MBA 753 : Causal Inference Methods for Business Analytics

Dr. Nivedita Bhaktha

29.08.2024

Agenda

- Event Studies
- Difference-in-Differences method

- time series data
- before & after intervention
- intervention is the only reason for effect

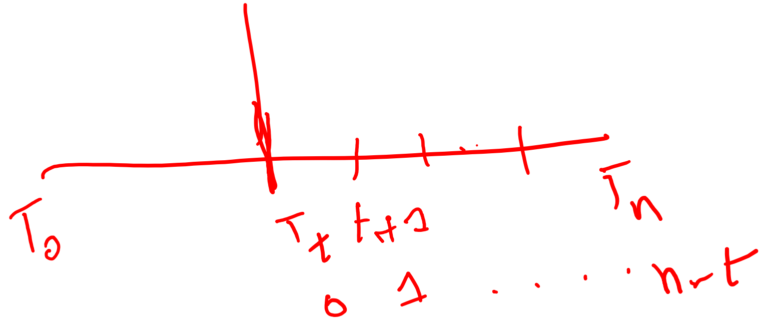
Event Studies

$$Y = \beta_0 + \beta_1 T + \beta_2 D + \beta_3 T*D + \varepsilon$$

Example – wellbeing classes

$$\rightarrow \quad Y = \beta_0 + \beta_1 T + \beta_2 D + \beta_3 P + \varepsilon$$

Column	Variable name	Description
Y	Wellbeing	Wellbeing index (from 0 to 300)
T	Time	Time (from 1 to 365)
D	Treatment	Observation post (=1) and pre (=0) intervention
P	Time Since Treatment	Time passed since the intervention



