Panel Data Analysis

Lecture 3: Synth, random effects and multilevel models

10 May, 2016

Prof. Andrew Eggers

Synthetic control methods (synth)

Synth: motivation

How could we use a Diff-in-Diff (DID) to measure the effect of

1. the reunification of Germany on the West German economy?
2. California's 1988 tobacco control program on cigarette sales in California?
3. terrorist conflict in the Basque region of Spain on the Basque economy?

Consider two alternatives:

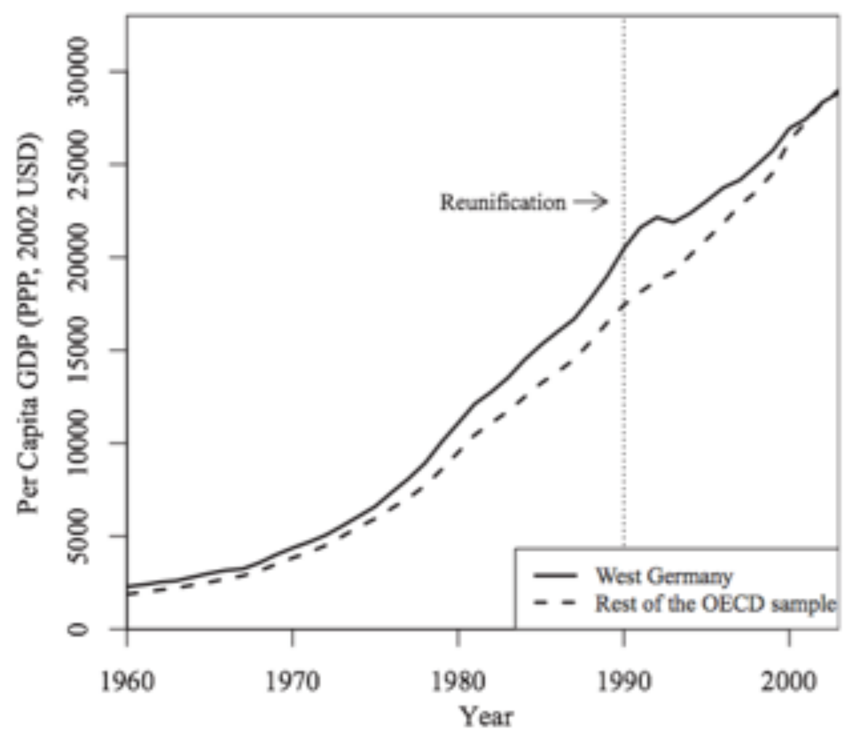
- collect data at disaggregated level, with some units being treated and others not
- collect data at aggregated level, with **one** unit being treated and others not

Synth: basic idea (1)

Consider this DID (Abadie, Diamond, Hainmueller, AJPS 2015):

- treatment group: West Germany (1 unit)
- control group: OECD countries
- treatment: reunification of Germany in 1990

