Difference-in-Differences

Data Science and Causal Inference Workshop 2025

Dr Christian Engels

START RECORDING NOW

University of St Andrews | FOUNDED 1413

4 September 2025

Importance of Empirical Research

University of St Andrews

- The availability of richer datasets and the advances in computational power have changed Social Sciences during the last 40 years.
- Currie, Kleven and Zwiers (2020) show that the fraction of empirical research keeps rising.
- A very common goal of empirical research is to uncover/highlight the causal effect of a given policy intervention.

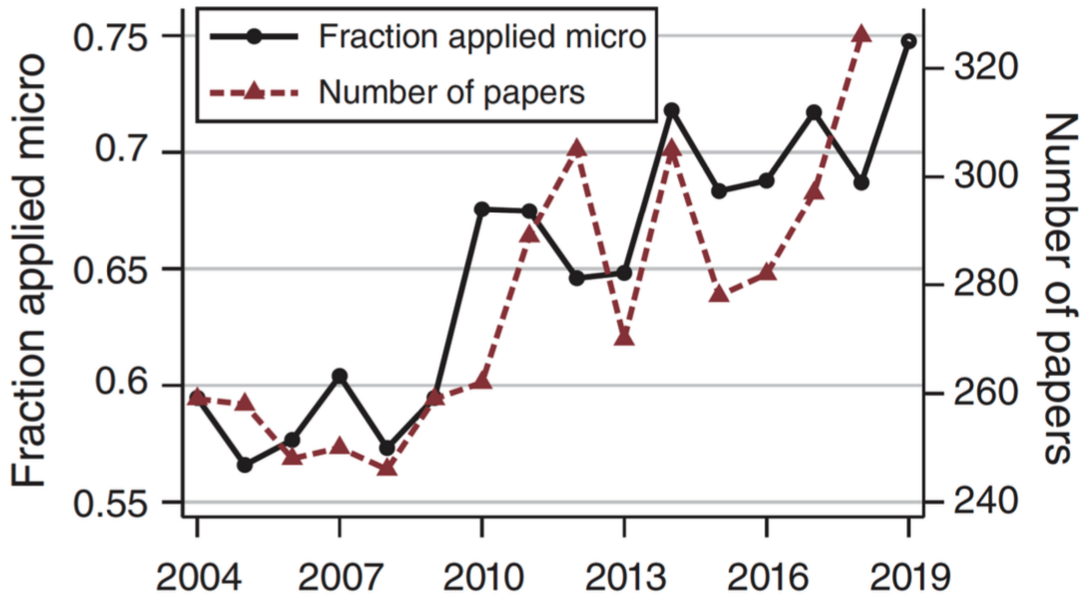


FIGURE 1. APPLIED MICROECONOMICS ARTICLES IN TOP-FIVE JOURNALS

Note: This figure shows the fraction of papers in top-five journals that report an applied microeconomics JEL code (left axis) and the total number of papers in the top-five journals (right axis).