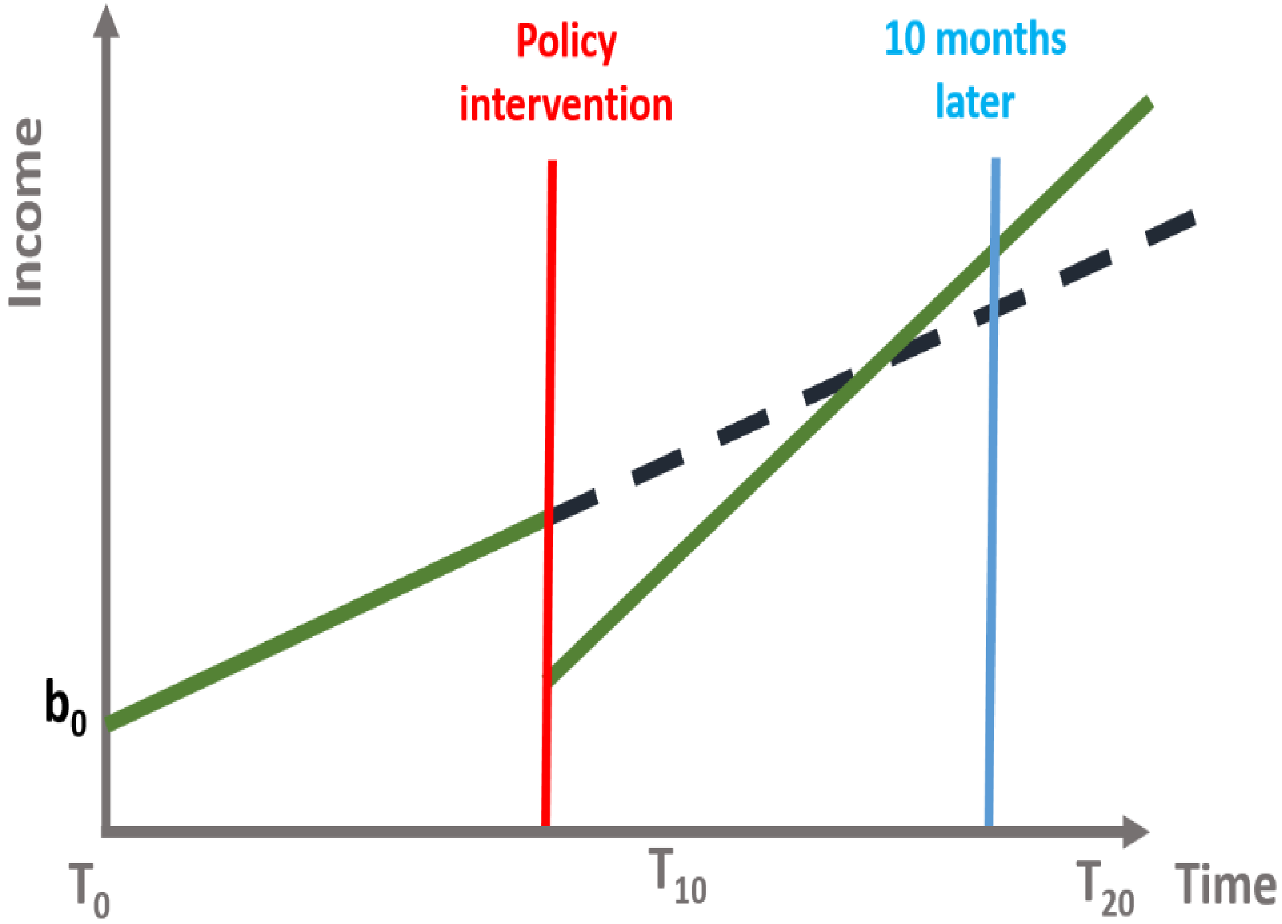
Autocorrelation

- Autocorrelation is a major issue when working with time series
  - Autocorrelation occurs when observation at one point in time depends from observations at another point in time
- OLS assumes that error terms associated with each observation are uncorrelated
  - Violated in presence of autocorrelation
- Impact - underestimated the standard errors
  - Overestimating the statistical significance

