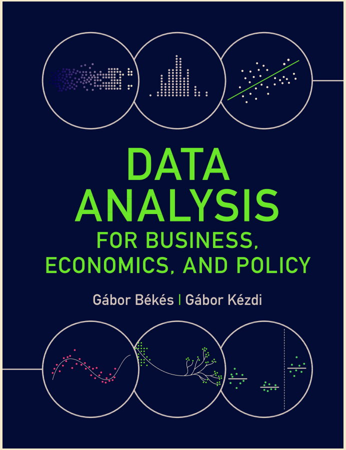
Békés-Kézdi: Data Analysis, Chapter 24: Appropriate Control Groups for Panel Data



Data Analysis for Business, Economics, and Policy

Gábor Békés (Central European University)
Gábor Kézdi (University of Michigan)

Cambridge University Press, 2021

gabors-data-analysis.com

Central European University

Any comments or suggestions:
gabors.da.contact@gmail.com

When and Why to Select a Control Group in xt Panel Data

- Observational xt panel data,
- Diff-in-diffs or FD or FE regression to estimate the effect of an intervention
- Say, binary treatment variable, this means that we use all observations with $x = 0$ to learn about the counterfactual.
  - some subjects were never treated
  - as well as time periods of treated subjects in which they were not treated.

- Big question: are including all observations for all subjects the best we can do to get the best control group?

When and Why to Select a Control Group in xt Panel Data

- Observational xt panel data,
- Diff-in-diffs or FD or FE regression to estimate the effect of an intervention
- Say, binary treatment variable, this means that we use all observations with $x = 0$ to learn about the counterfactual.
  - some subjects were never treated
  - as well as time periods of treated subjects in which they were not treated.

- Big question: are including all observations for all subjects the best we can do to get the best control group?
- Seen before: matching

When and Why to Select a Control Group in xt Panel Data

- Appropriate control group: parallel trend - ie same in terms of difference not level
- Can select a subset of the $x = 0$ observations that have a better chance of satisfying the parallel trends assumption.

- Selecting the control group: untreated observations to learn about the counterfactual average outcome of the treatment group.

- Two different setups, two methods

Synthetic control

Comparative Case Studies - with synthetic control

- An intervention happens for one unit
  - Typically, one country or one part of a country
    - A state, a region, a county
- Want to measure the effect of the intervention
  - On that unit
- Only one treated unit: can one make statistical inference?
  - Generalization of results to potential other intervention under very similar circumstances ("represented by our data")
  - External validity requires more: whether circumstances are similar

Similar to Diff-in-diffs

- Diff-in-diffs compares change of outcomes
  - Average change in treated units
  - Average change in comparison units
- Average change in comparison units estimate expected counterfactual change in treated units
  - How outcomes would have been expected to change without the treatment
- One treated unit here
- How to select comparison unit(s)?

The Synthetic Control Idea

- Start with a set of potential control units
  - The "donor pool"
- Create a single synthetic control unit by combining *all* or *some* of units from the donor pool
- This synthetic control unit estimates the counterfactual outcome in the treated unit
  - How the outcome would have changed in the treated unit if it was not treated
- Find best combination via algorithm
  - The one that is most similar to treated unit
  - In terms of pre-treatment outcomes and other covariates

Weighted Average from Donor Pool

- Synthetic control is a weighted average of control units
- Combine outcomes of untreated units

$$\hat{Y}(0) = \sum_{j=2}^{n} w_j Y_j$$

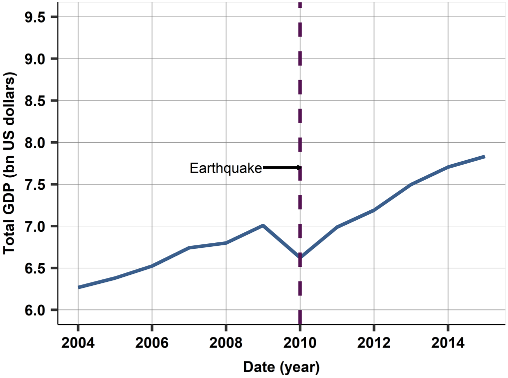
- The w_j are the optimal weights
  - Many weights can be zero
- Weights are time-invariant (same before/after)

Estimating the Weights

- The essential part of the method
- Use all pre-intervention outcome variables, X
  - Plus, potentially time-invariant variables
- Find weights that give a weighted average of all those
  - That is closest to their values for the treated unit

- By collecting all key outcome variables and choosing weights for these outcome variables, we also hope to take care of the common trend assumption

Case study - Total GDP in Haiti

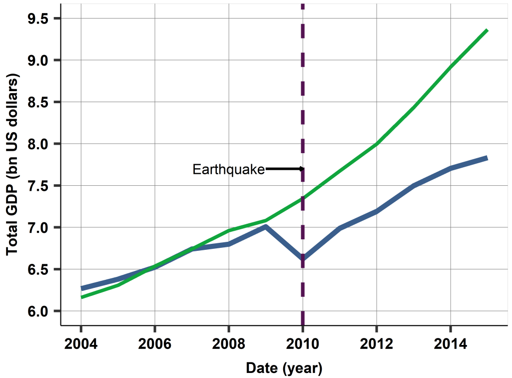


- Note: Total GDP, US dollars, constant prices of 2010, in billions.
- Source: haiti-earthquake dataset.

