- One of the main **design choices** in an RDD is choosing the bandwidth size h
- It arbitrates a **bias-variance trade-off** $\Rightarrow$ pick the h that minimises the MSE
- Select an **optimal bandwidth** (Imbens and Kalyanaraman 2012)
- h_{MSE} is optimal for the **point estimate**, not for inference: conventional CIs **undercover**
 - Use the **cer**rd bandwidth, or robust bias-corrected intervals
- Observations outside the bandwidth are typically manually filtered out pre-estimation. They can also be dropped by adding **observation-level** fixed effects.

Statistical power of RD designs

- Identification rests on observations **within the bandwidth**
- The effective sample can be small
- Stommes, Aronow, and Sävje (2023) study RD in political sciences and show that they:
 - **Bunch just above** conventional significance thresholds
 - Are **underpowered** to detect anything but large effects
 - Lead to **exaggeration**
- Conventional confidence intervals are **too narrow**
 - They ignore the bias of the local polynomial fit
 - Use **bias-corrected robust** intervals

Addressing heaping

- Running variables are often **rounded**
- eg birth weight piles up at round 100g values
- Observations exactly at round values may be different
- $\Rightarrow$ the units at the cutoff can a **selected** group
- Fix: **donut hole** RD, drop a small window around c and re-estimate
 - If the estimate moves a lot, it might indicate that the heaped observations were driving it

Checking the design

- Make a **binscatter** of the outcome against the running variable
 - small bins, no fit line as fit lines make readers see discontinuities that are not there (Korting et al. 2023)
- Continuity cannot be tested, but can be made credible
- Evaluate **sorting**: no jump in the density of x at c
 - **Plot** the density of the running variable
- Pre-treatment covariates should not jump at c
 - **Plot** the **density** of the running variable, and the **covariates**, against it
- No jump at **placebo cutoffs**, and no jump in **placebo outcomes**
- Report the **first-stage jump** in a fuzzy RD (plot it)
- Summary: **make a lot of pictures**

Summaries

Strengths and weaknesses - IV

- + **Compelling** identification strategy
- + Can use IVs to address attenuation bias that may result from measurement error in d
- – β_{IV} has finite sample bias. Stems from randomness in estimates of d
- – β_{IV} **less efficient** than OLS, precision further $\searrow$ with weak instruments
- – In many settings (e.g., non-linear d), 2SLS can be very biased

Strengths and weaknesses - RDD

- + Approximates a **local randomized experiment** around the cutoff
- + Relies on relatively **weak continuity assumptions** (in the sharp case)
- + Discontinuities are **everywhere** in economics: administrative rules, eligibility thresholds, borders
- – Identification rests on the **few observations close to the cutoff** $\Rightarrow$ often **underpowered** (Stommes, Aronow, and Sävje 2023)
- – No overlap $\Rightarrow$ results depend on **functional form** and bandwidth choices
- – Estimates are **highly local**: limited external validity
- – Fuzzy RDs add the IV assumptions on top, and are especially imprecise when the first-stage jump is small

! Take away messages

- An IV only uses the variation in x **explained by the instruments**
- Both IV and RDD estimates are **local**
- Weak instruments should be avoided, and $F > 10$ is not a licence
- A weak instrument also **amplifies** any violation of the exclusion restriction
- With covariates, 2SLS is often **not** a LATE on compliers
- A fuzzy RDD is just an IV
- No overlap in an RDD $\Rightarrow$ bandwidth and polynomial choices strongly affect the estimate
- Both IV and RDD identify **local** estimands: always ask *local to whom?*