The boom of (quasi-)experimental methods

Currie et al. (2020) documented this change well

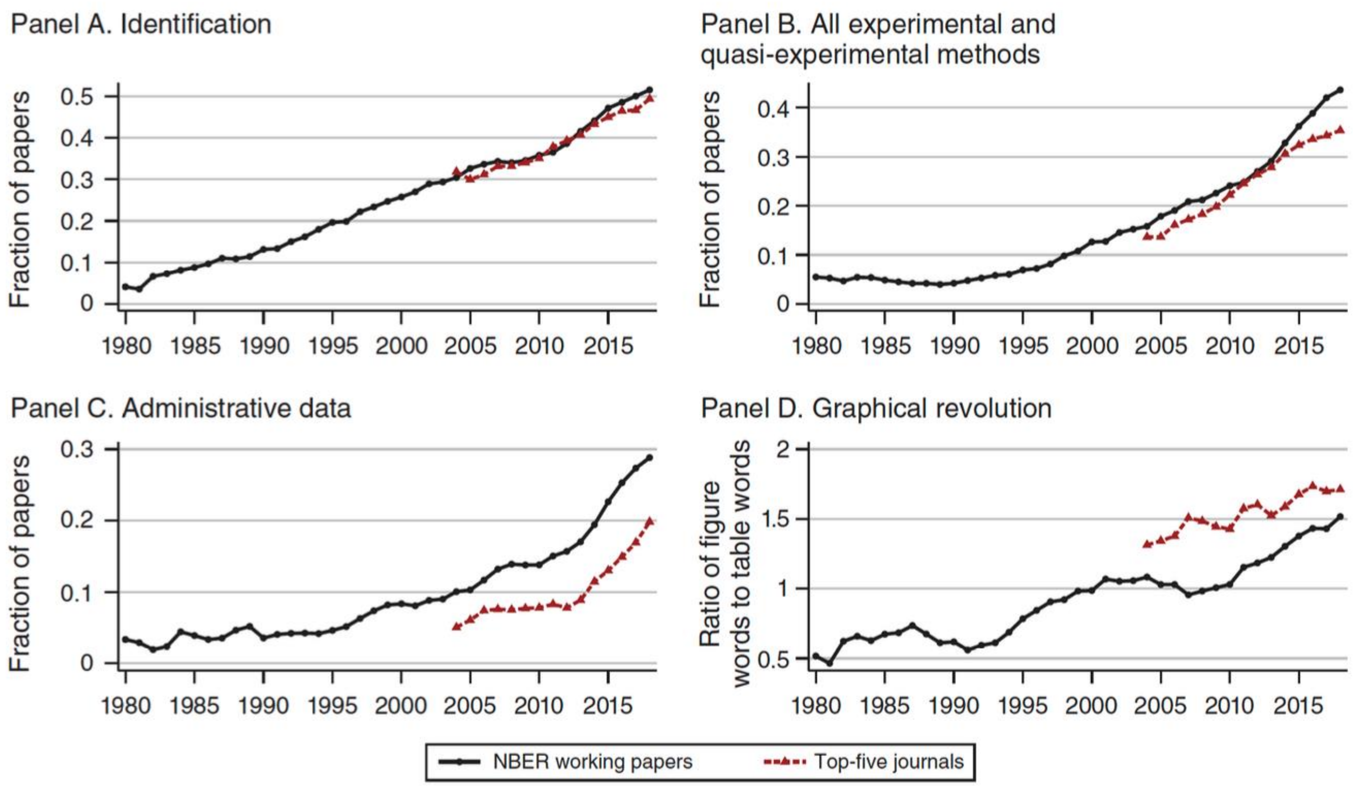


FIGURE 2. THE CREDIBILITY REVOLUTION

Notes: This figure shows different dimensions of the "credibility revolution" in economics: identification (panel A), all experimental and quasi-experimental methods (panel B), administrative data (panel C), and the graphical revolution (panel D). Panel D shows the ratio of the number of "figure" terms to the number of "table" terms mentioned. See Table A.I for a list of terms. The series show five-year moving averages.

What about experiments (or A/B tests)?

Currie et al. (2020)

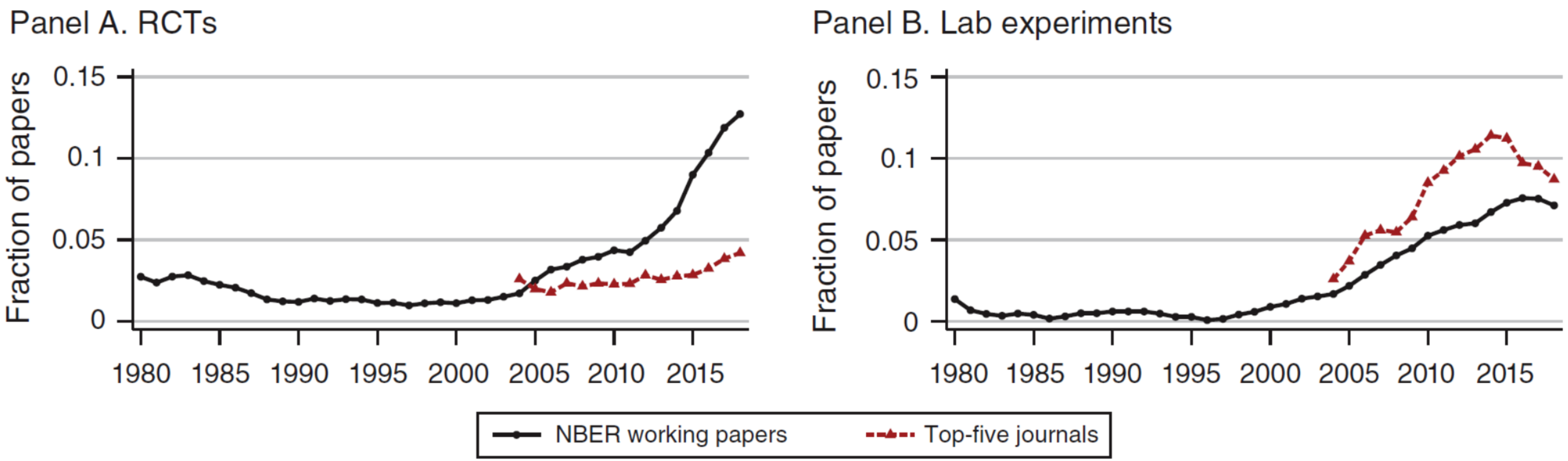


Figure 3. Experimental Methods

Notes: This figure shows the fraction of papers referring to each type of experiment. See Table A.I for a list of terms. The series show five-year moving averages.

Popularity of Difference-in-Differences methods

Currie et al. (2020)