- Check residuals plot and Durbin-Watson test

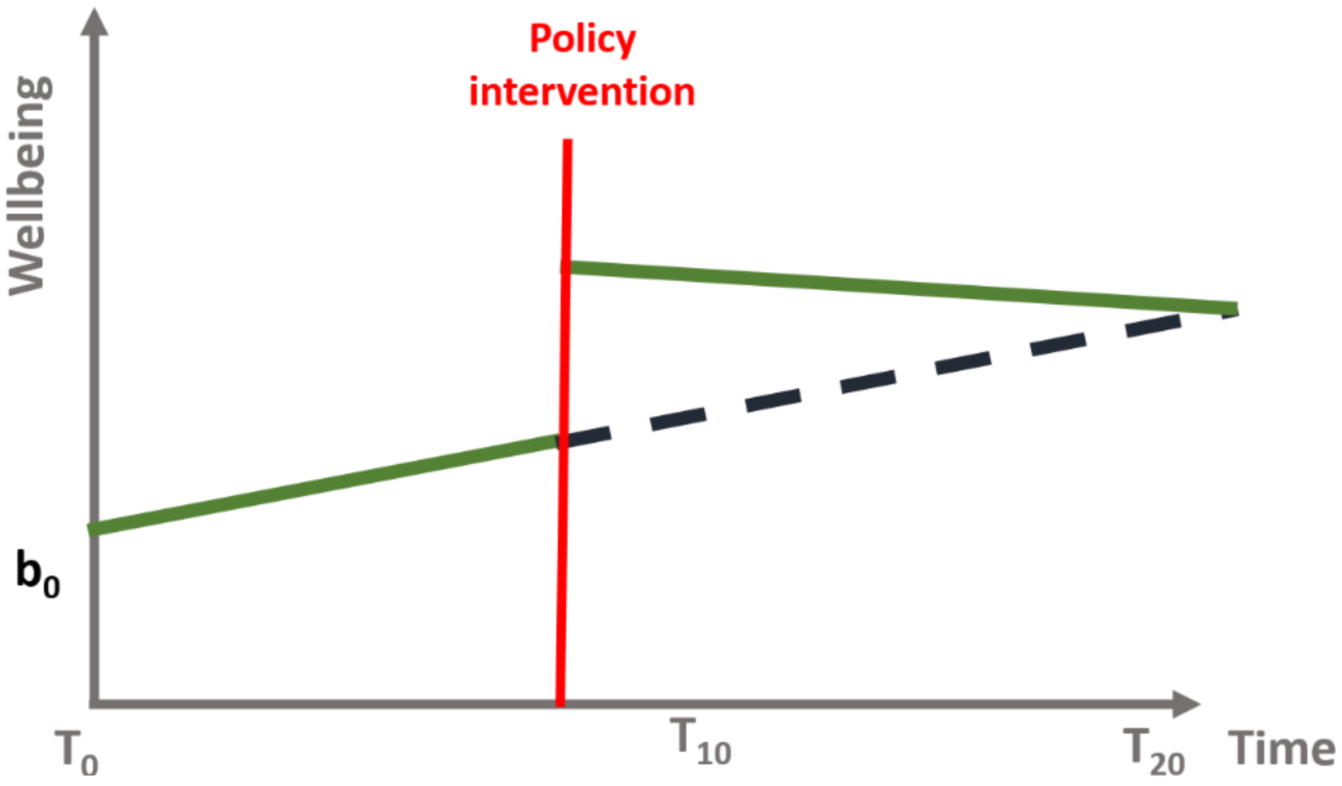
Issues with the design

- Delayed intervention effect
  - Misleading intervention effec
- look at immediate and sustained effects



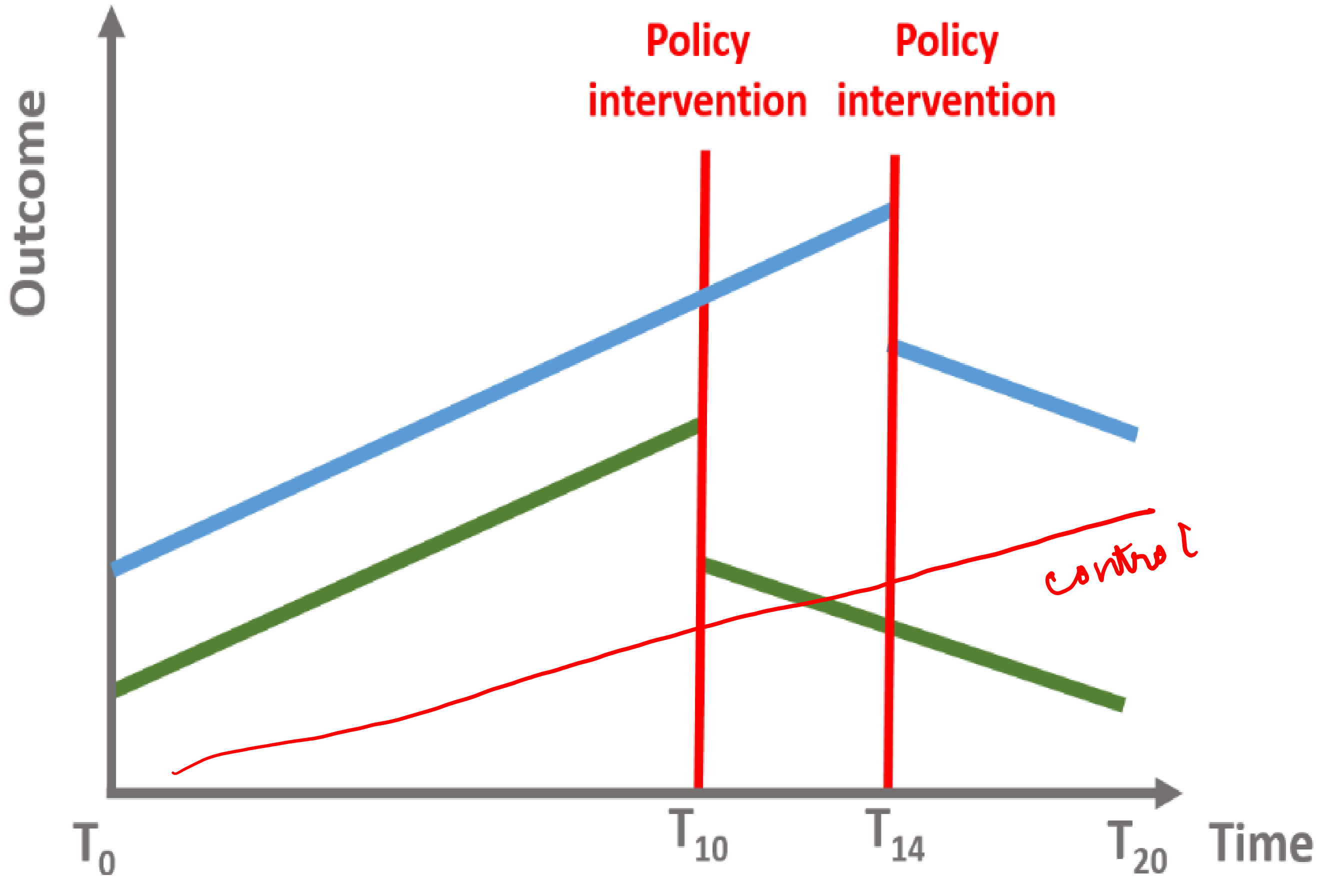
Issues with the design

- Regression to the mean
- How long to observe the effects of intervention?
  - how long the effect of an intervention will be sustained?