FIGURE 1 Trends in per Capita GDP: West Germany versus Rest of the OECD Sample



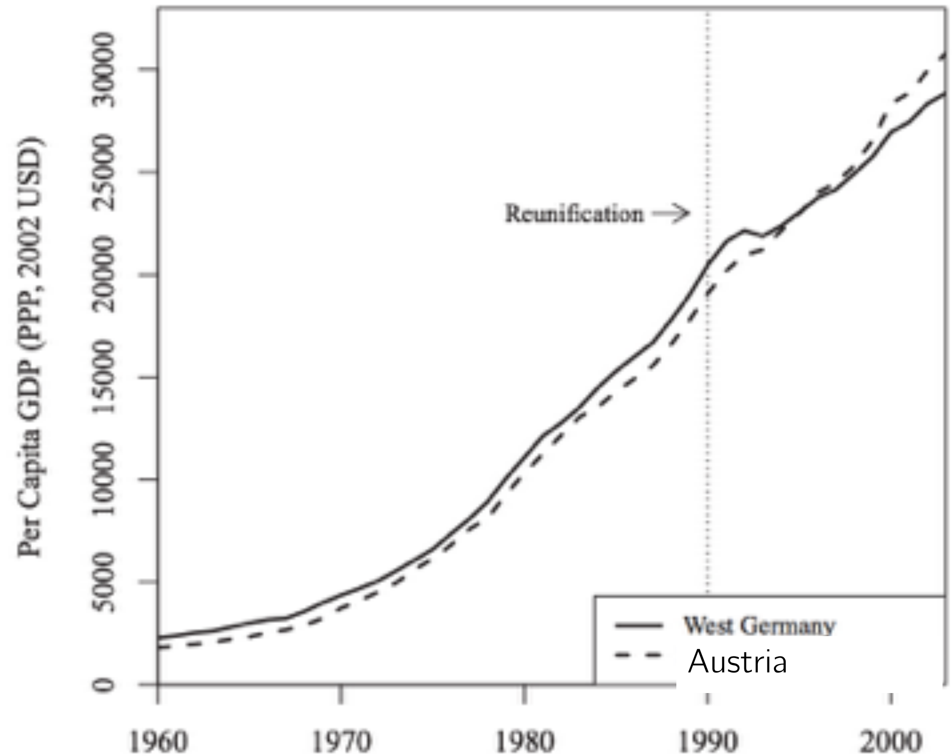
Was reunification good for West German economic growth?

Is the parallel trends assumption plausible?

Synth: basic idea (2)

Rather than comparing West Germany to the average of the whole OECD, what about comparing West Germany to Austria only?

Trends in per capita GDP: West Germany and Austria only



Was reunification good for West German economic growth?

Is the parallel trends assumption plausible?

Synth: basic idea (3)

What about comparing West Germany to *the weighted average of OECD countries that most closely tracks the pre-treatment outcomes of West Germany?*

Define weight vector $W=(w_1,\ldots,w_J)$, one weight for each country.

Define $X_1=(X_{11}, X_{12}\ \ldots\ X_{1k})$ as a $(k \times 1)$ vector of pre-treatment characteristics of **treated unit**.

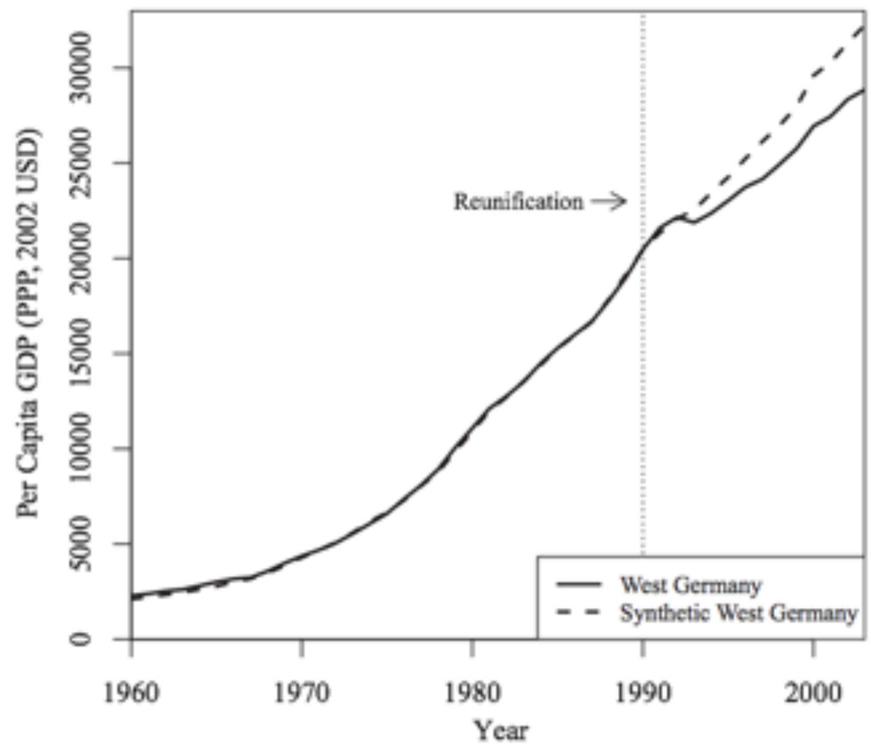
Define X_0 as a $(k \times J)$ matrix of pre-treatment characteristics of **control units** (one column per unit).