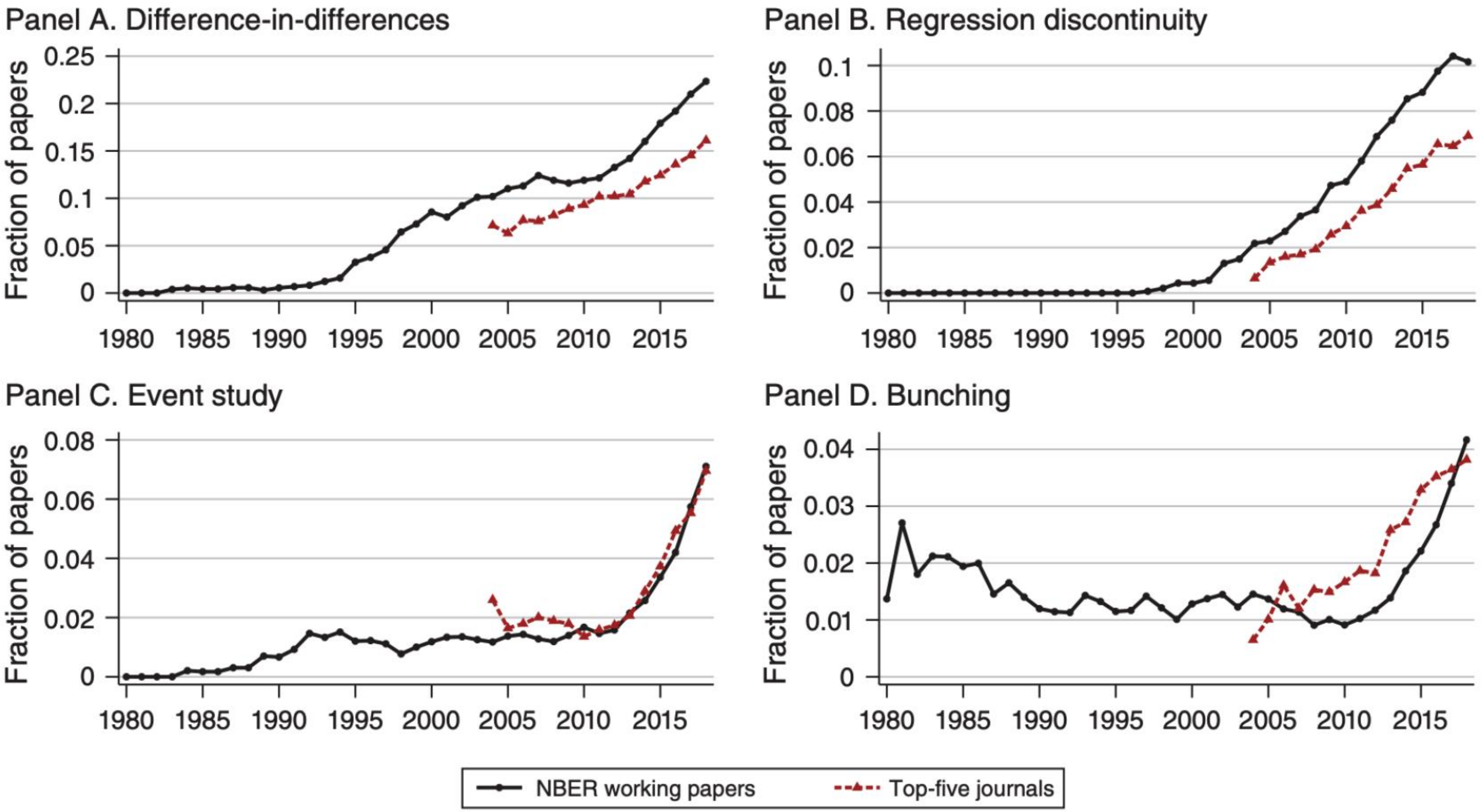


Figure 4. Quasi-experimental Methods

Recent popularity of DiD methods in empirical work

Goldsmith-Pinkham (2024) built on Currie et al. (2020) and updated the analysis using NBER working papers data that ends in May 2024.

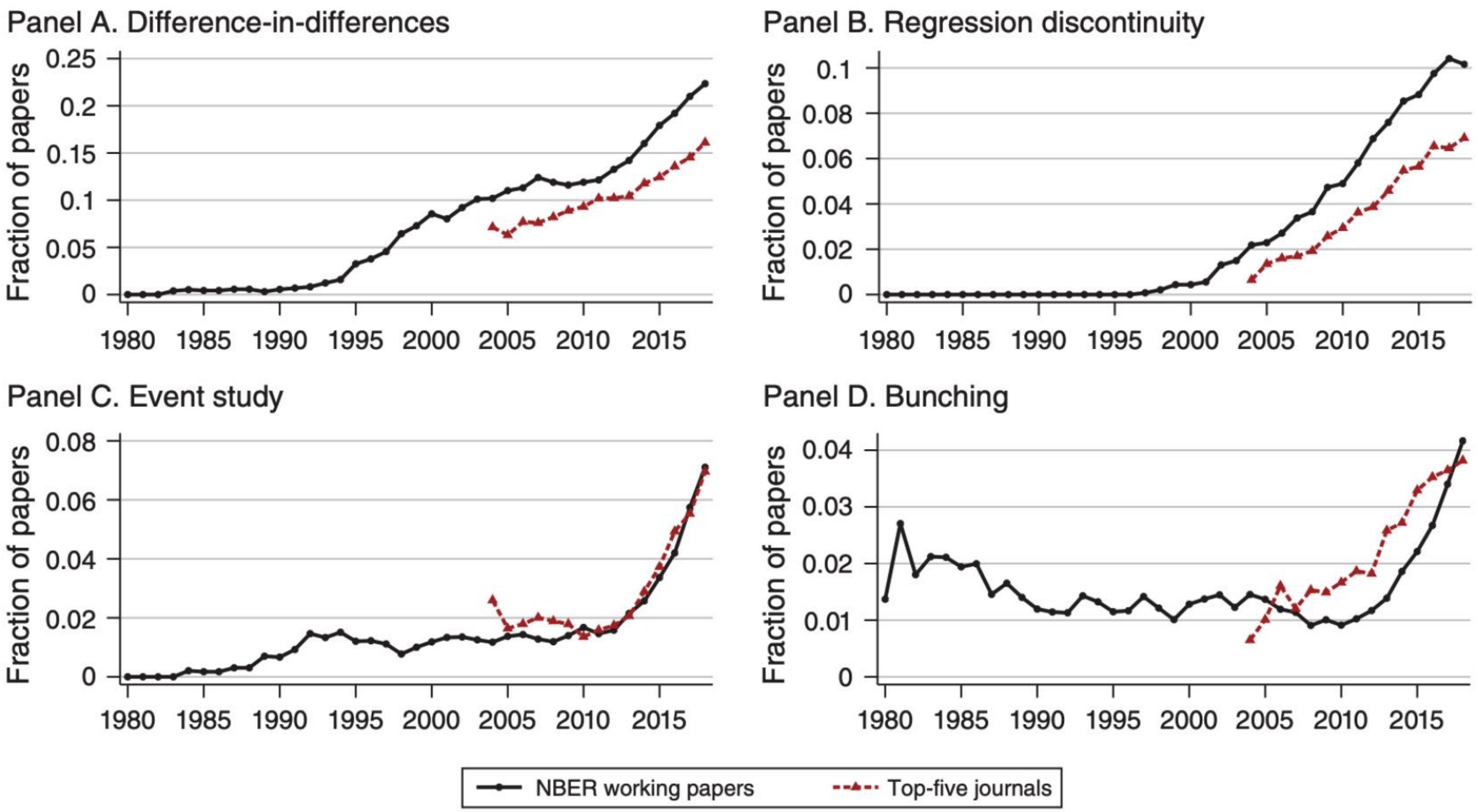
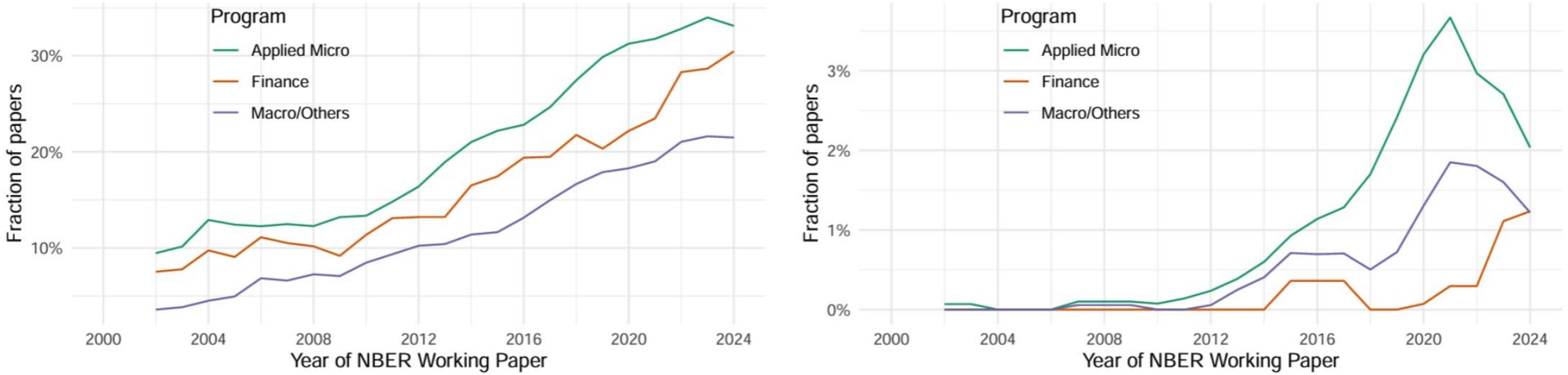