Issues with the design

- Validity Threats: selection bias and other related events
- Solution:
  - Use a control group
  - Study multiple effect time points

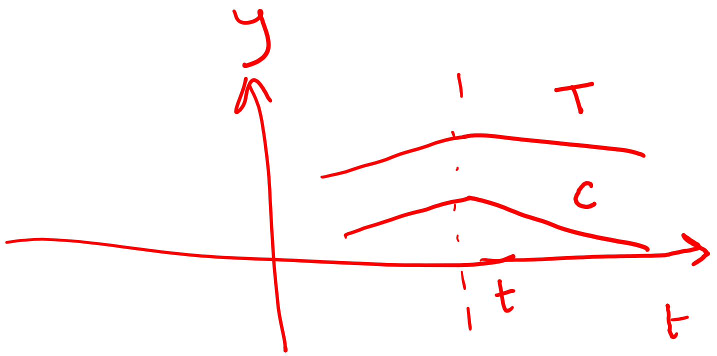


Difference-in-Differences Design

DID Concept

- Powerful quasi-experimental tool for assessing policy intervention
- It was introduced into economics by Orley Ashenfelter in the late1970s and then popularized through his student David Card (with Alan Krueger) in the 1990s
- Attempts to mimic random assignment with treatment and “comparison” sample
- Set-up: Policy intervention study – Event studies design
  - One group is ‘treated’ with intervention
  - Pre-post data for group receiving intervention
  - Can examine time-series changes but, unsure how much of the change is due to secular changes
  - There’s no counterfactual

Potential Solutions



- False counterfactuals
  - Before vs. After Comparisons:
    - Compares: same individuals/communities before and after program
    - Drawback: does not control for time trends
  - Participant vs. Non-Participant Comparisons:
    - Compares: participants to those not in the program
    - Drawback: selection bias

