xtevent: Estimation and Visualization in the Linear Panel Event-Study Design

Simon Freyaldenhoven[1] Christian Hansen[2]
Jorge Pérez Pérez[3] Jesse M. Shapiro[4]
Constantino Carreto[3]

[1]Federal Reserve Bank of Philadelphia

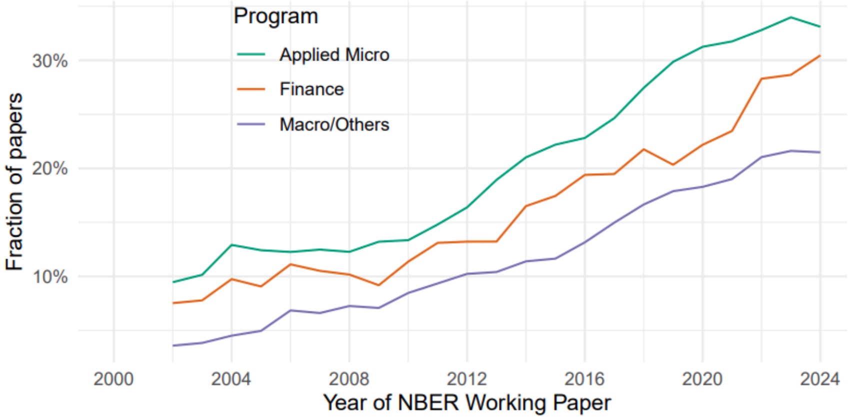
[2]University of Chicago

[3]Banco de México

[4]Harvard University and NBER

Motivation

- Event studies and DinD increasingly popular in economics. The plot shows the percentange of NBER working papers using DinD and event-study methods.



Source: Goldsmith-Pinkham (2024) Figure 5a

Today

- We introduce the `xtevent` package for estimation and visualization of event studies (Freyaldenhoven et al. 2025)
- The package implements the suggestions on the construction of event-study plots from Freyaldenhoven et al. (2021).
- `xtevent` offers the following features:
  - Estimation with general policy variables in both staggered and non-staggered adoption settings.
  - Adjustments for pre-event trends: trend extrapolation (Dobkin et al. 2018) or proxy variables (Freyaldenhoven, Hansen, and Shapiro 2019).
  - Enhancements to event-study plots.
- Our package complements many other recent contributions in Stata (`csdid`, `did_imputation`, `didmultiplegt`, etc.), as well as the official `xthdidregress` command.

Setup

Linear Panel Model

- Let z_{it} be an scalar policy, e.g., minimum wage
- Estimate

$$y_{it} = \alpha_i + \gamma_t + q_{it}'\psi + \sum_{m=-\infty}^{\infty} \beta_m z_{i,t-m} + C_{it} + \varepsilon_{it} \qquad \text{(linear panel model)}$$

- y_{it} scalar outcome, e.g., employment
- Unit fixed effects α_i and time fixed effects γ_t
- Observed controls q_{it}
- Unobserved confound C_{it} potentially related to policy z_{it}.
- Unobserved error ε_{it} unrelated to policy z_{it}
- $\{\beta_m\}_{m=-\infty}^{\infty}$ summarize the magnitude of the dynamic effects.

Estimating equation

- Estimate

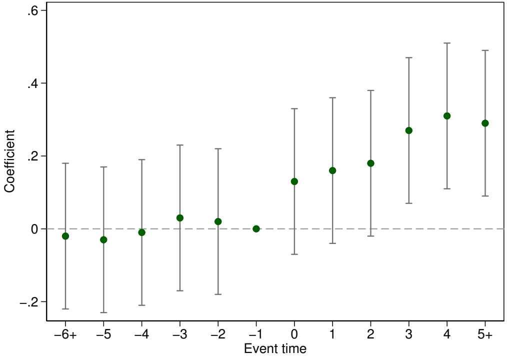
$$y_{it} = \sum_{k=-(B-1)}^{A-1} \delta_k \Delta z_{i,t-k} + \delta_A z_{i,t-A} + \delta_B (1 - z_{i,t+B-1}) + \alpha_i + \gamma_t + q'_{it}\psi + C_{it} + \varepsilon_{it}$$