Figure 4. Quasi-experimental Methods

Popularity of Difference-in-Differences methods: by fields

Goldsmith-Pinkham (2024): the popularity of DiD by fields

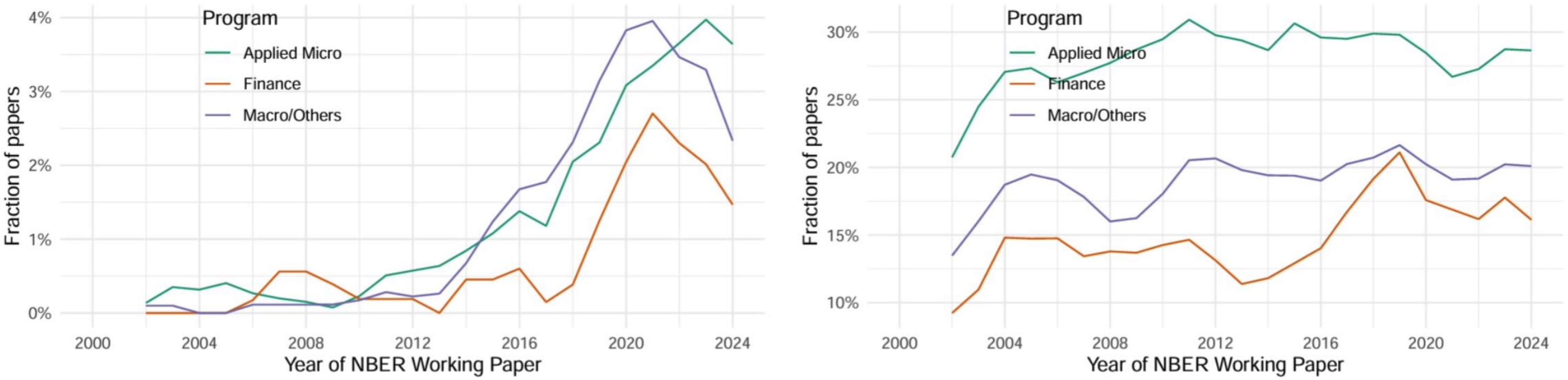


(a) Difference-in-differences (b) Synthetic controls

Figure 5: Panel (a) reports the share of papers that mention difference-in-differences or event studies. Figure (b) reports the share of papers that mention synthetic controls (this includes both synthetic difference-in-differences and synthetic control methods). See Table 2 for the breakdown of fields, and the Appendix for definitions on keywords.

Popularity of Difference-in-Differences methods

Goldsmith-Pinkham (2024): Compare previous plot with IV



(a) Bartik and shift-share instruments

(b) Instrumental variables

Figure 6: Panel (a) reports the share of papers that mention Bartik or shift-share instruments. Figure (b) reports the share of papers that mention instrumental variables. See Table 2 for the breakdown of fields, and the Appendix for definitions on keywords.

Causality with Observational Data: What can we do?

University of St Andrews

- In many applications, we do not have access to experimental data.
- Without an experiment, we will rely on observational data.
- With observational data, we have no choice but rely on assumptions to conduct causal inference.
- Different methods rely on different assumptions.
- Our job as researchers is to assess the pros and cons of each method in their ability to answer the questions we (and the business/policy makers/stakeholders) care about.

Causality with Observational Data: What else?

University of St Andrews

- DiD is very popular.
- Why?
- Data requirements, availability of tools to assess the plausibility of assumptions and easy-to use software.
- What are the main alternatives to DiD?
  1. Rely on unconfoundedness and leverage **regression, matching, re-weighting or double machine learning**.