- DID combines pre vs. post and participant vs. non-participant approaches

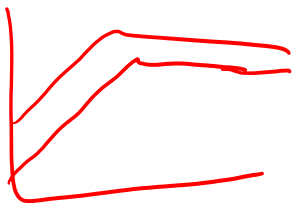
DID Framework

Before	After
T_1	T_2
C_1	C_2

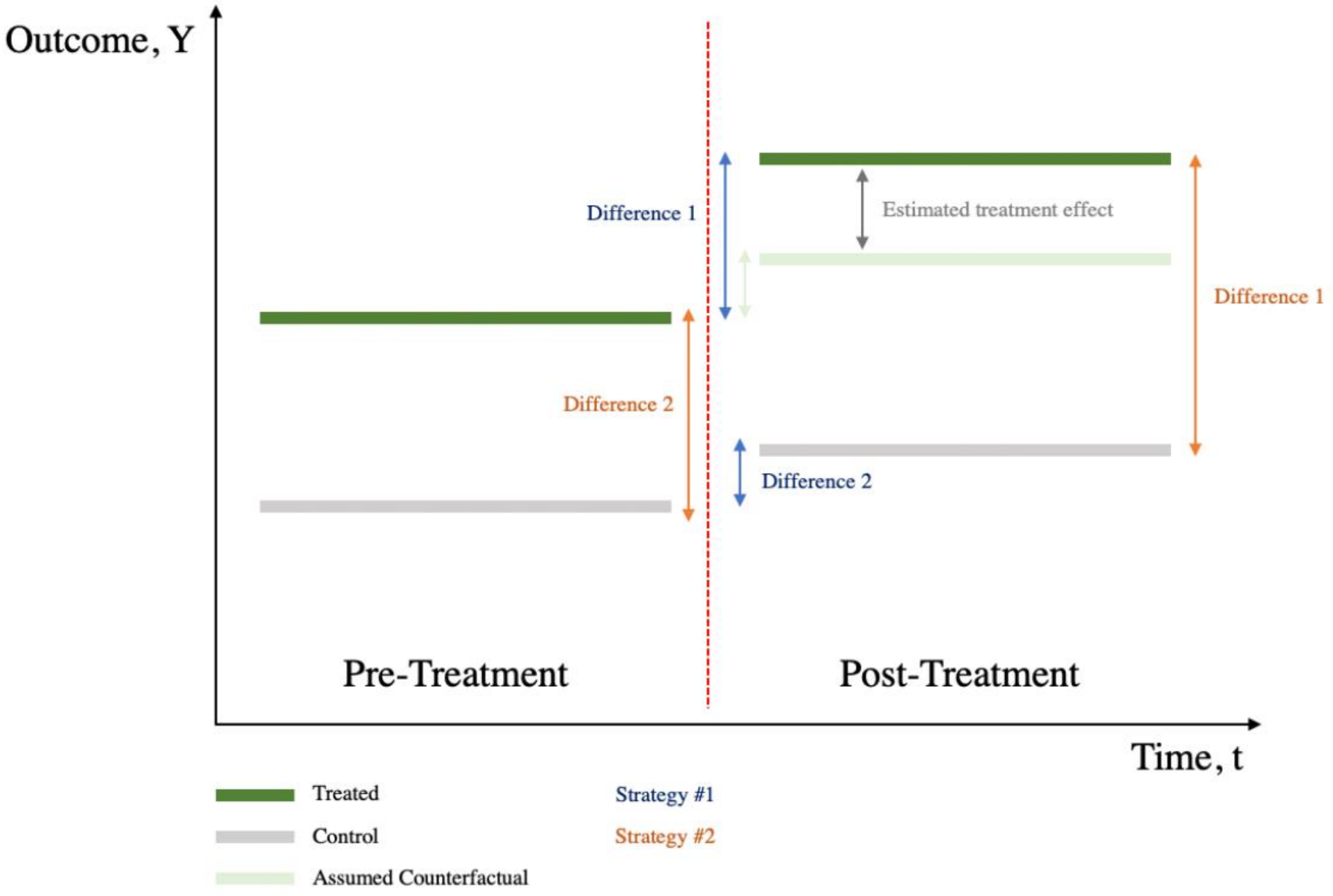
- Observe the (self-selected) treatment group and a (self-selected) comparison group before and after the program
- Key elements
  - **Time:** a before and after period. We could also have multiple periods before and after
  - **Comparison groups:** one group receives the intervention or is subjected to the policy change only in the post-period. These groups do not need to be comparable
  - **Fixed factors:** We assume that important factors that explain the outcome Y are fixed during the pre and post periods. If observed, we can control for those factors that could affect trends

DID Assumptions

- Three assumptions must hold: exchangeability, positivity, and Stable Unit Treatment Value Assumption (SUTVA) $E(Y_i | D_i) = E(Y_i)$
  - Intervention unrelated to outcome at baseline (allocation of intervention was not determined by outcome) $E(Y_i | D_i) \neq E(Y_i)$
  - Treatment/intervention and control groups have **Parallel Trends** in outcome: In absence of treatment, the unobserved differences between treatment and control groups are the same over time
  - Composition of intervention and comparison groups is stable for repeated cross-sectional design (part of SUTVA)
  - No spillover effects (part of SUTVA)