Define $v=(v_1, \ldots, v_k)$ as a $(k \times 1)$ vector indicating how important each characteristic is in weighting.

Choose W* to minimize

$$\sum_{m=1}^{k} v_m (X_{1m} - X_{0m} W)^2$$

FIGURE 2 Trends in per Capita GDP: West Germany versus Synthetic West Germany



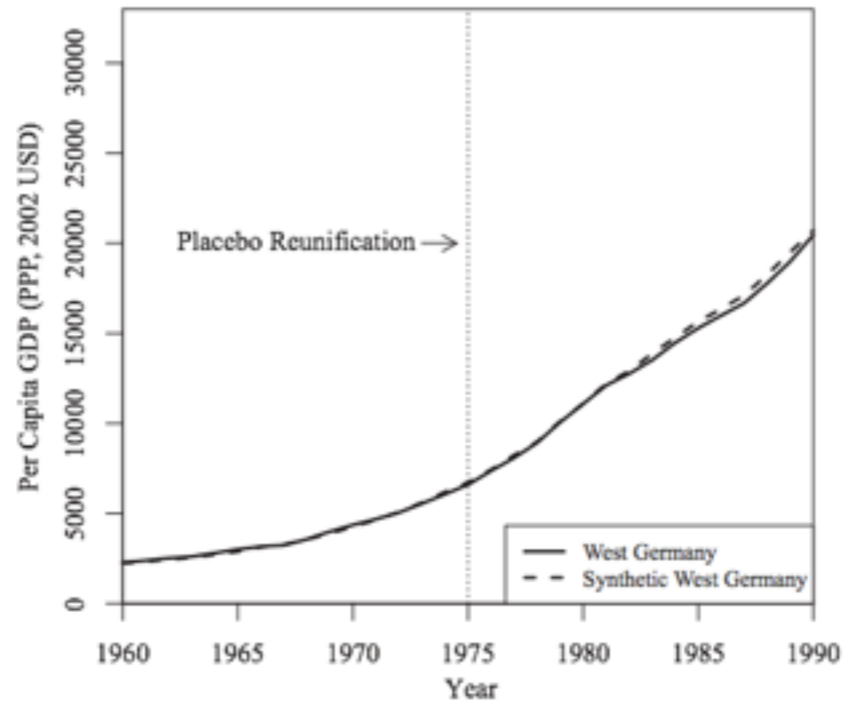
Was reunification good for West German economic growth?

Is the parallel trends assumption plausible?

Synth: falsification tests (1)

Given a long enough period before treatment, can apply same method to a date when treatment **did not** occur.

FIGURE 4 Placebo Reunification 1975–Trends in per Capita GDP: West Germany versus Synthetic West Germany