     Drawback: Rule out selection on unobservables.

     We need to have data on everything that affects treatment timing and outcome of interest (unconfoundedness assumption).

Causality with Observational Data: What else?

University of St Andrews

- What are the main alternatives to DiD?

  2. Rely on **Pre-Post analysis**

     Drawback: Does not account for potential trends in outcomes.

     This is more reasonable if we study very short-run effects, but that is not usually the case.

The appeal of Difference-in-Differences

- DiD methods exploit variation in time (before vs. after) and across groups (treated vs. untreated) to recover causal effects of interest.
- **DiD combines previous approaches to avoid their pitfalls.**
- Advantage: Allow for selection on unobservables and time-trends.

  Not magic: We need to assume that, absent the treatment and conditional on covariates (features), the outcome of interest would evolve similarly across groups/cohorts — **Parallel Trends assumption**.

Parallel Trends needs to be discussed and its plausibility assessed!

- **Data Requirements:** We need data from periods before and after treatment to use DiD (and some periods where no unit is treated).

Some DiD Examples

Beck, Levine and Levkov (2010): **Effect of bank branching deregulation on income distribution in the US**.

- Exploit staggered bank deregulation across states to understand its effect on the Gini index (among other outcomes); see also Baker, Larcker and Wang (2022).

Venkataramani, Shah, O'Brien, Kawachi and Tsai (2017): **Effect of US Deferred Action for Childhood Arrivals (DACA) immigration program on health outcomes**.

- Compared changes in health outcomes among individuals who met key DACA eligibility criteria (based on age at immigration and at the time of policy implementation) before and after program implementation versus changes in outcomes for individuals who did not meet these criteria.

Some DiD Examples

Card and Krueger (1994): **Effect of minimum wage on employment.**

- Compared the changes in wages, employment, and prices at stores in New Jersey (increased minimum wage) relative to stores in Pennsylvania (minimum wage remained fixed).

Dube, Lester and Reich (2010); Dube, William Lester and Reich (2016), Callaway and Sant'Anna (2021) and many others: **Effect of minimum wage on different measures of employment**

- Callaway and Sant'Anna (2021) exploit variation in the timing of state minimum wage changes to understand its effect on teen employment.

Some DiD Examples

Meyer, Viscusi and Durbin (1995): **Effect of weekly benefit amount on time out of work due to injury**.

- They compared high-earnings (affected by the policy change) and low-earnings (not affected by the policy change) individuals injured before and after increases in the maximum weekly benefit amount. Estimated effects in Kentucky and Michigan.

Malesky, Nguyen and Tran (2014): **Effect of government re-centralization in Vietnam on public services**.

- They compared provinces (and districts) that abolished elected councils in Vietnam to other provinces that did not abolish them, before and after the re-centralization. Analyzed 30 outcomes.

Some DiD Examples

Carey, Miller and Wherry (2020): **Effect of Medicaid expansion on access to care and utilization for those who are already insured**.

- They compare different insurance coverage and health care utilization measures among states that opted to expand Medicaid eligibility in 2014 or 2015 with those that did not expand by 2015, before and after the expansion.

Assunção, Gandour, Rocha and Rocha (2020): **Effect of rural credit on deforestation**.

- Compared municipalities within the Amazon biome (concession of subsidized rural credit for them are conditional on stricter requirements since 2008), with municipalities outside the border of the Amazon biome (not affected by the policy change), before and after the policy.

The canonical Difference-in-Differences estimator

The **canonical DiD** estimator is given by:

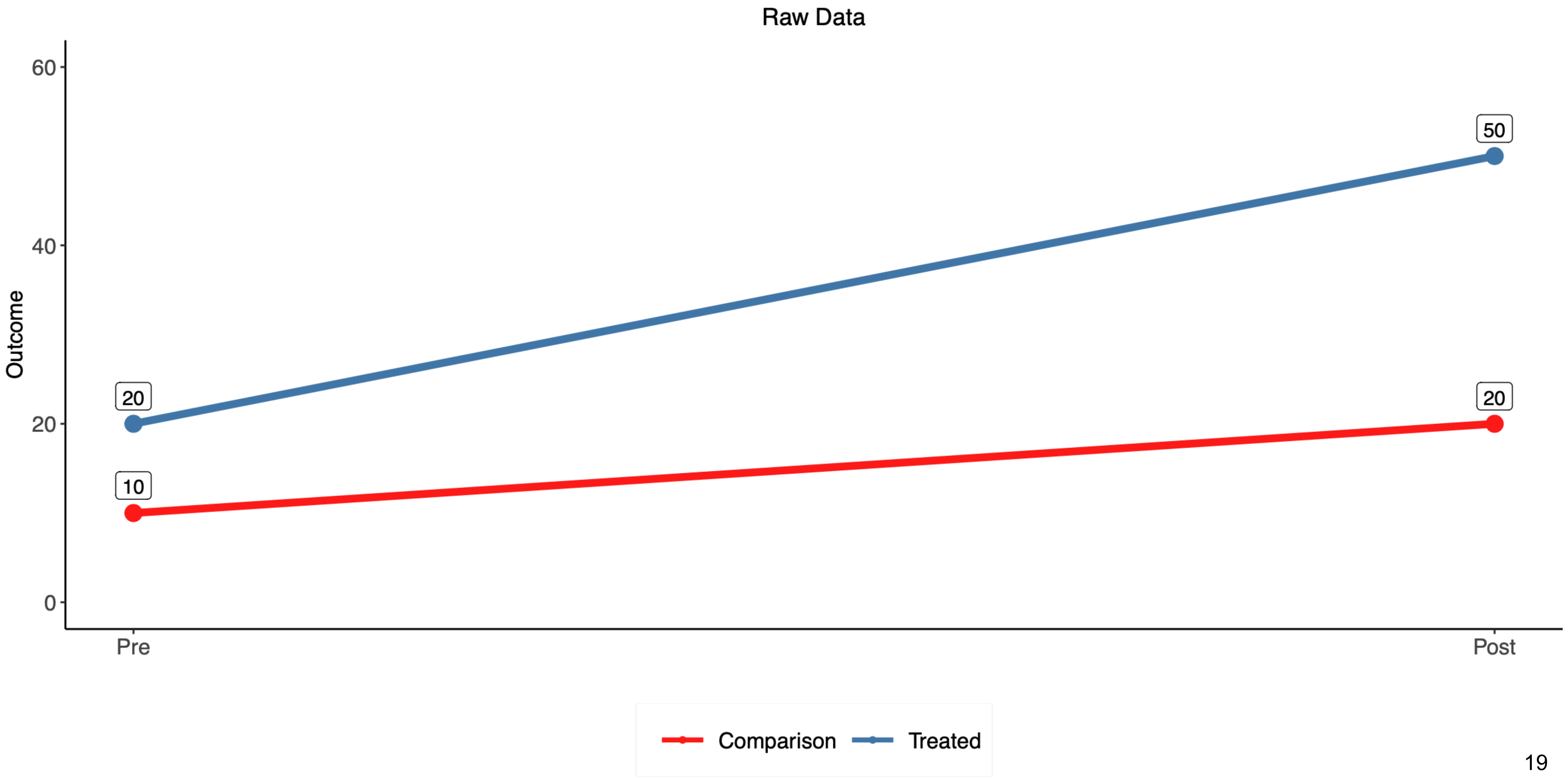
$$\hat{\theta}^{DiD} = \left(\bar{Y}_{Treated,Post} - \bar{Y}_{Treated,Pre}\right) - \left(\bar{Y}_{Untreated,Post} - \bar{Y}_{Unreated,Pre}\right)$$

