

# Event Studies

Christoph Hanck

Summer 2023





# Event studies

## An event as old as time

The event study is probably the oldest and simplest causal inference research design.

### Idea

- At a certain time, an event occurred, leading a treatment put into place at that time
- Whatever changed from before the event to after is the treatment effect



# Event studies

## Examples: Event studies

- Someone has a late-night beer and immediately falls asleep, and concludes the next day that beer makes them sleepy: an event study
- When someone puts up a “no solicitors” sign on the door and then notices fewer solicitors afterwards, and concludes the sign worked is again an event study
- When a dog is hungry, then finds its master and whines, and becomes fed and full, then concludes that whining leads to getting fed is also an event study



# Event studies

## Scenario for event study

Event studies are designed to work in cases where the relevant causal diagram looks like the following:

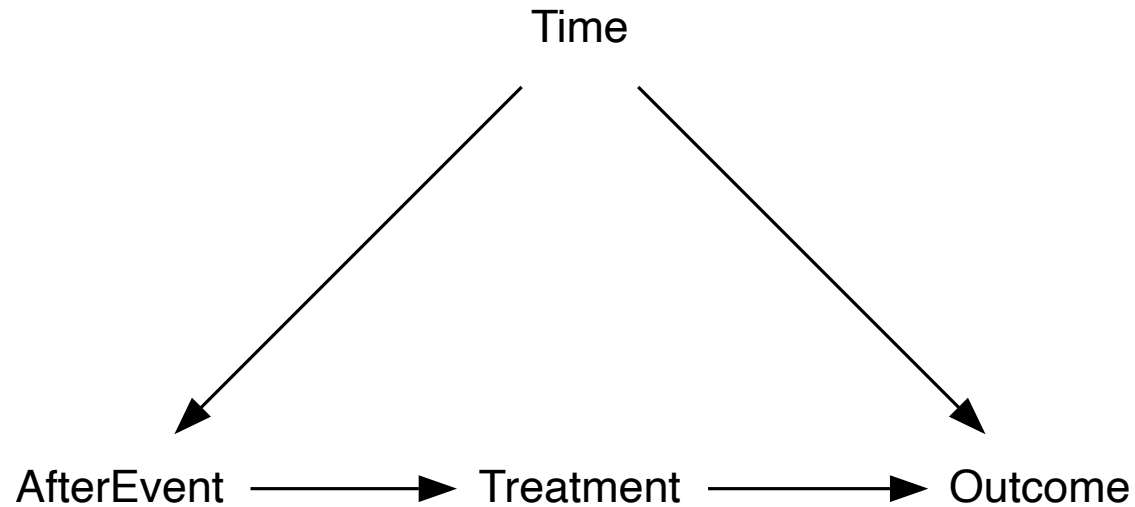


Figure 2: Causal diagram that event studies work for



# Event studies

## Scenario for event study

- At some given time period, when moved forward from “before-event” to “after-event”, some treatment goes into effect
- By comparing the outcome when the treatment is in place to the outcome when it is not, the effect of  $Treatment \rightarrow Outcome$  is estimated
- A back door is on *everything that changes with Time*:

$$Treatment \leftarrow AfterEvent \leftarrow Time \rightarrow Outcome$$



# Event studies

## Scenario for event study

- For an event study to work, the backdoor needs to be closed
- *AfterEvent* cannot be controlled for as that would in effect control for *Treatment*
- Because *Time* only affects *Treatment* through *AfterEvent* there is the possibility that we can close the back door by controlling for some elements of *Time* while still leaving variation in *AfterEvent*



# Event studies

## Scenario for event study

- The whole design of an event study is based around the idea that something changes over time — the event occurs, the treatment goes into effect.
- Then before-event and to after-event is compared and get the effect of treatment.
- *Treatment needs to be the only thing that is changing*
- If the outcome would have changed for some reason which cannot be accounted for, the event study will go poorly



# Prediction and deviation

## Treatment effect

We need to *predict the counterfactual*:

- Often we have plenty of information from the before-event times.
  - Predict what we would see without treatment, and then see how the actual outcome deviates from that prediction
  - (Under the assumption that whatever was going on before would have continued without treatment)
- The *extent of the deviation* is the treatment effect





