5. Event studies

LPO 8852: Regression II

Sean P. Corcoran

Introduction

Event studies

Event studies are a class of models that attempt to estimate the effect of an "event" (or treatment) over time on some outcome. This usually involves imputing or modeling a counterfactual time trend. Event studies are often interested in the *dynamic* aspects of treatment effects.

Event studies have a long history, especially in finance, where "event studies" estimate the effect of a notable event (e.g., new CEO, interest rate announcement) on the market or on an individual stock.

Lecture 4 considered the **interrupted time series** (single unit, before and after) design, which can be considered a type of event study.

Single unit event studies

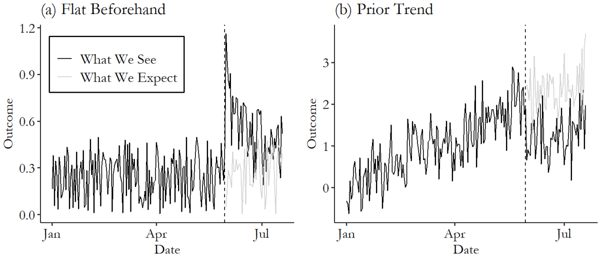
The single unit event study includes one unit observed "before" and "after" some event. To estimate the effect of the event, we need some way to predict the counterfactual.

Our best prediction may be to assume whatever we observed "before" would continue "after" in the absence of intervention. For example, we could fit a linear or nonlinear time trend before the event, or use other informative variables for prediction.

Note: you are unlikely to be estimating many single unit event studies! However, they are useful for building intuition.

Single unit event studies

Extrapolating pre-event trends into the post period:

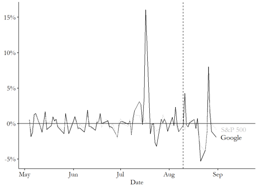


Source: Huntington-Klein chapter 17

Single unit event studies

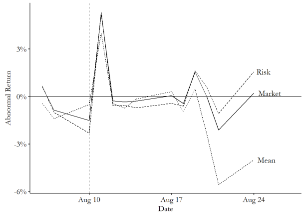
This approach could work well in contexts in which a trend is well established and expected to continue. Example: high-frequency stock price data observed over a relatively short period of time.

H-K: Google announces the creation of Alphabet on August 10, 2015. How did investors value this change? If markets are efficient, prices adjust quickly to new information. Graph: daily price change Google and S&P500



Single unit event studies

In the graph below, H-K plots "abnormal returns" for Google: the difference between the Google daily price change and (1) the *mean* price change in the pre period; (2) the mean price change for the *market* in the pre-period; and (3) the predicted Google price change based on the market price change (estimated in the pre-period; "*risk*").



Source: Huntington-Klein chapter 17

Single unit event studies

In the Google example, the effect was **transitory**. For some outcomes you might expect to see an effect that persists and perhaps changes the time trend thereafter. A simple interrupted time series model for this scenario:

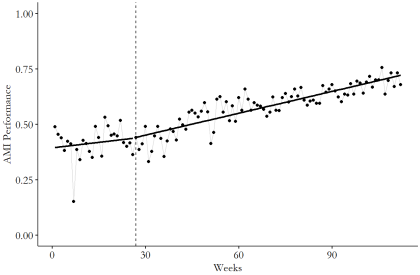
$$Y_t = \beta_0 + \beta_1 t + \beta_2 POST_t + \beta_3 (t \times POST_t) + e_t$$

where t is centered at $t = 0$ when the event occurs. This model assumes a linear time trend. One could use higher-order terms if appropriate.

NOTE: autocorrelation is common with time series data. Use robust standard errors when estimating a model like this.

Single unit event studies

H-K: example estimating the effects of an ambulance health intervention in England in 2010 on heart attack (Taljaard et al 2014).



Source: Huntington-Klein chapter 17

Multiple treated units: common treatment timing

Suppose you have multiple units (i) treated in the same period (and all units are treated). Now with multiple observations per time period, one can estimate separate time fixed effects:

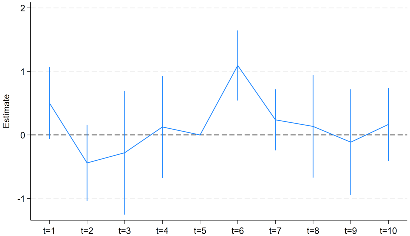
$$Y_{it} = \beta_0 + \gamma_t + e_t$$

where the last time period before the event is usually the omitted reference category. The estimated $\hat{\gamma}_t$ trace out the time path of Y relative to the reference period.