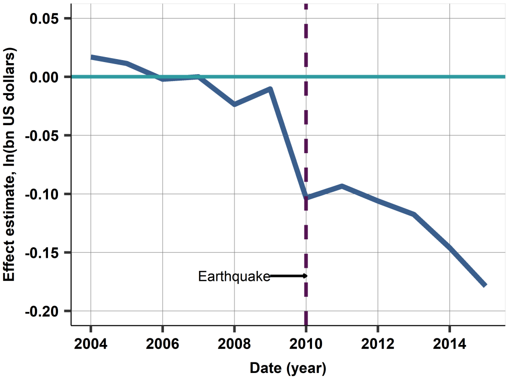
Case study -Synthetic country weights

- Synthetic Haiti:
  - 47% Togo
  - 23% Burundi,
  - 21%Cameroon,
  - 9% Moldova,
  - 0.2% Liberia

Case study - The effect of the 2010 earthquake on the total GDP of Haiti. Synthetic control estimate



Total GDP in Haiti and synthetic Haiti



Log GDP difference between Haiti and synthetic Haiti

Take-Away

- Synthetic controls is a potentially powerful method
  - In comparative case studies
  - That analyze the effect of an intervention in one unit
- Creates single synthetic control unit
  - As a weighted average of (some) comparison units
    - from the "donor pool"
  - Weights to minimize distance between treated and controls
    - In pre-treatment outcomes + other covariates
    - Maximize covariate balance
- Based on the assumption of selection on observables
  - Including common trend
- Can be generalized to multiple treated units
- A new method
  - Many potential applications ahead
  - Caveats and weaknesses not fully understood

Event study

Event studies: transformation of time

- Closer focus on treatment
- That happens at multiple time periods = staggered treatment
  - This could be an issue = conditions vary across periods = confounding
- Does not happen to some units

Event studies: transformation of time

- FD models are nice, but with leads and lags, not easy to interpret.
  - Great deal of new research potential problems with FD and FE models with time dummies
- We show an alternative way of thinking about the panel: reorganize it around the treatment

Event studies: Thinking carefully re control group

- In xt panel methods, all units included, even those never treated
  - Never treated: accidentally vs not suitable Why? Examples?

- Twist: thinking about how explicitly select a control group
- Finding untreated units that are similar to treated ones just before the treatment
  - Other options exist.

Event studies: transformation of time

- The event study method means reorganizing the data to focus on time periods before and after the intervention.
  - then using panel data methods, like FE or FD.
  - We change how we write down a model, not what we estimate
- Key issue: different subjects becoming treated at different times, x_{it} turns from zero to one at different values of t.
- Re-define time in the data. Instead of natural time, we define and use event time: time around the intervention.
  - Pre-intervention time periods: 1 stands for one time period before the intervention
  - The time period when the intervention happens is labeled as event time zero
  - Post-intervention periods : +1 for one time period after the intervention

Event studies: transformation of time

- To estimate the effect of the intervention, we need to specify an xt panel regression in first differences (FD) or fixed effects (FE).
- Let's write down an FD event study regression without confounder variables.
- Using data with observations between event time periods $-T_{before}$ to $+T_{after}$, and only observations that become treated, it is

$$\Delta y_{it}^{E} = \alpha + \sum_{0}^{s_{max}} \beta_s D_{is} + \sum_{s_{min}}^{1} \gamma_s D_{i(-s)} \qquad (1)$$

  - In this regression, the D_{is} variables are binary indicators for each event time period s (D stands for dummies).
  - So, for example, $D_{i1} = 1$ if the observation is for subject i and event time 1.
  - Just as with any other xt panel regression, we can add other right-hand-side variables if we want to.

Selecting a Control Group in Event Studies

- Parallel trends assumption: need a control group with similar pre-intervention changes in y.
- But the observations that could make up the control group don't have an intervention.
- How can we find subjects with pre-intervention changes that are similar to what we see in the treated group, if those subjects didn't experience an intervention?
- With a trick.