# Prediction and deviation

Example: Extent of the deviation is the effect of treatment

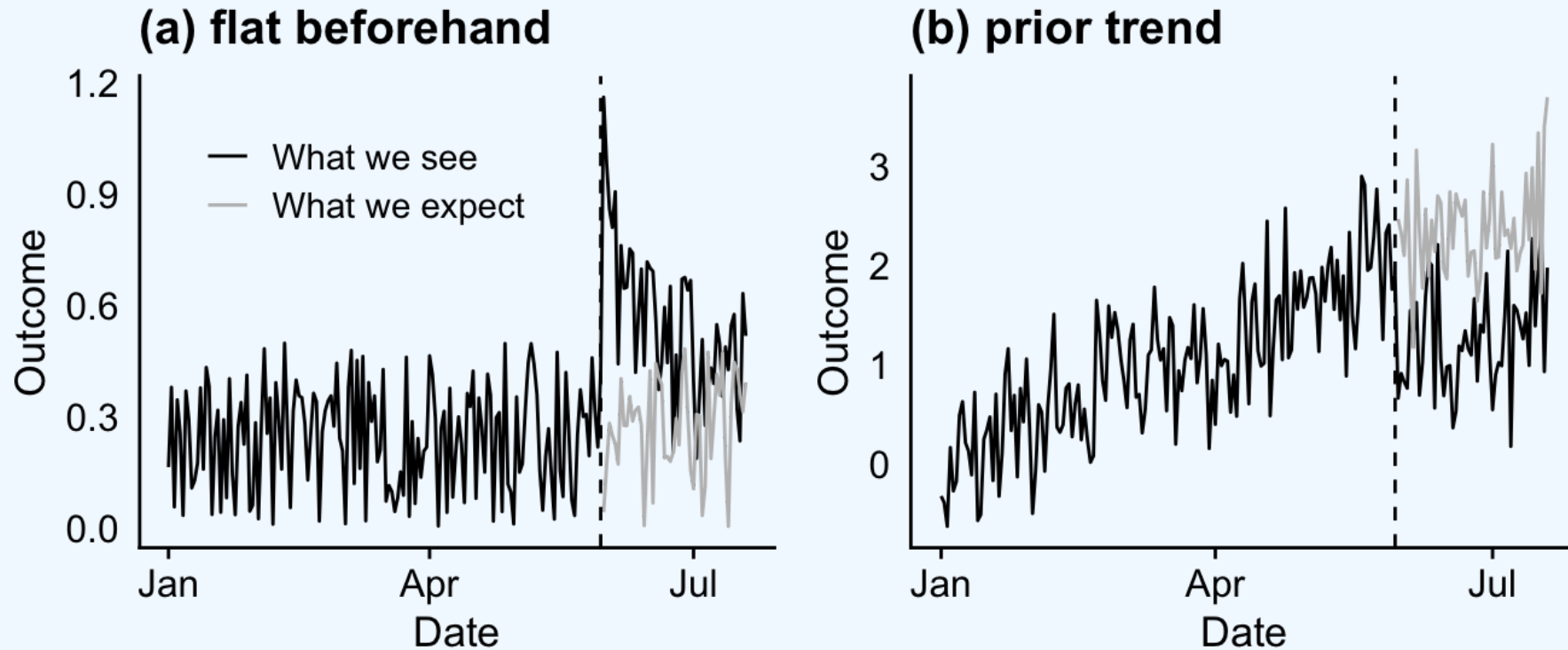


Figure 3: Two examples of graphs representing event studies



# Prediction and deviation

## Example: Extent of the deviation is the effect of treatment.

- To get the event study estimate, just compare the black line of real data to the gray line of predicted counterfactual data after the event goes into place.
- It looks like the treatment had an immediate positive effect on the outcome, with the black line far above the gray right after the treatment, but then that effect fades out over time and the lines come back together.
- The same idea applies to the second figure as well.



# Prediction and deviation

## How to make predictions?

There are three main approaches.

### Approach 1: Ignore possible changes

- A very common approach is to simply compare after-event to before-event without worrying about changes over time at all
- This works best if there is no time effect to worry about
- It also works well when looking at an extremely short time span



# Prediction and deviation

## How to make predictions?

### Approach 2: Predict after-event data using *before*-event data

Use patterns in the outcome data up to the event to predict what the outcome would be afterwards.

### Approach 3: Predict after-event data using *after*-event data

- Like in approach 2 use before-event data to establish some pattern in the data. Additionally look for relationships between the outcome and some other variables. E.g., stock price and market index.
- Extrapolate to the after-event period to get the predictions. Prediction can incorporate how those other variables change.