DID Model



Source: Jose Fernandez, 2024

Recap

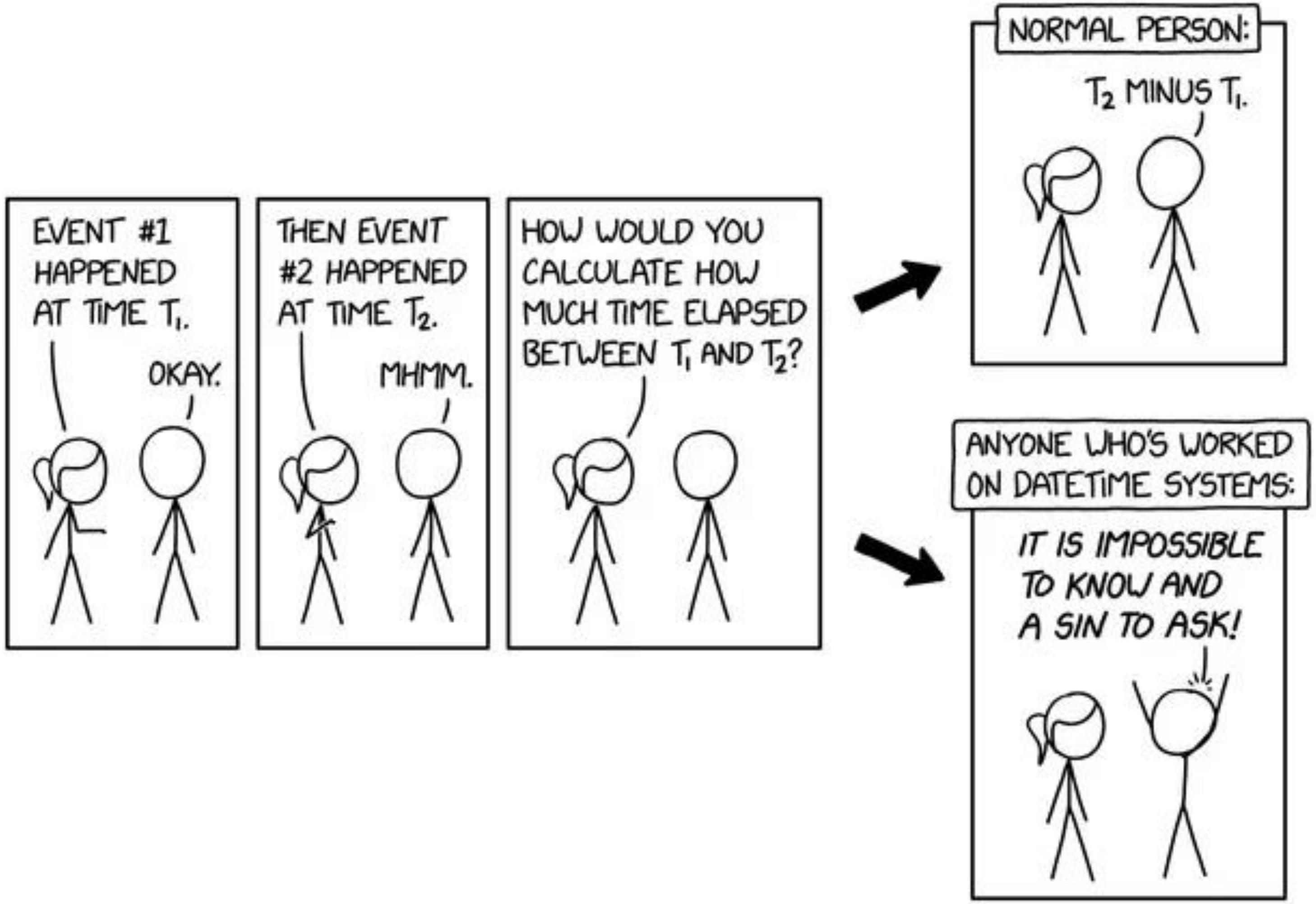
Summary

- When we want to study the effect on an intervention at a particular time point, we use time series data
  - Event studies - compare before event and after event
  - There are no alternative explanations for change
- DID is a widely used alternative to RCT
  - DID Setup: two or more groups, with units observed in two or more periods.
- Objectives achieved:
  - Can understand time series design, its advantages and disadvantages
  - Can interpret an event studies regression model
  - Can understand the advantages of DID design

References

- Scott Cunningham, Causal Inference: The Mix Tape, Yale University Press.
- Cook, T. D., & Campbell, D. T. (2007). *Experimental and quasi-experimental designs for generalized causal inference*.
- Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.
- Pearl, J., Glymour, M., & Jewell, N. P. (2016). *Causal inference in statistics: A primer*. John Wiley & Sons.

Thank You ☺



EVENT #1 HAPPENED AT TIME T_1.
OKAY.
THEN EVENT #2 HAPPENED AT TIME T_2.
MHMM.
HOW WOULD YOU CALCULATE HOW MUCH TIME ELAPSED BETWEEN T_1 AND T_2?
NORMAL PERSON:
T_2 MINUS T_1.
ANYONE WHO'S WORKED ON DATETIME SYSTEMS:
IT IS IMPOSSIBLE TO KNOW AND A SIN TO ASK!