References

- Aguilar-Gomez, Sandra. 2025. "Adaptation and Mitigation of Pollution: Evidence from Air Quality Warnings." *Journal of Environmental Economics and Management* 134 (November): 103215. <https://doi.org/10.1016/j.jeem.2025.103215>.
- Andrews, Isaiah, James H. Stock, and Liyang Sun. 2019. "Weak Instruments in Instrumental Variables Regression: Theory and Practice." *Annual Review of Economics* 11 (1): 727–53. <https://doi.org/10.1146/annurev-economics-080218-025643>.
- Arceo, Eva, Rema Hanna, and Paulina Oliva. 2016. "Does the Effect of Pollution on Infant Mortality Differ Between Developing and Developed Countries? Evidence from Mexico City." *The Economic Journal* 126 (591): 257–80. <https://doi.org/10.1111/eoj.12273>.
- Bartik, Timothy J. 1991. *Who Benefits from State and Local Economic Development Policies?* W.E. Upjohn Institute. <https://doi.org/10.17848/9780585223940>.
- Blandhol, Christine, John Bonney, Magne Mogstad, and Alexander Torgovitsky. 2026. "When Is TSLS Actually LATE?" *The Review of Economic Studies*, May, rdag029. <https://doi.org/10.1093/restud/rdag029>.
- Bound, John, David A. Jaeger, and Regina M. Baker. 1995. "Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogeneous Explanatory Variable Is Weak." *Journal of the American Statistical Association* 90 (430): 443–50. <https://doi.org/10.2307/2291055>.
- Card, David. 1993. "Using Geographic Variation in College Proximity to Estimate the Return to Schooling." Working Paper 4483. Working Paper Series. National Bureau of Economic Research. <https://doi.org/10.3386/w4483>.
- Deryugina, Tatyana, Garth Heutel, Nolan H. Miller, David Molitor, and Julian Reif. 2019. "The Mortality and Medical Costs of Air Pollution: Evidence from Changes in Wind Direction." *American Economic Review* 109 (12): 4178–4219. <https://doi.org/10.1257/aer.20180279>.
- Finkelstein, Amy, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Heidi Allen, Katherine Baicker, and Oregon Health Study Group. 2012. "The Oregon Health Insurance Experiment: Evidence from the First Year." *The Quarterly Journal of Economics* 127 (3): 1057–1106. <https://doi.org/10.1093/qje/qjs020>.
- Gelman, Andrew, and Guido Imbens. 2019. "Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs." *Journal of Business & Economic Statistics* 37 (3): 447–56. <https://doi.org/10.1080/07350015.2017.1366909>.
- Imbens, Guido, and Karthik Kalyanaraman. 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *The Review of Economic Studies* 79 (3): 933–59. <https://www.jstor.org/stable/23261375>.
- Keane, Michael P., and Timothy Neal. 2024. "A Practical Guide to Weak Instruments." *Annual Review of Economics* 16 (Volume 16, 2024): 185–212. <https://doi.org/10.1146/annurev-economics-092123-111021>.
- Korting, Christina, Carl Lieberman, Jordan Matsudaira, Zhuan Pei, and Yi Shen. 2023. "Visual Inference and Graphical Representation in Regression Discontinuity Designs." *The Quarterly Journal of Economics* 138 (3): 1977–2019. <https://doi.org/10.1093/qje/qjad011>.

- Lee, David S., Justin McCrary, Marcelo J. Moreira, and Jack Porter. 2022. "Valid t-Ratio Inference for IV." *American Economic Review* 112 (10): 3260–90. <https://doi.org/10.1257/aer.20211063>.
- Mellon, Jonathan. 2024. "Rain, Rain, Go Away: 194 Potential Exclusion-Restriction Violations for Studies Using Weather as an Instrumental Variable." *American Journal of Political Science* n/a (n/a). <https://doi.org/10.1111/ajps.12894>.
- Moretti, Enrico, and Matthew Neidell. 2011. "Pollution, Health, and Avoidance Behavior: Evidence from the Ports of Los Angeles." *Journal of Human Resources* 46 (1): 154–75. <https://doi.org/10.1353/jhr.2011.0012>.
- Olea, José Luis Montiel, and Carolin Pflueger. 2013. "A Robust Test for Weak Instruments." *Journal of Business & Economic Statistics* 31 (3): 358–69. <https://www.jstor.org/stable/43702731>.
- Schlenker, Wolfram, and W. Reed Walker. 2016. "Airports, Air Pollution, and Contemporaneous Health." *The Review of Economic Studies* 83 (2): 768–809. <https://doi.org/10.1093/restud/rdv043>.
- Stommes, Drew, P. M. Aronow, and Fredrik Sävje. 2023. "On the Reliability of Published Findings Using the Regression Discontinuity Design in Political Science." *Research & Politics* 10 (2). <https://doi.org/10.1177/20531680231166457>.