where the subscripts refer to:

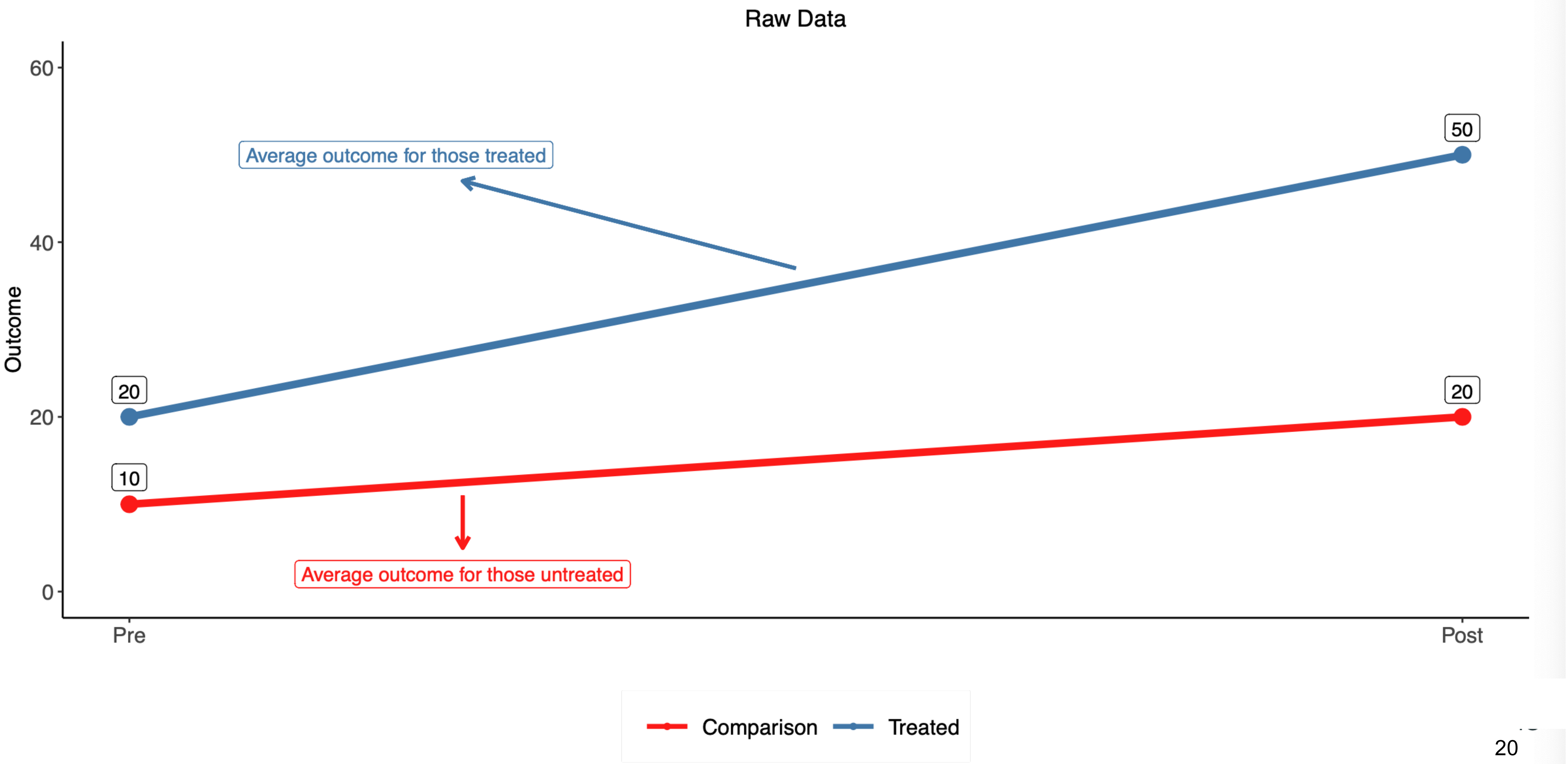
- Treated, Post = average outcome for the group that received the treatment after it happened
- Treated, Pre = average outcome for the group treated group before the treatment
- Untreated, Post = average outcome for the group that did not receive the treatment after the time that the treatment happened
- Untreated, Pre = average outcome for the group that did not receive the treatment before the treatment happened