Selecting a Control Group in Event Studies

- Select the control group by focusing on the average patterns of pre-intervention outcome changes in the treatment group
  - and possibly other variables.
- Define pseudo-interventions: points in time for untreated subjects that are preceded by changes in y that are, on average, similar to actual pre-intervention changes among treated subjects.
- Control group = counterfactual – very similar to treated just before the treatment

Selecting a Control Group in Event Studies

1. Transform panel data into event time, with intervention at 0
2. Define a criterion of what similar means
3. Use an algorithm to identify these control units.
4. Add them to the regression

Estimation

- With the appropriate control group, and a pseudo-intervention time for each control subject, define event time just as we did for treated subjects.
- Event-time regression to estimate average changes in y before and after the intervention, separately for treated and untreated subjects.

- Building on the event study regression we specified for treated subjects only, we would add an interaction term to the dummy variables of pre-intervention and post-intervention time periods:

Estimation equation

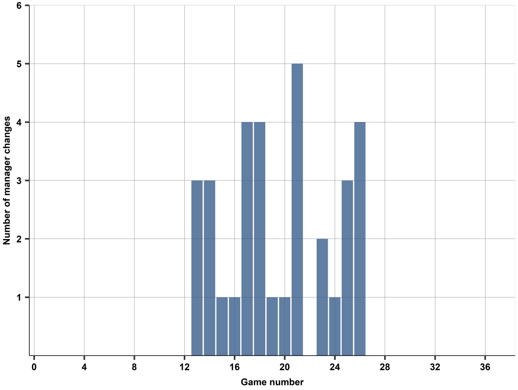
$$\Delta y_{it}^{E} = \alpha + \sum_{0}^{s_{max}} \beta_s D_{is} + \sum_{s_{min}}^{1} \gamma_s D_{i(-s)}$$
$$+ \eta treated_i + \sum_{0}^{s_{max}} \delta_s treated_i \times D_{is} + \sum_{s_{min}}^{1} \phi_s treated_i \times D_{i(-s)} \tag{2}$$

- β coefficients show the average post-intervention changes among untreated subjects – for them, these are the average changes after their designated pseudo-interventions.
- The effect estimates are the δ coefficients: show the average difference in post-intervention changes between treated and untreated subjects, – the time period of the treatment (δ_0), the first time period after the treatment (δ_1), etc.

Case study – Estimating the Impact of Replacing Football Team Managers

- Football managers may be sacked during the season
- Source: football dataset. English Premier League, 11 seasons from 2008–2009 to 2018–2019.
- N=33 manager changes.
- The numbers of interventions by game week. Interventions are manager changes with 12 games before and after in the season without another manager change.

Case study - The numbers of manager changes by game number



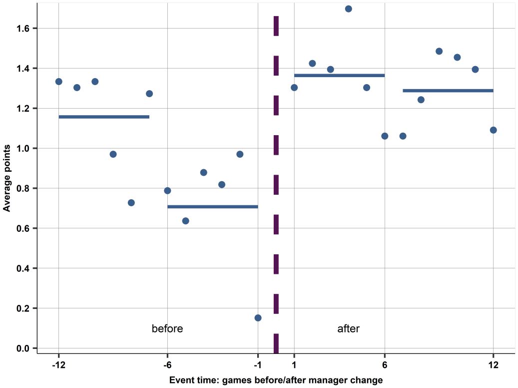
- Note: The numbers of interventions by game week. Interventions are manager changes with 12 games before and after in the season without another manager change.
- Source: football dataset. English Premier League, 11 seasons from 2008–2009 to 2018–2019. N=33 manager changes.

Case study – Estimating the Impact of Replacing Football Team Managers

- Average points before and after manager change and pseudo-intervention
- Average points (wins 3 pts, draws 1 pt, losses 0 pt) by event time (games before/after the management change).
- Consider six-week averages
- Interventions: when a manager is sacked
- Actual manager changes: $33 \times 3 = 99$ observations.
- Pseudo-interventions: $67 \times 3 = 201$ observations.

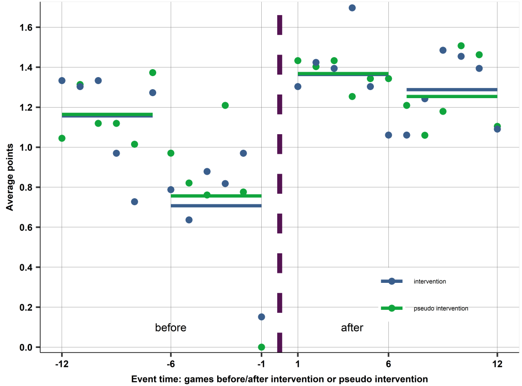
- Source: football dataset. English Premier League, 11 seasons, balanced panel of 12 games before and after 33 manager changes and 67 pseudo-interventions. N=2400 team-games.

Case study - Average points before and after manager replacement



- Note: Average points (wins 3 pts, draws 1 pt, losses 0 pt; over 33 events) before and after manager change by event time (games before/after the manager change).
- Six-week average lines added.

