

Causal Inference Review Section (1)

Shengqiao Lin

Jan 26 2023

Plan for this semester

- For each method
 - Replication codes/Best practice
 - Presentation/Visualization
 - Pitfalls/robustness
- Supplementary materials
- Assignments
- Final Presentation

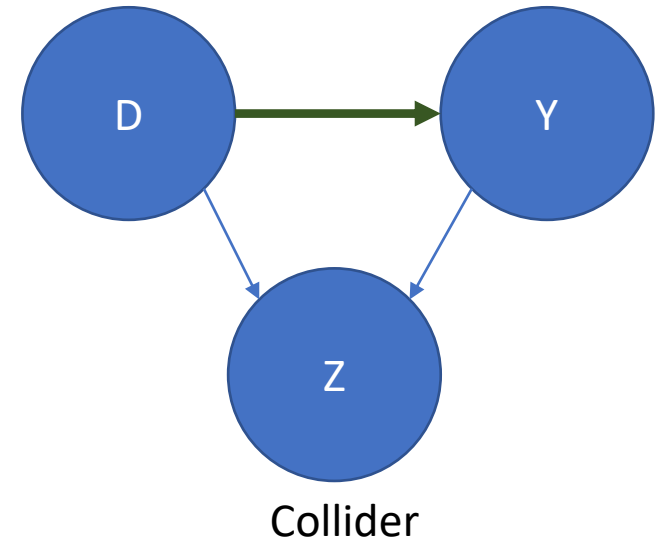
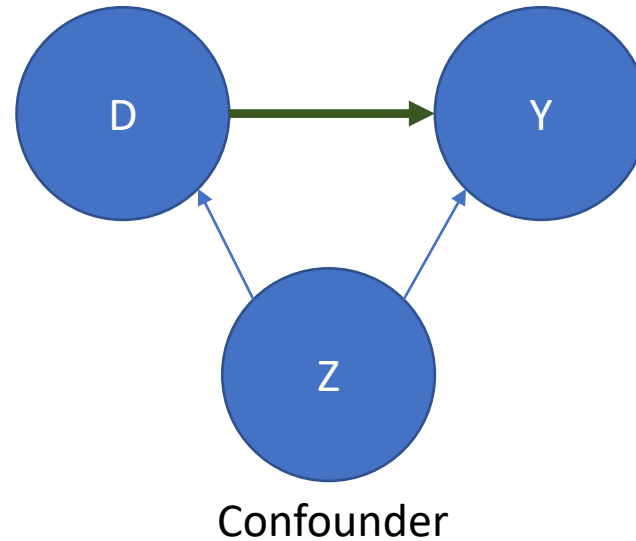


Plan for Today

- Covariates: What to control in Regressions (or not)
 - Confounders ✓
 - Colliders ✗
- Common mistakes: IV, RD, DID