Difference-in-Differences via graphs

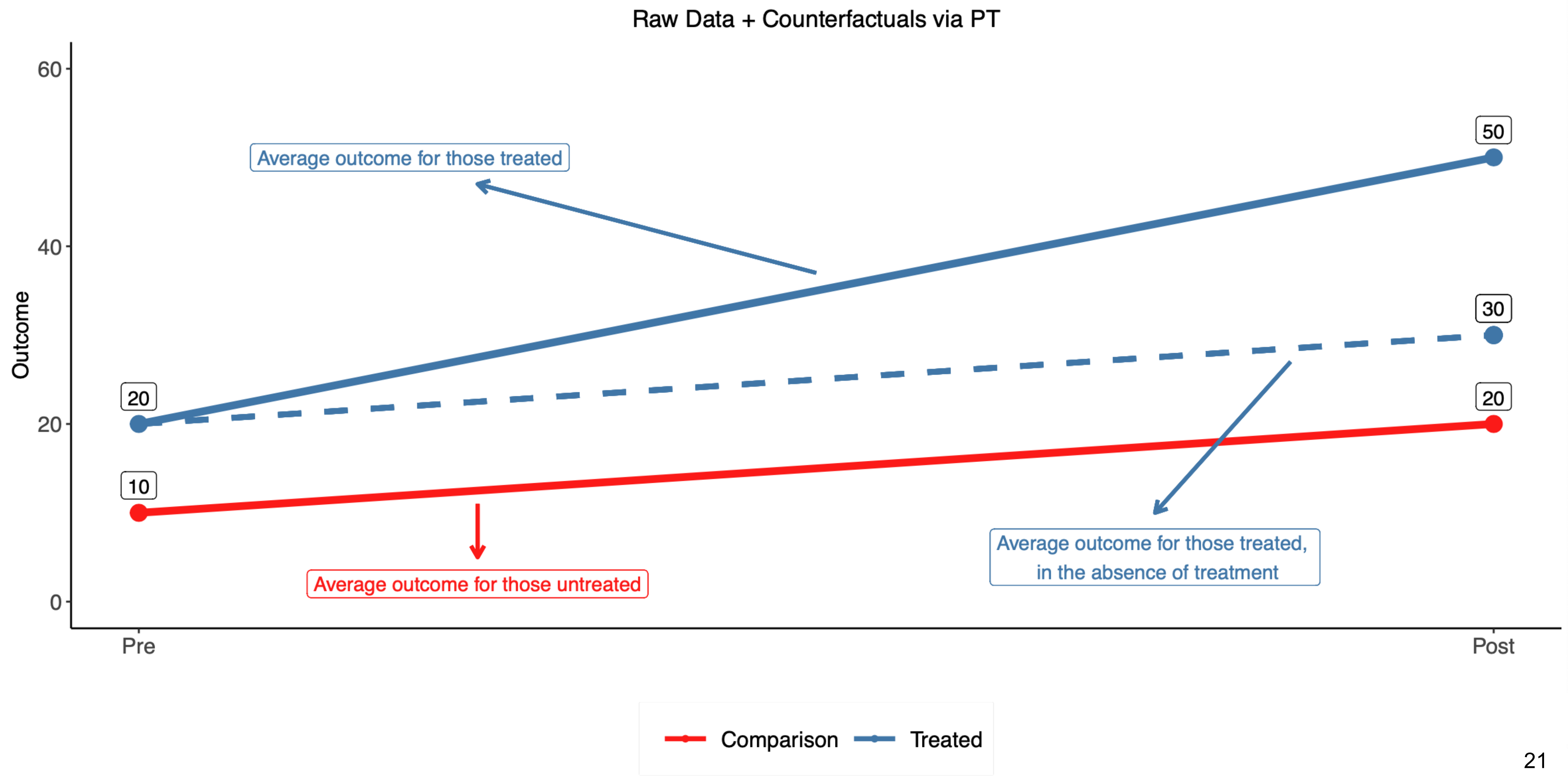
University of St Andrews



Difference-in-Differences via graphs



Difference-in-Differences via graphs



Difference-in-Differences via graphs

University of St Andrews

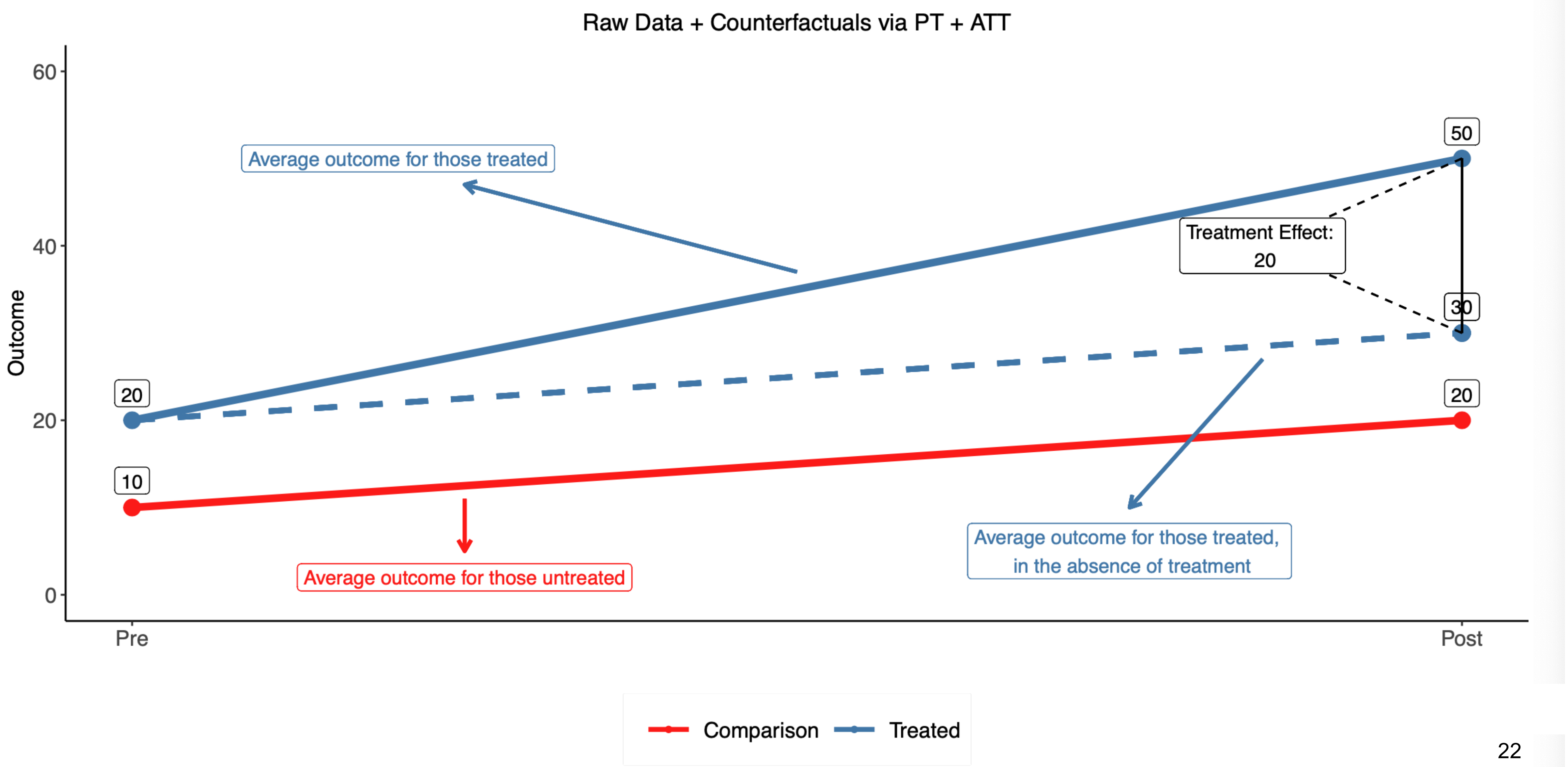


Illustration with R

Staggered difference-in-differences

Exploring two Github Repos

University of St Andrews

https://github.com/borusyak/did_imputation
https://github.com/andrewchbaker/JFE_DID

On Codespaces, create a new folder called “repos” and use “git clone” to download the above repos

References

Assunção, Juliano, Clarissa Gandour, Romero Rocha, and Rudi Rocha, “The Effect of Rural Credit on Deforestation: Evidence from the Brazilian Amazon,” Economic Journal, 2020, 130 (626), 290–330.

Baker, Andrew C., David F. Larcker, and Charles C.Y. Wang, “How much should we trust staggered difference-in-differences estimates?,” Journal of Financial Economics, 2022, 144 (2), 370–395.

Beck, Thorsten, Ross Levine, and Alexey Levkov, “Big bad banks? The winners and losers from bank deregulation in the United States,” Journal of Finance, 2010, 65 (5), 1637–1667.

Callaway, Brantly and Pedro H. C. Sant’Anna, “Difference-in-Differences with Multiple Time Periods,” Journal of Econometrics, 2021, 225 (2), 200–230.

Card, David and Alan Krueger, “Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania,” American Economic Review, 1994, 84 (4), 772–793.