# Prediction and deviation

## Notes on these approaches

- All three methods make it clear that the event studies are generally intended for a fairly short post-event window (*short horizon*)
- The farther out one gets from the event, the worse a job one's pre-event-based predictions are going to do, and the more likely it is that the *Time* back door will creep in, inspite of the controls!



# Event studies

## Some notes on terminology

Event studies were developed across many fields independently and so terminology may have different meaning (and/or may be confusing)

- In health fields, the term **statistical process control** is used if there are no apparent pre-event trends. These, and the wide family of other similar methods, can be thought of as a form of event study.
- **Interrupted time series** usually either refers to an event study that uses time series econometrics, or to a regression-based approach that fits one regression before-event and another after-event



# Event studies in the stock market

Event studies are highly popular in studies of the stock market when the outcome of interest is a stock's return.

- Because in any sort of efficient, highly-traded stock market, a stock's price should already reflect all the public knowledge about a company
- Returns might vary a bit from day to day, or might rise if the market goes up overall
- The only reason the price should have any sort of sharp change is if there is surprising new information revealed about the company
  - This can be considered a *treatment* !



# Event studies in the stock market

## Process for doing a stock market event study

- Pick an **estimation period** a fair bit before the event, and an **observation period** from just before the event to the period of interest after the event
- Use the data from the estimation period to estimate a model that can make a prediction  $\hat{R}$  of the stock's return  $R$  in each period. Three popular ways of doing this are:
- Means-adjusted returns models, market-adjusted returns models and risk-adjusted returns models





# Event studies in the stock market

## Process for doing a stock market event study

1. **Means-adjusted returns model:** Just take the average of the stock's return in the estimation period,  $\hat{R} = \bar{R}$
2. **Market-adjusted returns model:** Use the market return in each period,  $\hat{R} = R_m$
3. **Risk-adjusted returns model:** Use the data from the estimation period to fit a regression describing how related the return  $R$  is to other market portfolios, like the market return:

$$R = \alpha + \beta R_m$$

Then, in the observation period, use that model and the actual  $R_m$  to predict  $\hat{R}$  in each period.



# Event studies in the stock market

## Process for doing a stock market event study

- In the observation period, subtract the prediction from the actual return to get the **abnormal return**,

$$AR = R - \hat{R}$$

- Look at  $AR$  in the observation period:

Nonzero  $AR$  values *before* the event imply the event was anticipated.

Nonzero  $AR$  values *after* the event give the effect of the event on the stock's return.

For the same efficient-market reason that stock returns are a good candidate for event studies in the first place, effects will generally spike and then fade out quickly.



# Event studies with regression

## Idea

- Estimate one regression of the outcome on the time period before the event. Then estimate another regression of the outcome on the time period after the event. See how the two are different.
- This approach can be implemented with the simple use of an interaction term (**segmented regression**),

$$Outcome = \beta_0 + \beta_1 t + \beta_2 After + \beta_3 t \cdot After + \epsilon,$$

where  $t$  is the time period and  $After$  is a binary variable equal to 1 in any time after the event.



# Event studies with regression

## Advantage

A more-precise estimate of the time trend than going day by day (or second by second, etc.) can be obtained.

## Disadvantage

- The exact shape that the effect takes may not be captured. We need a decent (polynomial) representation of the time series on either side of the event. Results may be bad if the time trend is misspecified!
- For autocorrelated data there is a high chance of obtaining a statistically significant effect of the event, even if it truly had no effect!
  - Use heteroskedasticity- and autocorrelation-consistent (HAC) standard errors
  - Use a model that directly accounts for autocorrelation



# Event studies with regression

## Example: Effect of ambulance quality control policy on heart attack care

- Taljaard et al. (2014) looked at an English policy put in place in mid-2010 to improve quality of health care received in the ambulance on the way to the hospital on the chances of heart attack and stroke afterwards
- Quality-improvement teams were formed, they all collaborated, they shared ideas and they informed staff
- The aim was to determine if the quality improvement teams significantly improved heart attack care



# Event studies with regression

## Example: Effect of ambulance quality control policy on heart attack care

Their regression model is:

$$AMI = \beta_0 + \beta_1(Week - 27) + \beta_2 After + \beta_3(Week - 27) \cdot After + \epsilon$$

- *AMI* measures heart attack performance (Acute Myocardial Infarction)
- *Week - 27* (centered at the event period, which allows the coefficient to represent a jump in performance)
- *After* (indicator for being after the 27-week mark of the data where the policy was introduced)