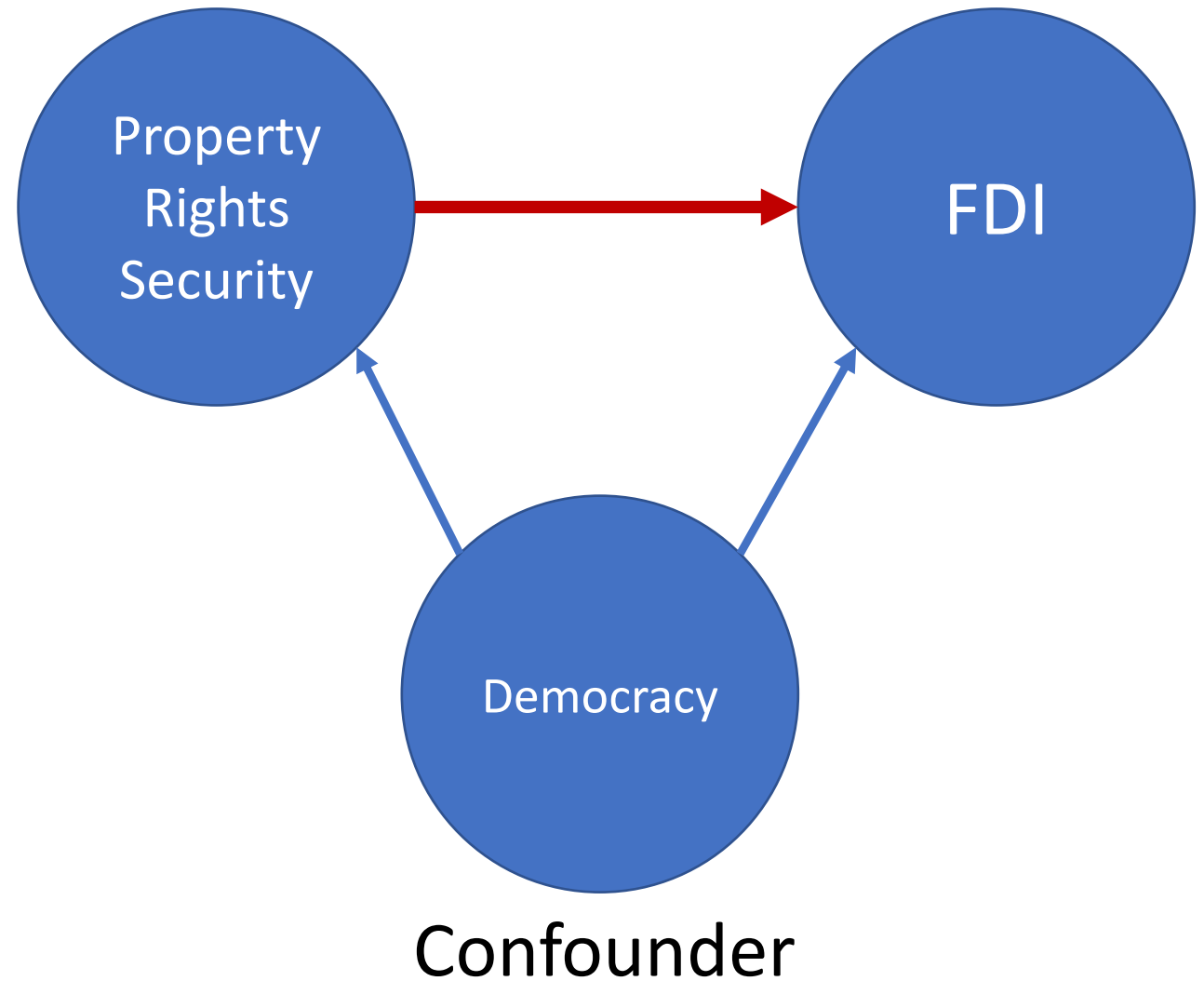
Two kinds of Covariates

- Covariates: associated with D and Y
- Confounder
 - $D \leftarrow Z \rightarrow Y$
- Collider
 - $D \rightarrow Z \leftarrow Y$