H-K: example using simulated data, 10 units and 10 time periods with treatment in $t = 6$.

Multiple treated units: common treatment timing

Each estimated year effect is relative to $t = 5$. (Treatment is in $t = 6$).



In this example with simulated data, all pre-treatment year effects should be zero. Some may differ from zero due to sampling variability, however.

Multiple treated units: common treatment timing

While simple to estimate, there are a few important limitations to this design:

- There is no untreated comparison group. It is impossible to separate treatment effects from other things changing over time.
- In the real word, events happen at different times.

Event study models

Event time

Define **event time** or **time since event** j as the number of periods since an event occurred for unit i.

$j = 0$ in the period in which treatment occurs

$j = -1$ in the period just before treatment

$j = 1$ in the period just after treatment, etc.

Let m represent the (max) number of periods observed before event

Let n represent the (max) number of periods observed after event

Can refer to E_i as the period in which the event occurs (if any) for unit i. So in time period t, $E_i = t - j$ (or, $j = t - E_i$). Note we are assuming at most <u>one</u> event per unit for now, and the event is "absorbing."

Core features of event study models

$$Y_{it} = \underbrace{\left(\sum_{j \in \{-m,\ldots,0,\ldots n\}} \gamma_j D_{i,t-j} \right)}_{\text{Event study terms}} + \underbrace{\alpha_i + \delta_t}_{\text{Panel fixed effects}} + \beta X_{it} + u_{it}$$

- $D_{i,t-j} = 1$ if the event occurred j periods before the time period t
- α_i is a unit fixed effect, δ_t is a fixed effect for **calendar time**
- Time-varying covariates X_{it} are optional
- Events can occur at higher levels of aggregation (e.g., states)
- The γ_j are the **event study coefficients**
- When estimating, $j = -1$ is usually the omitted reference period

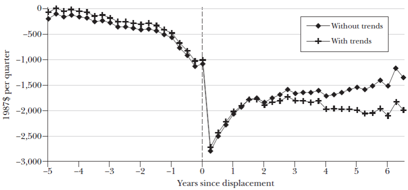
Core features of event study models

- You will notice this is our TWFE difference-in-differences model but with event time dummies for every pre and post period.
- The model can also accommodate "never treated" cases: they have $D_{i,t-j} = 0 \forall t$
- The main output—an **event study plot**—is just a plot of the estimated event study coefficients.

Example

An early (1993) example of an event study plot:

Figure 1
An Event Study Example: Loss of Income after Being Displaced from a Job



Source: Jacobson, LaLonde, and Sullivan (1993).
Note: Figure reproduced from Jacobson, LaLonde, and Sullivan (1993). The x-axis is measured in "event time." The y-axis show income for each period relative to a baseline comparison period more than five years prior to the job discplacement.

Source: Jacobson, LaLonde, and Sullivan (1993) reproduced in Miller (2023)

Uses of event study models

Event study models are used for several purposes:

- To estimate dynamic treatment effects—when the appropriate assumptions hold.
- To check for differences in pre-trends (e.g., due to anticipation, model mis-specification, omitted time-varying covariates) in a DD design.

Data structure

Miller (2023) characterizes data structures typically used in event studies using two criteria:

1. Are there "never-treated" units?
2. Is there considerable variation in the treatment date across units?

Table 1
Data Structures for Event Study Estimation

	Only Ever-Treated Units	*There are Never-Treated Units*
Common Event Date	N/A	DiD-type
Varying Event Date	Timing-based	Hybrid

Note: Author's proposed labels for event study data structures, based on whether the analysis data sample uses never treated units or not, and on whether treated units have a common event date or varying event dates. "DiD-type" = "Difference in Difference type."

You should be able to answer "yes" to one or both of these questions.

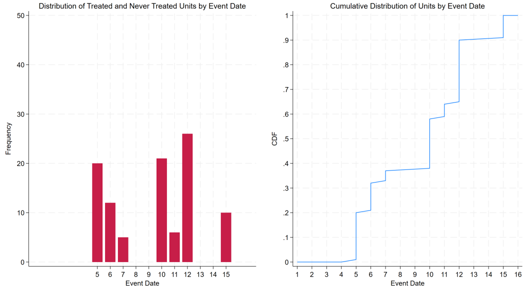
Data structure

- Top left: single group before and after; hard to identify treatment effects.
- "DiD type": treated and untreated units before and after.
- "Timing based": only treated units but treatment times vary. Assume timing is as good as random.
- "Hybrid": comparisons to untreated and treated earlier/later.