Case study - Average points before and after manager change and pseudo-intervention



- Note: Average points (wins 3 pts, draws 1 pt, losses 0 pt) by event time (games before/after the management change).
- Six-week average lines added.
- Interventions: blue dots and lines. Pseudo-interventions: green dots and lines.

Case study - The effect of replacing managers: FD regressions with pseudo ctrls

- Can estimate difference from graph with an equation
- Combine actual manager changes: $33 \times 3 = 99$ observations and pseudo-interventions: $67 \times 3 = 201$ observations.
- Change in points is regressed on pre-intervention change and two periods (6-game spells) of post intervention changes ...
- ... interacted with treatment.

$$\Delta y_{it}^{E} = \beta_0 + \beta_1 post_{1-6} + \beta_2 post_{7-12} + \beta_3 treat + \beta_4 treat \times post_{1-6} + \beta_5 treat \times post_{7-12} \tag{3}$$

Case study - The effect of replacing managers: FD regressions with pseudo ctrls

- First we estimated a regression separately for the actual manager changes: $33 \times 3 = 99$ observations.
- Then we estimated an analogous regression for the pseudo-interventions: $67 \times 3 = 201$ observations.
- The third regression combines them, including the interaction with treatment.
- The regression includes the following binary variables: *treat* for team-season with actual manager changes (the treatment group), $post_{1-6}$ for 1 to 6 games after the intervention, or pseudo-intervention, and $post_{7-12}$ for 7 to 12 games after the intervention, or pseudo-intervention.
- The formula for this combined regression is

$$\Delta y_{it}^{E} = \beta_0 + \beta_1 post_{1-6} + \beta_2 post_{7-12} + \beta_3 treat + \beta_4 treat \times post_{1-6} + \beta_5 treat \times post_{7-12} \tag{4}$$

Case study - The effect of replacing managers: FD regressions with pseudo ctrls

- The formula for this combined regression is

$$\Delta y_{it}^{E} = \beta_0 + \beta_1 post_{1-6} + \beta_2 post_{7-12} + \beta_3 treat + \beta_4 treat \times post_{1-6} + \beta_5 treat \times post_{7-12}$$

  - β_3 the difference between the treatment and control group in terms of average point change from 7–12 games before to 1–6 before. If we selected the control group well, this should be close to zero.
  - β_4, β_5 = effect estimates – show the difference between treated and control in average point changes from the six-game-interval before to the six-game-interval after, and the change from the six-game-interval after to the next six-game interval.

Case study - The effect of replacing managers: FD regressions with pseudo ctrls

- The formula for this combined regression is

$$\Delta y_{it}^{E} = \beta_0 + \beta_1 post_{1-6} + \beta_2 post_{7-12} + \beta_3 treat + \beta_4 treat \times post_{1-6} + \beta_5 treat \times post_{7-12}$$

  - Here the intercept, β_0, shows the average change in points in the reference time period, from 7–12 games before to 1–6 games before, for pseudo-interventions.
  - β_1 shows the average change in points from 1–6 games before to 1–6 games after, compared to the change in the reference time period, for pseudo-interventions.
  - β_2 shows the average change in points from 1–6 games after to 7–12 games after, compared, again, to the change in the reference time period, for pseudo-interventions.

Case study - The effect of replacing managers: FD regressions with pseudo ctrls

Variables	(1) treatment	(2) control	(3) treatment+control
$post_{1-6}$	1.11**	1.06**	1.06**
	(0.19)	(0.09)	(0.09)
$post_{7-12}$	0.37*	0.34**	0.34**
	(0.16)	(0.09)	(0.09)
treated			-0.00
			(0.10)
$treated \times post_{1-6}$			0.04
			(0.20)
$treated \times post_{7-12}$			0.04
			(0.18)
Constant	-0.45**	-0.45**	-0.45**
	(0.10)	(0.03)	(0.03)
Observations	99	201	300
R-squared	0.33	0.42	0.39

Note: Clustered standard error estimates in parentheses. ** p <0.01, * p <0.05. ** p<0.01, * p<0.05.

Case study - The effect of replacing managers: FD regressions with pseudo ctrls

- Teams perform better after new manager
- But even middle run, just slightly better than before the crisis period.
- When compared to similar teams, experiencing a dip, we see no difference: on average rebound even if manager is kept

- Key innovation: mixing panel data with picking control carefully.
  - Better than just looking at teams with a change
  - Better than including all team-games, ie controls where no manager change would have made sense.

Event studies - Take-away

- Thinking about selecting the appropriate control group
- Synthetic control
  - Finding a mix of untreated units, weight them
  - To yield a single synthetic control unit With similar combined pretend
- Event studies
  - Starting point is xt panel
  - Staggered intervention – create event time
  - Think explicitly about control group - not all units