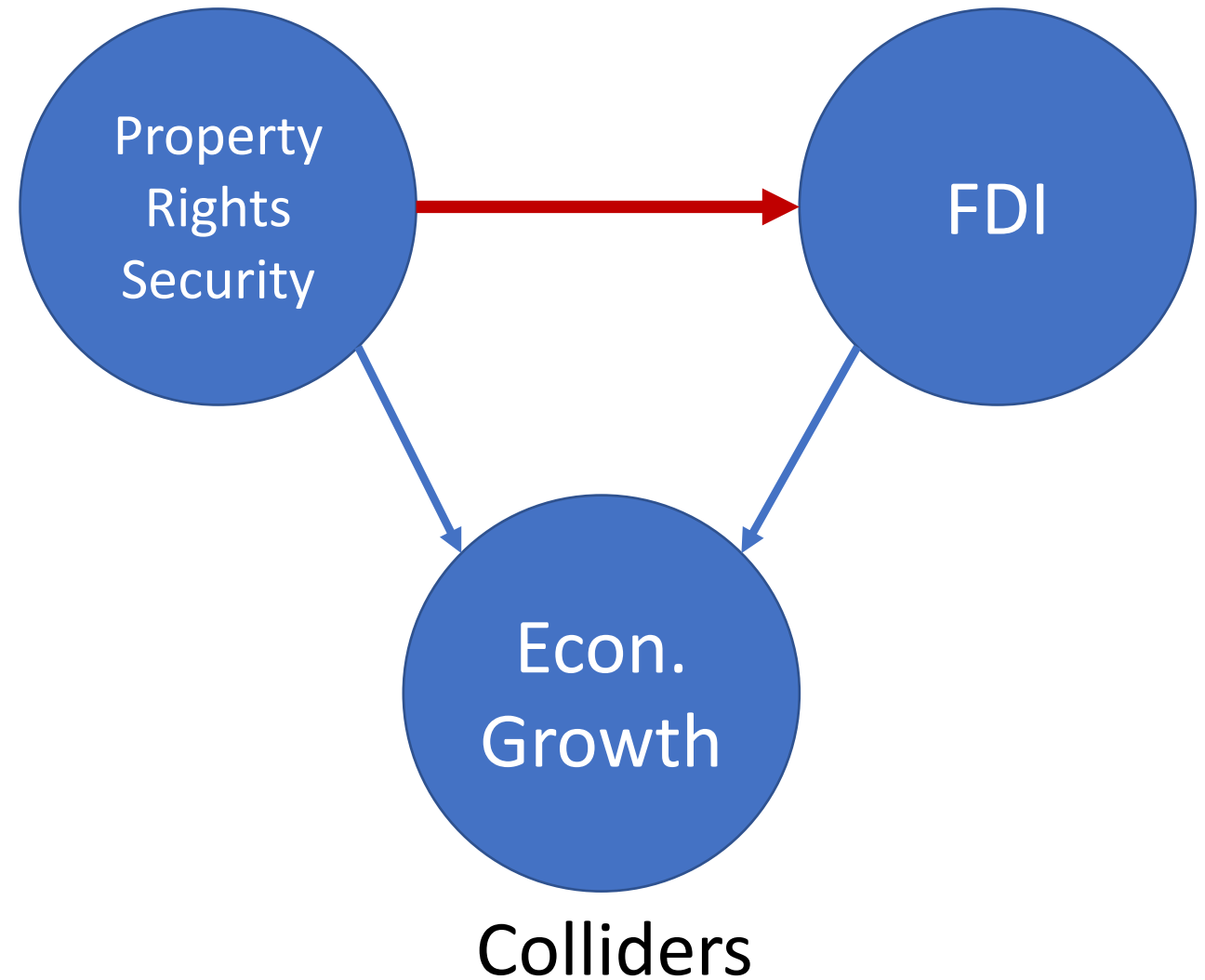
Example 1

- D: Property rights security
 - Z: Democracy
 - Y: FDI
-
- $Y \leftarrow 0.3 \cdot Z + 0.4 \cdot D + u_1$
 - $D \leftarrow 0.5 \cdot Z + u_2$