(estimating equation)

- Δ denotes the first difference operator
- Normalize $\delta_{-1} = 0$
- Plot $\{(k, \hat{\delta}_k)\}_{k=-B}^{A}$
  - $A =$ number of periods **A**fter to plot
  - $B =$ number of periods **B**efore to plot
- NB: For algebra, see Freyaldenhoven et al. (2021) or Schmidheiny and Siegloch (2020)

Trend adjustment | IV estimator | Sun and Abraham | Imputation of policyvar

Definition of plot



Points on plot correspond to $\{(k, \hat{\delta}_k)\}_{k=-B}^{k=A}$.

Default plot and recommendations

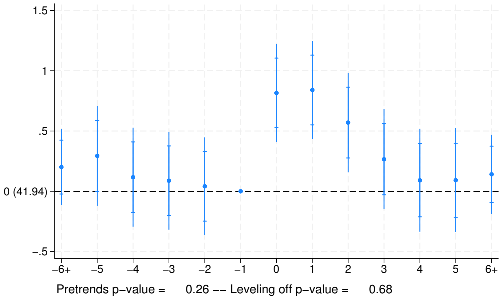
Default plot in `xtevent`

Using the dataset from Freyaldenhoven et al. (2025):

```
xtevent y x, panelvar(id)
 timevar(t) policyvar(z)
 window(5) impute(stag) plot
```

Where:

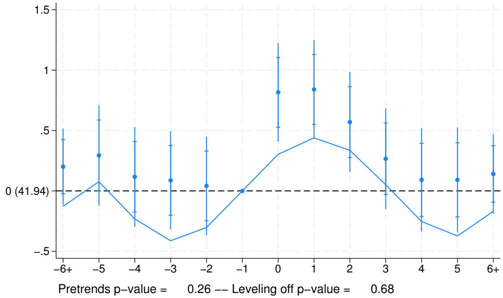
- `y` is the outcome variable
- `x` is a covariate
- `policyvar()` declares the policyvar. In staggered adoption, it is 1 when the observation is treated and onward, 0 otherwise.
- `impute()` indicates an imputation rule
- `plot` displays the event-study plot



Recomendations

The default plot includes our recommendations:

1. Normalization
2. Outcome variable level
3. Uniform inference
4. Overidentification and testing
5. Least wiggly path of confounds consistent with the estimates:
   `xteventplot, smpath(line)`



Trend adjustment

Trend adjustment

Estimating equation

C_{it} can be written as

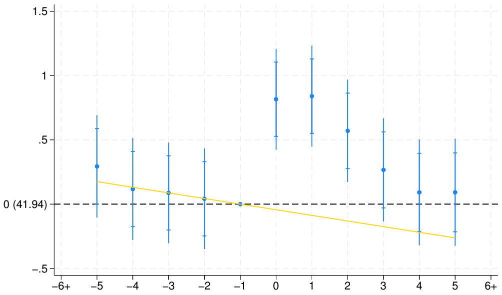
$$C_{it} = \tilde{\alpha}_i + \tilde{\gamma}_t + q_{it}'\tilde{\psi} + \sum_m \phi' f(m) z_{i,t-m}$$

- for a known set of basis functions $f(\cdot)$ and unknown parameters $\tilde{\alpha}_i, \tilde{\gamma}_t$ and $\tilde{\psi}$
- then, the estimation equation may be estimated by including the appropriate terms from the equation above directly in a regression model or by GMM in a second step following the estimation equation via two-way fixed effects.
- Intuitively, suppose that a trend in event time can approximate the confound. In that case, we can learn about the trend in periods where the policy is inactive and extrapolate it to later periods. The differences between the outcome variable and the extrapolated trend are then informative of the policy effects (Dobkin et al. 2018).