It is important to be clear with the reader about which data structure you have, and the distribution of observations across event times. See following examples of graphical ways to communicate this.

Data structure: event timing

All units are treated but at varying times:

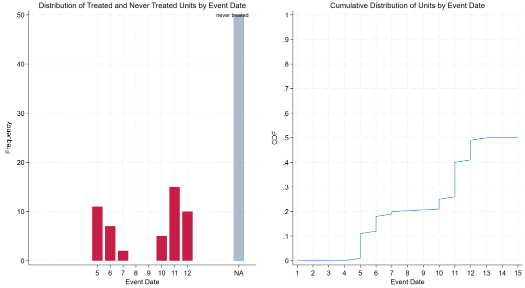


Probability and cumulative distribution functions

Source: Miller (2023) appendix

Data structure: event timing

Data has “never treated” units:



Probability and cumulative distribution functions

Source: Miller (2023) appendix

Estimation and interpretation: example

The event study model above can be estimated using OLS and `regress` in Stata.

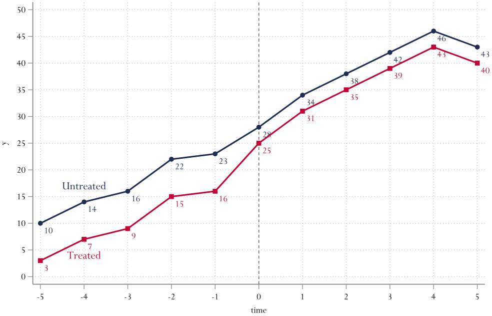
Example: Stylized data with multiple groups, a common treatment period, and a "never treated" group.

```
reg y2 i.treated##ib4.etime2
```

This is a full factorial of treatment group status and time period—includes main effects for *treated* and each year, and their interaction. You will need to specify the omitted time period. Here the omitted reference group is time period 4 (the year before treatment in this case).

Estimation and interpretation: example

Stylized example: suppose these are the mean observed Y in each time period for two groups:



Estimation and interpretation: example

```
. reg y2 i.treated##ib4.etime2

      Source |       SS           df       MS      Number of obs   =     1,100
-------------+----------------------------------   F(21, 1078)     =    193.37
       Model |  182849.286        21  8707.10886   Prob > F        =    0.0000
    Residual |  48540.4525     1,078  45.0282491   R-squared       =    0.7902
-------------+----------------------------------   Adj R-squared   =    0.7861
       Total |  231389.739     1,099  210.545713   Root MSE        =    6.7103

------------------------------------------------------------------------------
          y2 | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
   1.treated |  -8.379426   1.342062    -6.24   0.000    -11.01278   -5.746076
             |
      etime2 |
         -5  |  -13.99077   1.342062   -10.42   0.000    -16.62412   -11.35742
         -4  |  -9.457968   1.342062    -7.05   0.000    -12.09132   -6.824618
         -3  |  -5.494279   1.342062    -4.09   0.000    -8.127629    -2.86093
         -2  |   1.748657   1.342062     1.30   0.193    -.8846928    4.382006
          0  |   3.477879   1.342062     2.59   0.010     .8445296    6.111229
          1  |   7.835866   1.342062     5.84   0.000     5.202516    10.46922
          2  |   16.10816   1.342062    12.00   0.000     13.47481    18.74151
          3  |   16.83639   1.342062    12.55   0.000     14.20304    19.46974
          4  |   22.24522   1.342062    16.58   0.000     19.61187    24.87857
          5  |   19.24763   1.342062    14.34   0.000     16.61428    21.88097
             |
treated#etime2 |
       1#-5  |   2.958434   1.897962     1.56   0.119    -.7656849    6.682552
       1#-4  |   1.876838   1.897962     0.99   0.323     -1.84728    5.600957
       1#-3  |   .0605446   1.897962     0.03   0.975    -3.663574    3.784663
       1#-2  |  -1.621474   1.897962    -0.85   0.393    -5.345592    2.102645
        1#0  |   7.718263   1.897962     4.07   0.000     3.994145    11.44238
        1#1  |   7.082635   1.897962     3.73   0.000     3.358517    10.80675
        1#2  |   4.351253   1.897962     2.29   0.022     .6271342    8.075371
        1#3  |   6.906317   1.897962     3.64   0.000     3.182199    10.63044
        1#4  |   6.213328   1.897962     3.27   0.001     2.489209    9.937447
        1#5  |   5.132189   1.897962     2.70   0.007     1.408071    8.856308
             |
       _cons |   23.43028    .948981    24.69   0.000     21.56822    25.29234
------------------------------------------------------------------------------
```