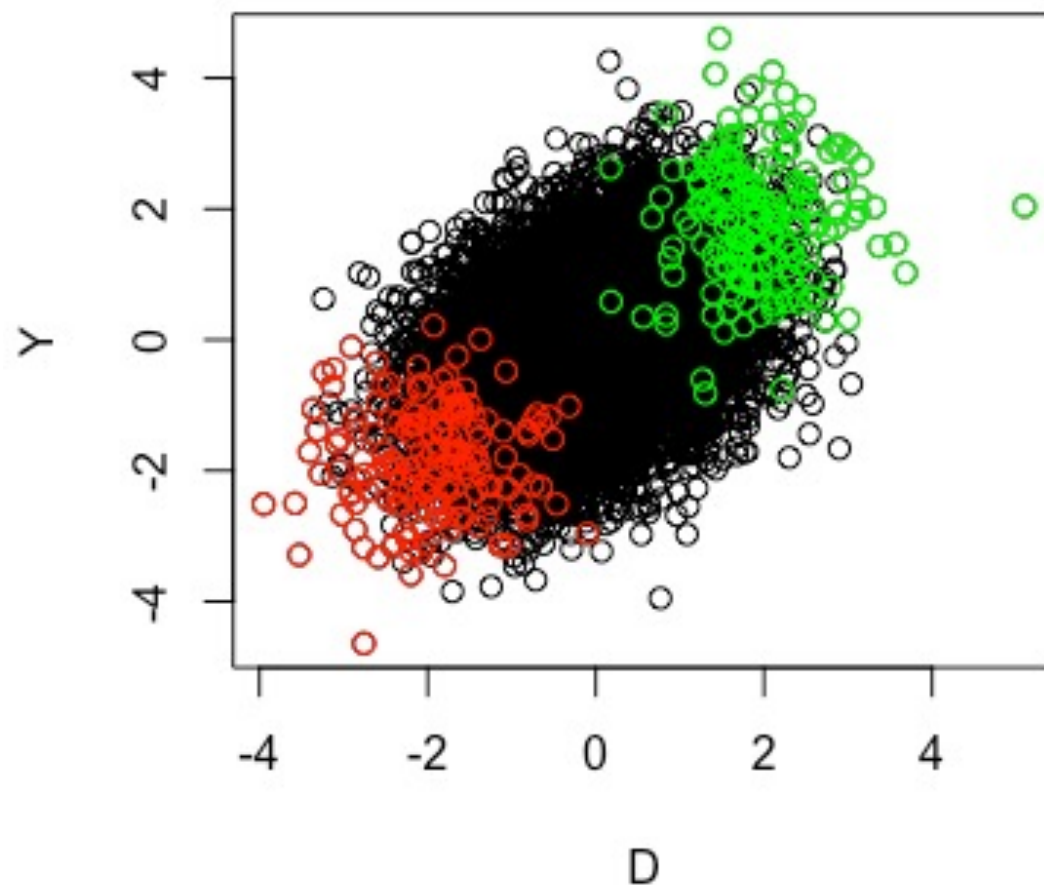
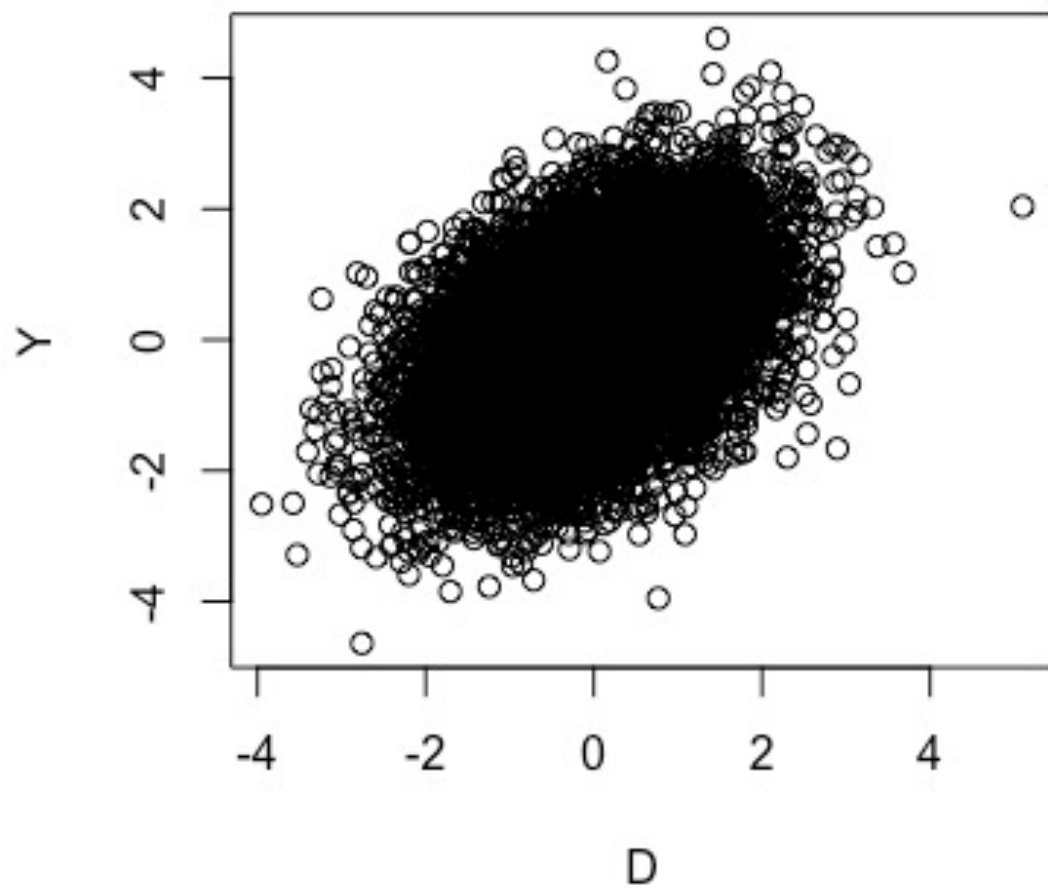


Example2

- D: Property rights security
 - Z: Economic growth
 - Y: FDI
-
- $Y \leftarrow 0.4 * D + u1$
 - $Z \leftarrow 0.5 * D + 0.3 * Y + u2$
-
- *Post-treatment outcomes