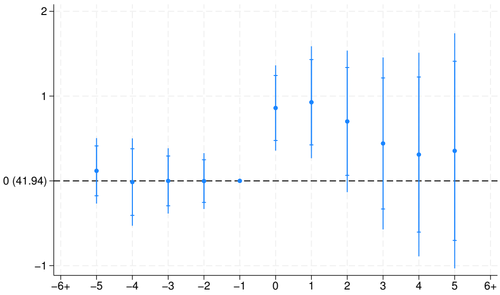
Trend adjustment

```
. xtevent y x, panelvar(id) timevar(t) policyvar(z)
  window(5) impute(stag) trend(-3, method(gmm) saveoverlay)
```

`xteventplot, overlay(trend)`



`xteventplot`



IV estimator

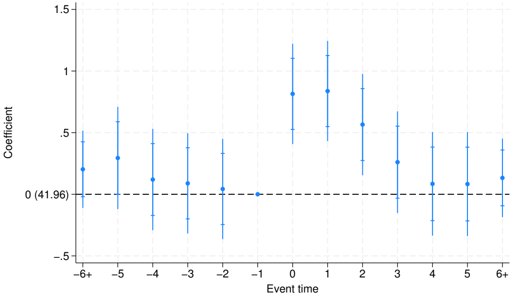
IV estimator

- `xtevent` allows estimation when a pre-trend is present using the instrumental variables estimator of Freyaldenhoven et al. (2019).
- Consider a design in which unobserved confounds may be related both to the outcome and to the policy variable of interest. Estimating equation
- Option `proxy()` to indicate the proxy variable(s) for the confound.
- `xteventplot` create several plots to illustrate this estimator:
  - `xteventplot, y` to create an unadjusted event-study plot for the outcome.
  - Option `proxy()` to illustrate the dynamics of the proxy by creating an event-study plot for the proxy.
  - Option `overlay(iv)` to create a plot that aligns the dynamics of the proxy and the outcome between the coefficient used as an instrument and the normalized coefficient.
  - `xteventplot` with no additional options, we create a plot that shows the coefficients of the outcome after subtracting the rescaled event-study coefficients for the proxy.

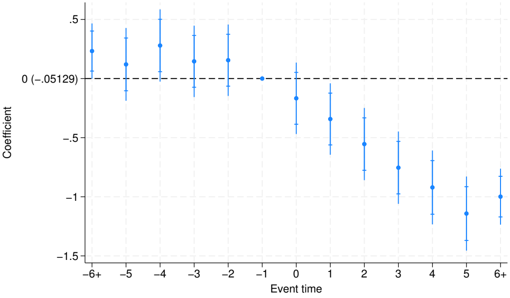
IV estimator

```
xtevent y x, panelvar(id) timevar(t) policyvar(z)
   window(5) impute(stag) proxy(x)
```

```
xteventplot, y
ytitle("Coefficient")
xtitle("Event time")
```

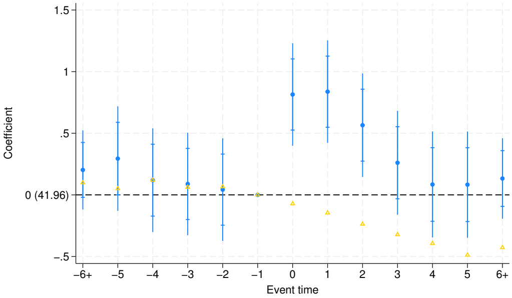



```
xteventplot, proxy
ytitle("Coefficient")
xtitle("Event time")
```

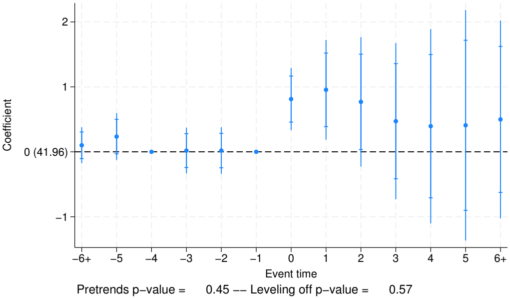


IV estimator

```
xteventplot, overlay(iv)
ytitle("Coefficient")
xtitle("Event time")
```




```
xteventplot,
ytitle("Coefficient")
xtitle("Event time")
```