Estimation and interpretation: example

Interpreting the event study coefficients:

- Constant term in the reference year: mean outcome for untreated in the base period.
- Coefficient on "treated": *difference* in mean outcome between the treated and untreated groups in the base period.
- Coefficients on event time (-5 to +5): different between time period k and the base period, for the *untreated* group.
- Coefficients on event time $\times$ treated interaction: the difference in mean outcomes between the treated and untreated groups in period k relative to their prevailing difference in the base year.

With parallel trends, would expect the coefficients on the pre-treatment interactions to be zero.

Estimation using eventdd

The user-written Stata package `eventdd` is a flexible solution that automatically generates the needed variables, estimates the regression, and produces a graph. Example syntax:

`eventdd` *y x1 x2 i.group*, `timevar(`*eventtime*`)`

This syntax estimates an OLS model with the *group* main effect included in the covariates (see also next slide). The key variable here is *eventtime* (a name you provide), defined as the **relative time** to treatment. 0 corresponds to the first year of treatment, -1 refers to the first lead, and so on. This variable should be **missing for groups that are never treated**. See Clarke and Schythe (2020).

Estimation using eventdd

With lots of groups (or panel data) you can have `eventdd` estimate a fixed effects model specification:

`eventdd` *y x1 x2*, `timevar(`*timetoevent*`) method(fe, absorb(state))`

Here the command uses `xtreg` where the variable *state* is used as the fixed effect.

Event study example: Miller et al (2021)

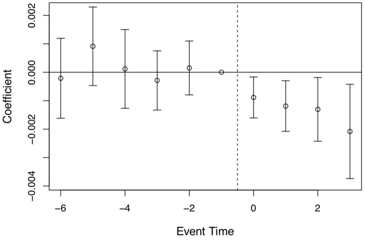
The following figures are from Miller et al. (QJE 2021), via the *Mixtape*. The authors estimate the impact of state expansion of Medicaid under ACA on the annual mortality rates of older persons under 65 in the U.S.

A causal interpretation of DD assumes changes over time in states that did *not* expand Medicaid provide the counterfactual for those that did.

They find a 0.13 percentage-point decline in annual mortality, a 9.3% reduction over the sample mean, as a result of Medicaid expansion.

Event study example: Miller et al (2021)

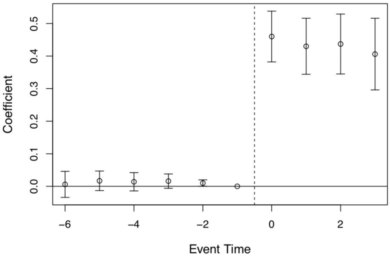
Plotted points are event study coefficients, shown with 95% confidence intervals. (Time zero is the first year of expansion). Outcome: mortality rate



There is no evidence these states' mortality rates were on different trajectories prior to Medicaid expansion.

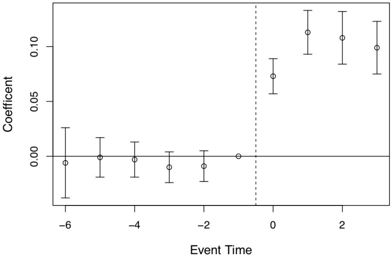
Event study example: Miller et al (2021)

The authors first look for a "first stage": did the expansion of Medicaid actually increase rates of eligibility for Medicaid? Did it increase Medicaid coverage? Did it lower the uninsured rate? Here: **eligibility**



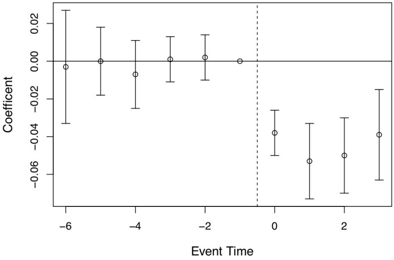
Event study example: Miller et al (2021)

Here: Medicaid **coverage rates**



Event study example: Miller et al (2021)

Here: **uninsured rates**



Taken together, these graphs are compelling: Medicaid expansion increased eligibility and coverage, and reduced the uninsured. One would hope to see these first stage effects before expecting an effect on health outcomes.