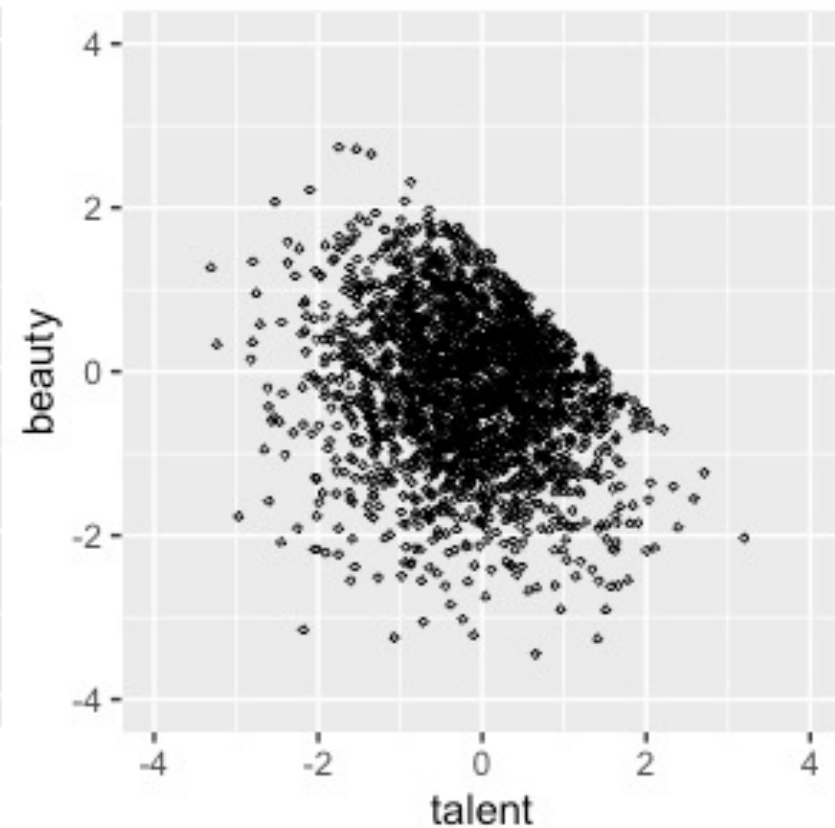
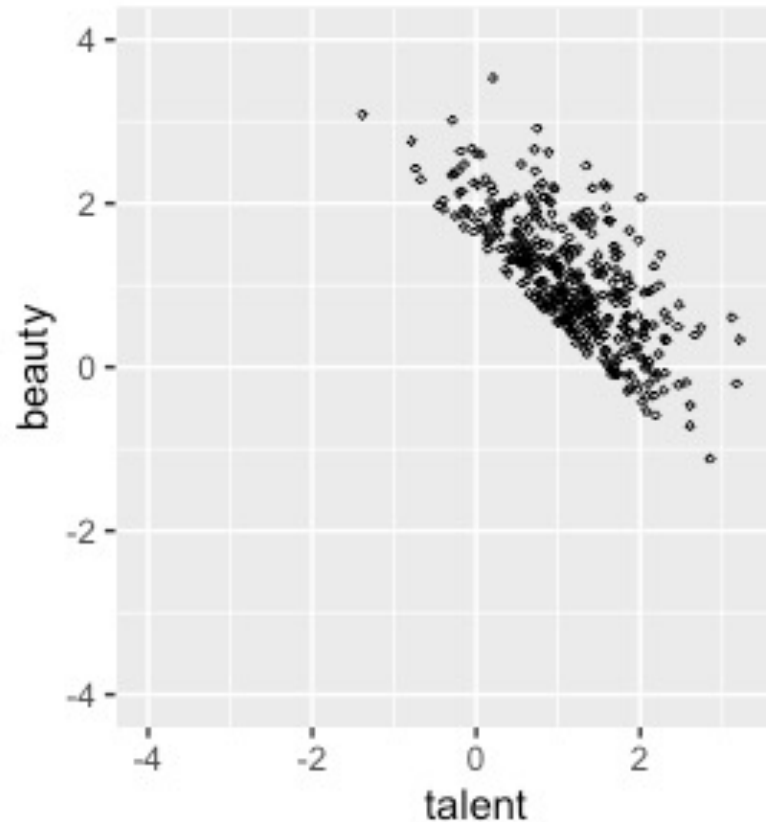
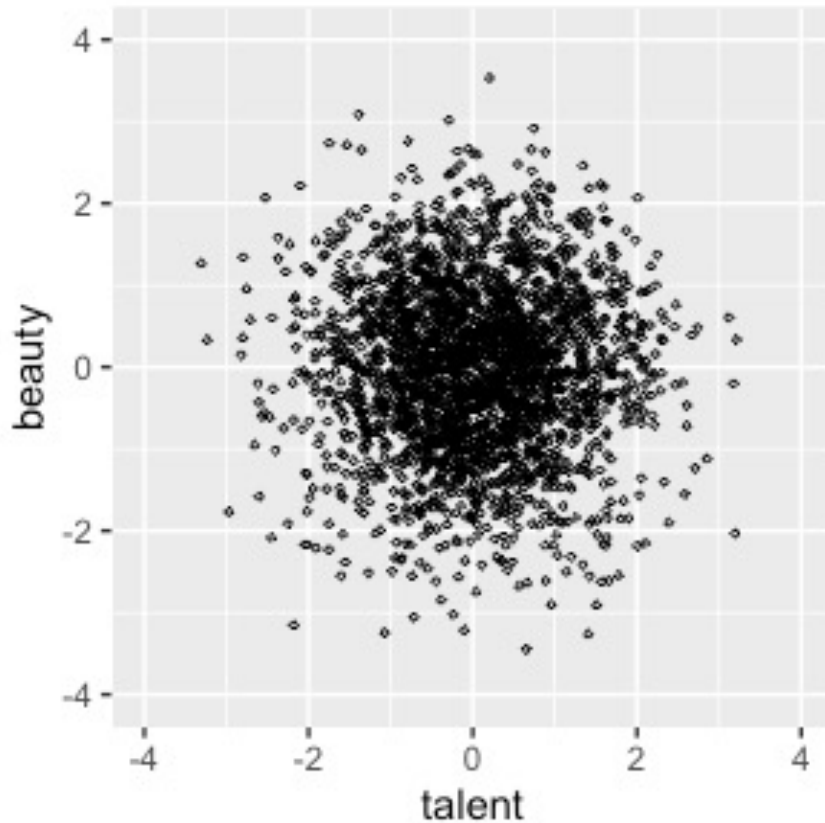