Carey, Colleen M., Sarah Miller, and Laura R. Wherry, “The Impact of Insurance Expansions on the Already Insured: The Affordable Care Act and Medicare†,” American Economic Journal: Applied Economics, 2020, 12 (4), 288–318.

Currie, Janet, Henrik Kleven, and Esmée Zwiers, “Technology and Big Data Are Changing Economics: Mining Text to Track Methods,” AEA Papers and Proceedings, May 2020, 110, 42–48.

Dube, Arindrajit, T William Lester, and Michael Reich, “Minimum Wage Effects across State Borders: Estimates Using Contiguous Counties,” The Review of Economics and Statistics, 2010, 92 (4), 945–964.

——, T. William Lester, and Michael Reich, “Minimum wage shocks, employment flows, and labor market frictions,” Journal of Labor Economics, 2016, 34 (3), 663–704.

Goldsmith-Pinkham, Paul, “Tracking the Credibility Revolution across Fields,” arXiv:2405.20604, 2024.

Malesky, Edmund J., Cuong Viet Nguyen, and Anh Tran, “The impact of recentralization on public services: A difference-in-differences analysis of the abolition of elected councils in Vietnam,” American Political Science Review, 2014, 108 (1), 144–168.

Meyer, Bruce D., W. Kip Viscusi, and David L. Durbin, “Workers’ Compensation and Injury Duration: Evidence from a Natural Experiment,” The American Economic Review, 1995, 85 (3), 322–340.

Venkataramani, Atheendar S., Sachin J. Shah, Rourke O’Brien, Ichiro Kawachi, and Alexander C. Tsai, “Health consequences of the US Deferred Action for Childhood Arrivals (DACA) immigration programme: a quasi-experimental study,” The Lancet Public Health, 2017, 2 (4), e175–e181.