# Event studies with regression

Example: Effect of ambulance quality control policy on heart attack care

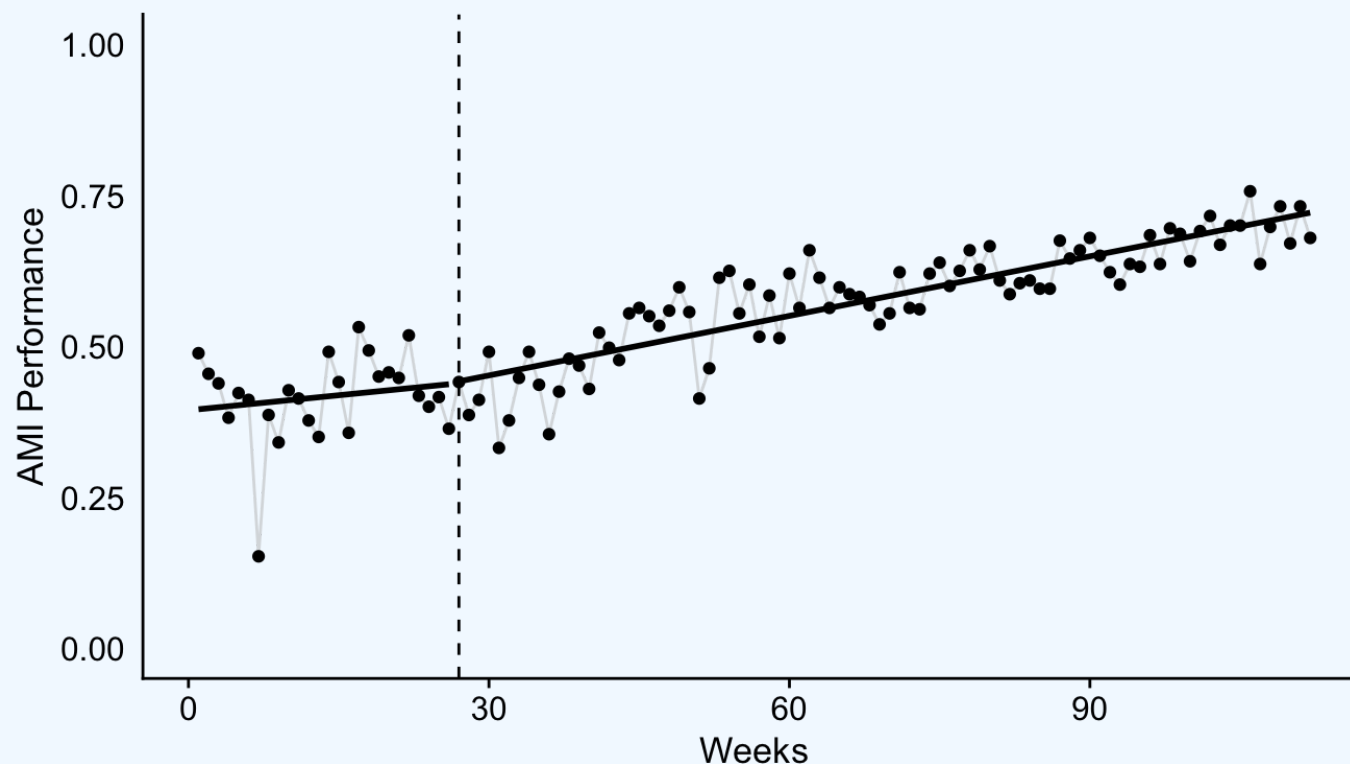


Figure 4: Effect of an Ambulance Quality Control Cooperative Policy on Heart Attack Care in Taljaard et al. (2014)



# Event studies with regression

## Example: Effect of ambulance quality control policy on heart attack care

- There is no jump at the cutoff. So no immediate effect of the policy on outcomes.
- The jump is represented by the change in prediction at the event time,  $Week = 27$
- Comparing the lines on the left and the right at that time, a prediction of  $\beta_0$  on the left and  $\beta_0 + \beta_2$  on the right is obtained.