Example2



Example 3: Beauty or Talents

- Beauty and talents are randomly assigned
- A star means a combination of both $>85\%$



Some common mistakes

- IV:
 - Weak Instrument
 - Violating exclusion restriction
- RDD:
 - Sorting behavior/Self-selection
- DID
 - Problematic parallel trend
 - Multiple treatments
 - Sorting behavior

Instrumental variable

- $Z \rightarrow X \rightarrow Y$
- Weak Instrument
 - Z cannot predict X (well)
 - Rule of Thumb: $F > 10$
- Exclusion Restriction
 - Z can only affect Y through X
 - Justification/checks needed
- Source: Lal et al. (2022) How Much Should We Trust Instrumental Variable Estimation in Political Science? Practical Advice Based on Over 60 Replicated Studies*
- Replicated articles: APSR, AJPS, JOP, 2010-2020

TABLE 3. SUMMARY OF REPLICATION RESULTS

(%)	APSR (13)	AJPS (23)	JOP (25)	All (61)
<i>Panel A</i>				
First-stage F Statistic Unreported	13.3	20.8	28.0	21.9
F Statistic Miscalculated	15.4	47.4	50.0	40.0
Replicated F Statistic ≥ 10.0	86.7	79.2	80.0	81.2
Replicated F Statistic ≥ 104.7	33.3	37.5	24.0	31.2
<i>Panel B</i>				
Reported SE $< 0.8 \times$ Bootstrapped SE	26.7	33.3	40.0	34.4
Reported SE $< 0.5 \times$ Bootstrapped SE	26.7	20.8	28.0	25.0
Reported SE $< 0.3 \times$ Bootstrapped SE	13.3	16.7	12.0	14.1
Replicated $ z \geq 1.96$	53.3	62.5	60.0	59.4
Replicated $ z \geq 3.43$	33.3	29.2	8.0	21.9
<i>Panel C</i>				
$ \hat{\beta}_{2SLS}/\hat{\beta}_{OLS} > 1$	86.7	91.7	96.0	92.2
$ \hat{\beta}_{2SLS}/\hat{\beta}_{OLS} > 3$	46.7	41.7	52.0	46.9
$ \hat{\beta}_{2SLS}/\hat{\beta}_{OLS} > 5$	33.3	33.3	32.0	32.8
$ \hat{\beta}_{2SLS}/\hat{\beta}_{OLS} > 10$	13.3	16.7	16.0	15.6

Note: “ F Statistic Miscalculated” is based on a subset of studies that report first-stage F statistics and defined as reported first-stage partial F being at least 30% smaller than the bootstrapped F statistics. Lee et al. (2020) show that, to have a correct test size at 5%, researchers need to require either $F \geq 104.7$ or $|z| \geq 3.43$.

TABLE 2. TYPES OF IVs

IV Type	#Papers	Percentage%
Theory	40	62.5
Geography/climate/weather	10.5	16.4
History	10	15.6
Treatment diffusion	2.5	3.9
Others	17	26.6
Experiment	12	18.8
Rules & policy changes	5	7.8
Change in exposure	3	4.7
Fuzzy RD	2	3.1
Econometrics	7	10.9
Interactions/“Bartik”	5	7.8
Lagged treatment	1	1.6
Empirical test	1	1.6
Total	64	100.0

