

Identification and Collinearity in Panel Event Studies

Bas Machielsen

Setup

Let G_i denote unit i 's first treatment period and define event time as $k_{it} = t - G_i$. For an event time k , let

$$D_{it}^k = \mathbf{1}\{t - G_i = k\}.$$

A conventional two-way fixed-effects event-study regression is

$$Y_{it} = \alpha_i + \lambda_t + \sum_{\substack{k=-K \\ k \neq -1}}^L \delta_k D_{it}^k + u_{it},$$

where event time $k = -1$ is the omitted reference period.

Three distinct questions matter:

1. Are the columns of the regression matrix linearly independent?
2. Are untreated comparison observations available for a given cohort and period?
3. Does the regression coefficient identify a causally meaningful average effect?

These questions should not be conflated.

Why the Earlier Collinearity Argument Was Incorrect

For example, the event-time-zero dummy can be written as

$$D_{it}^0 = \sum_g \mathbf{1}\{G_i = g\} \mathbf{1}\{t = g\}.$$

The terms on the right are **interactions** between treatment-cohort dummies and calendar-time dummies. A model containing unit and time fixed effects contains their main effects, not all unit-by-time or cohort-by-time interactions. Therefore, this identity does not establish that D_{it}^0 is a linear combination of the included fixed-effect columns.

With staggered adoption, event time and calendar time generally vary independently enough for event-time coefficients to be computed. A never-treated group is consequently not required merely to make the design matrix full rank.

Comparison Support Under Staggered Adoption

Suppose one cohort is treated in 2023 and another in 2024. In 2023, the 2024 cohort is not yet treated and can serve as a contemporaneous comparison for the 2023 cohort, under the relevant parallel-trends and no-anticipation assumptions.

This comparison support changes over time:

- Never-treated units can provide comparisons throughout the sample.
- Not-yet-treated units can provide comparisons only until their own treatment.
- After the last cohort is treated, no untreated comparison observations remain.

A standard TWFE regression may nevertheless produce coefficients after that point by using already-treated units as comparisons. When treatment effects vary across cohorts or over event time, these “forbidden comparisons” need not identify the desired average treatment effect. Cohort-time estimators avoid this problem by explicitly using never-treated or not-yet-treated observations and by reporting effects only where suitable support exists.

Cases Where Collinearity Does Occur

Common treatment date

If every unit is treated in the same period G , then

$$D_{it}^k = \mathbf{1}\{t = G + k\}.$$

Each event-time dummy is then exactly a calendar-time dummy. With time fixed effects, the event-time coefficients are perfectly collinear and cannot be separated from common time shocks. No panel DiD comparison identifies the treatment effect in this design.

Saturated event time when every unit is eventually treated

There can also be an additional normalization problem when all units are eventually treated and event-time dummies exhaust every observed relative period. Because $k_{it} = t - G_i$,

$$\sum_{k \neq -1} (k + 1) D_{it}^k = t - G_i + 1.$$

The right-hand side is an additive combination of a time component, a unit component, and a constant. It is therefore absorbed by unit and time fixed effects. Omitting the $k = -1$ indicator handles the usual reference-category normalization but does not, by itself, resolve this linear-trend dependence in a fully saturated all-treated design. Software must impose another normalization or omit another column. Using a finite event window, binning endpoints, or including never-treated observations changes the exact rank condition, but it does not create causal comparison support where none exists.

Practical Guidance

- Diagnose matrix rank separately from causal identification.
- State which observations act as comparisons for each cohort and event time.
- Prefer never-treated or not-yet-treated comparisons in staggered-adoption DiD.
- Do not report event times beyond the available untreated support as though they were identified by clean DiD comparisons.
- Use cohort-robust event-study estimators when effects may be heterogeneous.
- Keep finance-style event studies conceptually separate: they construct a counterfactual return from a pre-event market or factor model rather than from untreated units in a panel DiD design.