# Event studies with regression

## Example: Effect of ambulance quality control policy on heart attack care

- The difference is

$$(\beta_2 + \beta_0) - \beta_0 = \beta_2$$

- $\beta_2$  is the jump in the prediction at the moment the treatment goes into effect
- This is the **event study estimate** of the treatment effect



# Event studies with multiple affected groups

- All the methods that are discussed so far have assumed an event that affects only one group
- However, there are many events that can be applied to lots of groups.
- To run an event study in a case like that, the following can be done:
  - Pretending not having many groups
  - Each group can be treated separately
  - Groups can be aggregated with regression



# Event studies with multiple affected groups

## Pretending not having many groups

- If a method is designed for a single time series, but lots of different time series are present that matter. Then they all can be squeezed together.
  - Average a bunch of time series.
- One disadvantage is that a lot of information can be lost by doing this and the estimate will be less precise



# Event studies with multiple affected groups

## Each group can be treated separately

- Running one event study per group if study methods for single groups are present and there are lots of groups will give a different effect for each group.
- At this point we have effects for each group — this could be an immediate effect of
  - a single post-event time period or
  - the coefficient of *After* on slide 19, or
  - an average effect over the entire post-event period, or a cumulative effect.
- From there the distribution of effects across all the groups can be inspected (and summarized, think averages)



# Event studies with multiple affected groups

## Groups can be aggregated with regression

- The regression approach to event study estimation allows to have more than one observation per time period
- The regression model is nearly the same as when there is just one time series:

$$Outcome = \beta_i + \beta_1 t + \beta_2 After + \beta_3 t \cdot After + \epsilon$$

- The only difference is the switch from  $\beta_0$  to a group to a group fixed effect of  $\beta_i$
- This is not always necessary, but if the different groups have different baseline levels of *Outcome* the fixed effect will help deal with that



# Event studies with multiple affected groups

## Groups can be aggregated with regression

- This regression will produce results very similar to what can be expected if we first average all the data together and then estimate a regression on that single time series.
- If an event matters only for a day or the interest is to know the full shape of the event, a different type of regression is required:

$$Outcome = \beta_0 + \beta_t + \epsilon$$

(regression on a set of *time-period fixed effects*)



# Event studies with multiple affected groups

## Groups can be aggregated with regression

- The fixed effects are done in a particular way. The last time period before the event should have an effect is selected as the reference category.
- The fixed effect for a given period is then just an estimate of the mean outcome in that period relative to the period just before the event
- If time-period fixed effects themselves are plotted, they give a single time series
- The only difference is that now standard errors for each of the periods are present, and everything is made relative to the period just before the event



# Forecasting properly with time series methods

- Remember that an event study requires that data in the after-event period is present and that we can conduct a comparison and a prediction of what would have happened in the after-event period if the event had not occurred
- The prediction is invariably based on the idea that, *without the event*, whatever was going on before the event would have continued
- In other words, time series data from before the event is used to predict beyond the event. This is **time series forecasting**.





# Forecasting properly with time series methods

## The ARMA Model

- The ARMA model is an acronym for two common features that can be seen in time series data.
  - **AR stands for autoregression**, meaning that the value of something today influences its value tomorrow.
  - **MA stands for moving-average**, meaning that transitory effects on something last a little while and then fade out.



# Forecasting properly with time series methods

## The ARMA Model

The ARMA model consists of an autoregressive (AR) and a moving-average (MA) component

### AR(1) model

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \epsilon_t$$

$t$  and  $t - 1$  indicate that the predictor for  $Y$  in a given period is  $Y$  from the previous period: the **lag** of  $Y$ .

### MA(1) model

$$Y_t = \beta_0 + (\epsilon_t + \theta \epsilon_{t-1})$$

$\epsilon_t$  and  $\epsilon_{t-1}$  are error terms at times  $t$  and  $t - 1$ . Since  $\epsilon_t$  and  $\epsilon_{t-1}$  cannot be observed,  $\theta$  cannot be estimated by regular regression (there are alternative methods).



# Forecasting properly with time series methods

## Example: ARMA(2,1) model

An ARMA(2,1) model includes two autoregressive terms and one moving-average term. The current value of the time series depends on the previous two values, and the transitory effect in one previous period,

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + (\epsilon_t + \theta \epsilon_{t-1}).$$