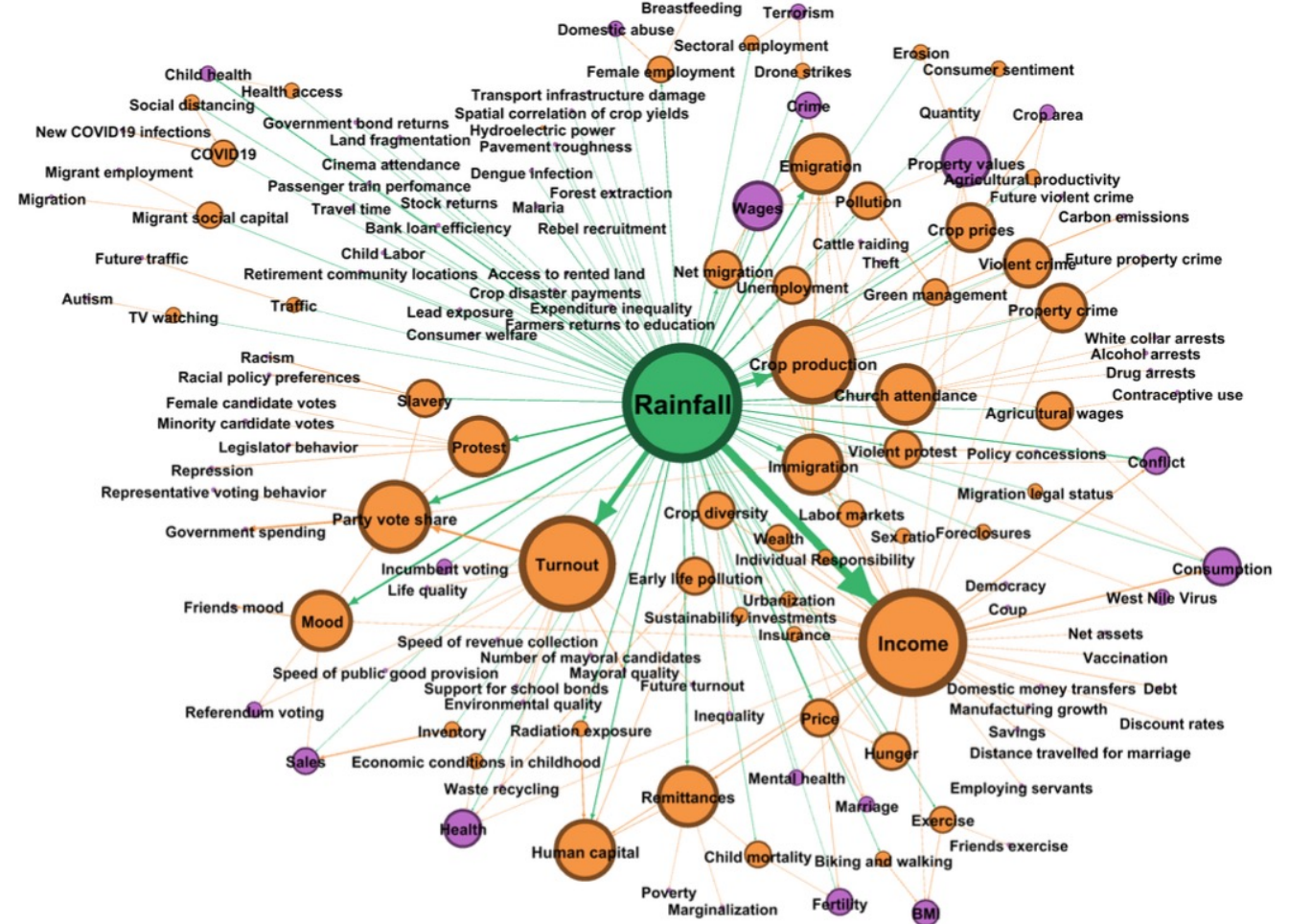
Note: One paper uses both geography-based instruments and an instrument based on treatment diffusion from neighbors. We count 0.5 for each category.

Some Examples for IV

- Geography:
 - Chen, T., Kung, J.K.S. and Ma, C., 2020. Long live Keju! The persistent effects of China's civil examination system. *The Economic Journal*, 130(631), pp.2030-2064
 - Ancient civil exam system -> human capital today
 - Distance to the nearest pine and bamboo forests (ingredient for textbooks/pens)
- History:
 - Guiso, Luigi, Paola Sapienza and Luigi Zingales. 2016. *Long-term persistence*. *Journal of the European Economic Association* 14(6):1401–1436.
 - Free city-state in the Middle Ages -> social capital today
 - Presence of a bishop -> free city-state
 - Rogowski et al. 2021. Public Infrastructure and Economic Development: Evidence from Postal Systems. *AJPS*.
 - Copartisan candidates with the president won -> postal systems (as robustness checks)
- But these are time-invariant -> That's why weather becomes popular IVs.