Sun and Abraham estimator

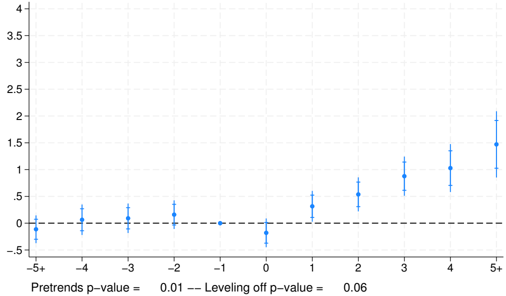
SA estimator

- The model in Estimating equation assumes that the causal effect of the policy is homogeneous over units *i*.
- Recent literature has highlighted that if treatment effects are heterogeneous by treatment time, then the effects estimated may not be properly weighted averages of the cohort-level treatment effects.
- Sun and Abraham (2021) propose estimating event studies for each treated cohort separately, comparing each one with an untreated cohort and then averaging the effects, weighting by the percentage of treated units in each cohort to arrive at a treatment effect on the treated.
- `xtevent` implements this estimator with the option `sunabraham` or with the options `cohort()` and `control_cohort()`.

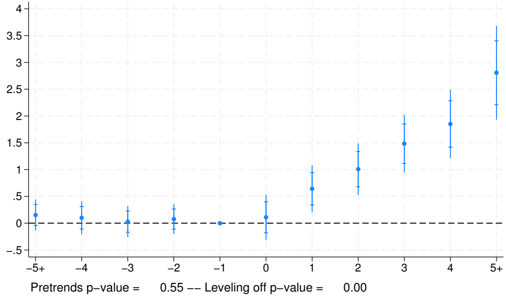
SA estimator

Using simulated data from Borusyak et al. (2024).

```
xtevent Y, panelvar(i) timevar(t)
policyvar(D) impute(stag)
window(4) plot
```




```
xtevent Y, panelvar(i) timevar(t)
policyvar(D) impute(stag)
window(4) plot sunabraham
```



Thank you!

Scan QR to download the paper

Scan QR to download the latest version of the package

Imputation of policyvar

- Because Estimating equation includes leads and lags of the policy variable z_{it} and its first difference, the estimation sample may be smaller than the entire sample available.
- This may imply that we must restrict the estimation window to calculate the necessary leads and lags of z_{it}.
- However, the user may have additional information that allows imputation of the policy variable. `xtevent` allows for several imputation schemes by the `impute` option.
- For instance, to impute the policy variable under staggered adoption, we use the `impute(stag)` option. `xtevent` verifies that the policy variable follows staggered adoption. If so, `xtevent` imputes the policy variable outside the observed time range. Then, it uses the imputed policy variable to generate the event-time dummies and endpoints.

Imputation of policyvar

```
list id t z z_d v_eq_m6 - v_eq_m1 if id==19 & t>=29
```

id	t	z	z_d	v_eq_m6	v_eq_m5	v_eq_m4	v_eq_m3	v_eq_m2	v_eq_m1
19	29	0	0	1	0	0	0	0	0
19	30	0	0	1	0	0	0	0	0
19	31	0	0	1	0	0	0	0	0
19	32	0	0	0	1	0	0	0	0
19	33	0	0	0	0	1	0	0	0
19	34	0	0			0	1	0	0
19	35	0	0				0	1	0
19	36	0	0					0	1
19	37	1	1						0
19	38	1	0						
19	39								
19	40								

Imputation of policyvar

Using the `impute(stag)` option:

id	t	z	z_imputed	z_imputed_d	v_eq_m6	v_eq_m5	v_eq_m4	v_eq_m3
19	29	0	0	0	1	0	0	0
19	30	0	0	0	1	0	0	0
19	31	0	0	0	1	0	0	0
19	32	0	0	0	0	1	0	0
19	33	0	0	0	0	0	1	0
19	34	0	0	0	0	0	0	1
19	35	0	0	0	0	0	0	0
19	36	0	0	0	0	0	0	0
19	37	1	1	1	0	0	0	0
19	38	1	1	0	0	0	0	0
19	39		1	0	0	0	0	0
19	40		1	0	0	0	0	0

Default plot