# Forecasting properly with time series methods

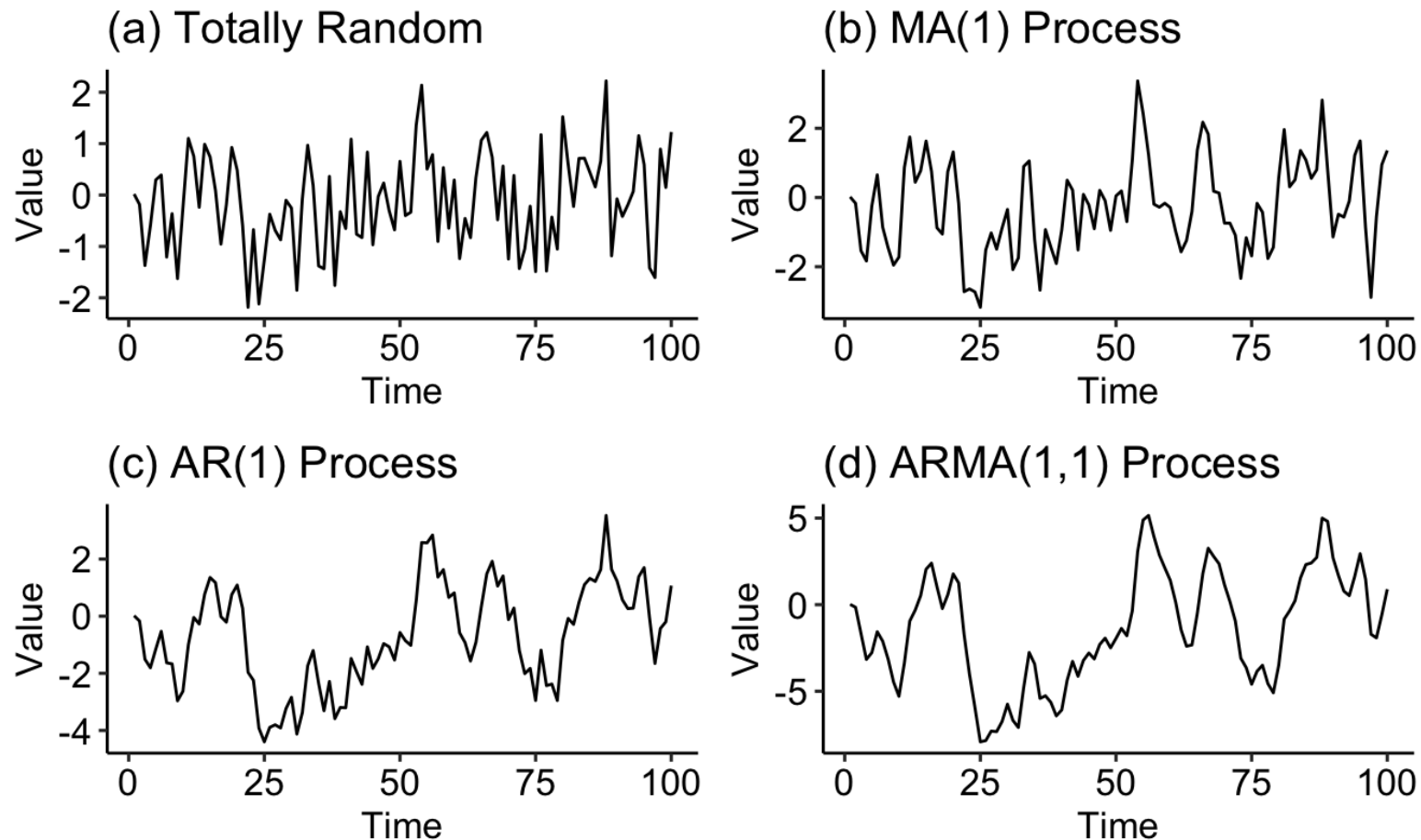


Figure 5: Different time series processes



# The joint-test problem

- Results obtained from an event study are a combination of two things:
  - The actual event study effect
  - The model which is used to generate the counterfactual prediction
- Event study designs are highly reliant on making *accurate predictions of the counterfactual*.

That **prediction is the only thing making the event study work!**

- Doing a significance test of that effect would imply
  - ...testing the true event-study effect.
  - ...doing a joint test of both the true effect and whether the predictive model is right.
- These two parts *cannot* be separated!
  - **The counterfactual-prediction model is made as well as possible**



# The joint-test problem

## Definition: Placebo test

A placebo test is a test performed under conditions where there should be zero effect. So if you find an effect, something is wrong.



# The joint-test problem

## Placebo testing

The application of placebo testing in event studies has been around a long time.

## Idea

- When there is supposed to be an effect, and tested for it, we cannot tease apart what part of the findings is the effect and what part is due to the counterfactual model being wrong.
- If it is known from before that there was no effect, then anything estimated could only be the counterfactual model being wrong
  - If an effect were obtained under those conditions, then the model was fishy and we would have to try something else