Rainfall as IV

- Miguel et al. (JPE2004, Economic Shocks and Civil Conflicts)
 - Rainfall-> Economic shock -> Civil conflicts
 - Discussion on potential violations
 - Stats/theory/history
- Wasow (APSR2020, Agenda Seeding: How 1960s Black Protests Moved Elites, Public Opinion and Voting)
 - Rainfall -> violent movement -> vote share
- Rainfall is super popular. If we link all together....
 - Rainfall -> many X -> many Y -> many other Y



Source: Mellon, 2022, Rain, Rain, Go Away: 192 Potential Exclusion-Restriction Violations for Studies Using Weather as an Instrumental Variable

Regression Discontinuity Design

- The continuity assumption:
 - Running factors -> Discontinuous jump (treatment status)
 - Sometimes it is not as-if random (de la Cuesta and Imai, ARPS 2016)
 - No incentive/ability to change the treatment status
 - Sorting behavior:
 - One might manipulate their running factors to change the treatment status
 - E.g., electoral frauds in close elections
 - Justification: incentive/ability

Difference-in-differences

- Parallel trends:
 - Treated group will evolve as the control group if being untreated
 - Pre-treatment trend as suggestive evidence for post-treatment trend
 - Violations: e.g., economic growth among countries
- Multiple treatments
 - A bunch of X: policy changes
- Sorting behavior/re-location
 - Incentive/capacity to change treatment status
 - E.g. Workers in specific industries
- *Recommend: Liu, L., Wang, Y. and Xu, Y. (2023), A Practical Guide to Counterfactual Estimators for Causal Inference with Time-Series Cross-Sectional Data. American Journal of Political Science. <https://doi.org/10.1111/ajps.12723>*

A few combination

- PSM+DID
- RD+IV
- RD+DID
- IV+DID
- Synthetic Control Method + DID

Try ChatGPT



Which is the most popular in top journals of political science



It is difficult to say which causal inference method is the most popular in top journals of political science, as the choice of method often depends on the specific research question and data at hand. However, some methods are more commonly used than others in political science research.



In recent years, the most popular causal inference method in political science has been the difference-in-differences (DiD) method. This method has been widely used to estimate the causal effect of various policies and interventions, such as the effects of welfare programs and education policies on economic outcomes.

Another popular method used in political science is the regression discontinuity design (RDD). This method is often used to estimate the causal effect of policies and interventions that are assigned based on a threshold or cutoff, such as the effects of campaign spending on election outcomes.

Instrumental variable (IV) method and Propensity score matching (PSM) are also commonly used in political science research. They are mostly used when there is a concern of bias caused by unobserved confounding factors.

It is worth noting that the choice of causal inference method also depends on the type of data available, and the type of question you want to answer. Researchers should carefully consider the assumptions and limitations of each method when choosing which one to use.



What are the most common mistakes that people made when using causal inference methods in political science research?



Causal inference methods can be complex and there are many potential pitfalls to be aware of when using them in political science research. Here are a few of the most common mistakes that researchers make when using causal inference methods:



1. **Not considering the assumptions of the method:** Many causal inference methods make specific assumptions about the data and the underlying causal relationships. Researchers should carefully consider these assumptions and be aware of any potential violations, as these can lead to biased or incorrect estimates of the treatment effect.
2. **Not accounting for confounding variables:** Confounding variables are variables that are related to both the treatment and the outcome, and can lead to biased estimates of the treatment effect if they are not controlled for. Researchers should be mindful of potential confounding variables and use appropriate methods to control for them, such as matching or stratification.
3. **Not considering the sample size and power:** Sample size and power are important considerations when using causal inference methods. Researchers should be aware of the number of units and the distribution of the data, and how this may affect the precision and accuracy of the estimates.
4. **Not using appropriate estimation techniques:** There are many different estimation techniques that can be used for causal inference, and researchers should choose the most appropriate one for their data and research question. For example, when there's a discontinuity in the assignment of treatment, RDD is appropriate, but if there's randomness, then propensity score matching (PSM) or inverse probability weighting (IPW) might be better.
5. **Not interpreting the results correctly:** The results of causal inference methods can be complex and difficult to interpret. Researchers should be careful not to overstate the conclusions or generalize the findings beyond the data and the assumptions of the method.

It's important to note that causal inference methods are a powerful tool for political science research, but they require careful consideration and a solid understanding of the underlying assumptions and limitations. Researchers should be mindful of potential mistakes, and be transparent about the assumptions, limitations and potential biases in their research.

Regenerate response