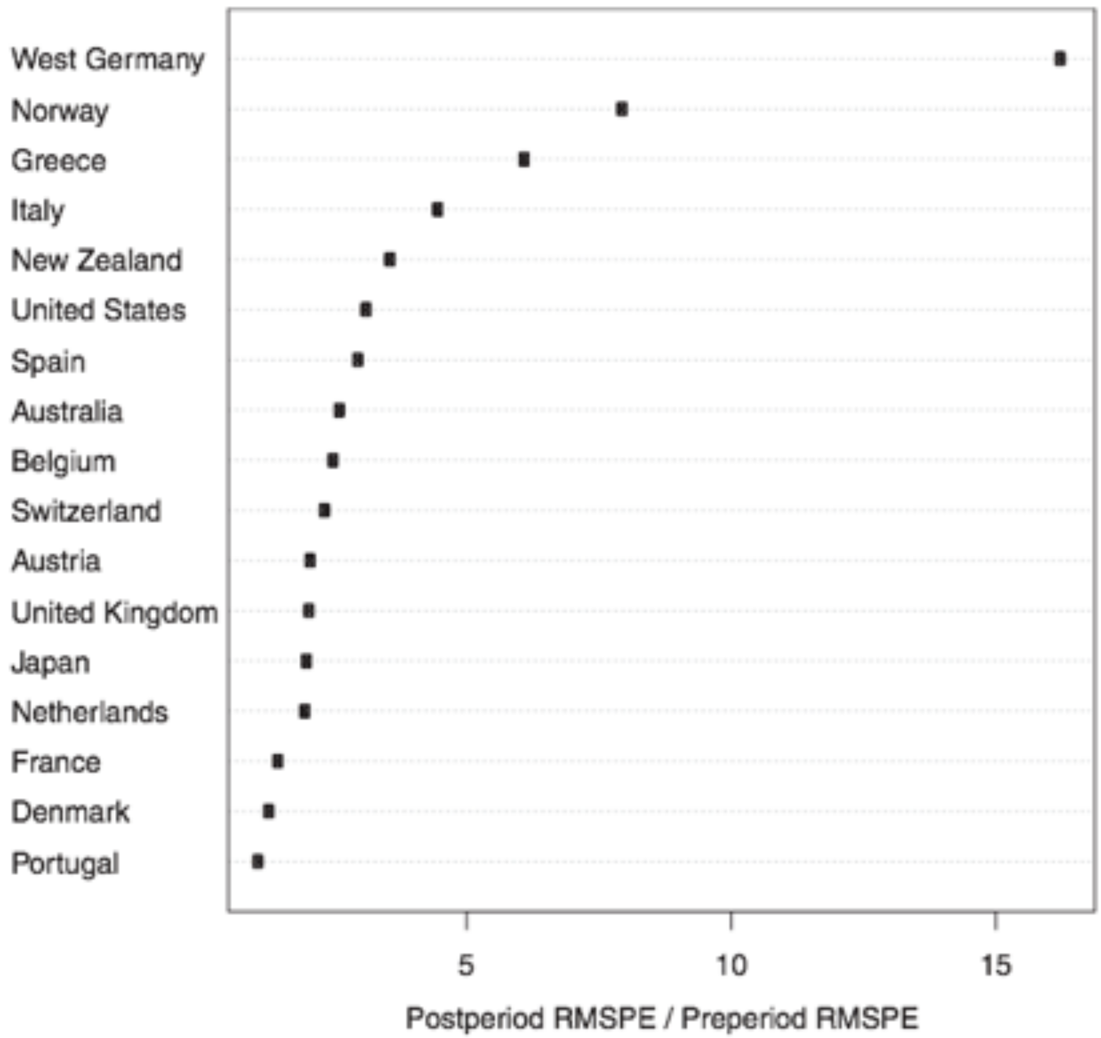


Synth: falsification tests (2)

Can apply same method to other countries where treatment **did not** occur.

Here: how much does a given country deviate from synthetic control group after 1990?

FIGURE 5 Ratio of Postreunification RMSPE to Prereunification RMSPE: West Germany and Control Countries



What about uncertainty/inference/significance/ standard errors?

Regression output provides standard errors, which we use to say whether a result is "significant". What do we do in this case? (1 treated country, 1 synthetic control unit made of 5 countries.)

Abadie et al (2010) response is that there are actually two kinds of uncertainty:

A. **Uncertainty due to sampling.** e.g. Card and Krueger don't observe all fast food restaurants in each state; maybe answers would be different with a different sample of restaurants.

B. **Uncertainty due to unobserved potential outcomes.** e.g. in W Germany example we don't know if W. Germany's GDP/cap would actually have progressed like synth control group in absence of treatment.

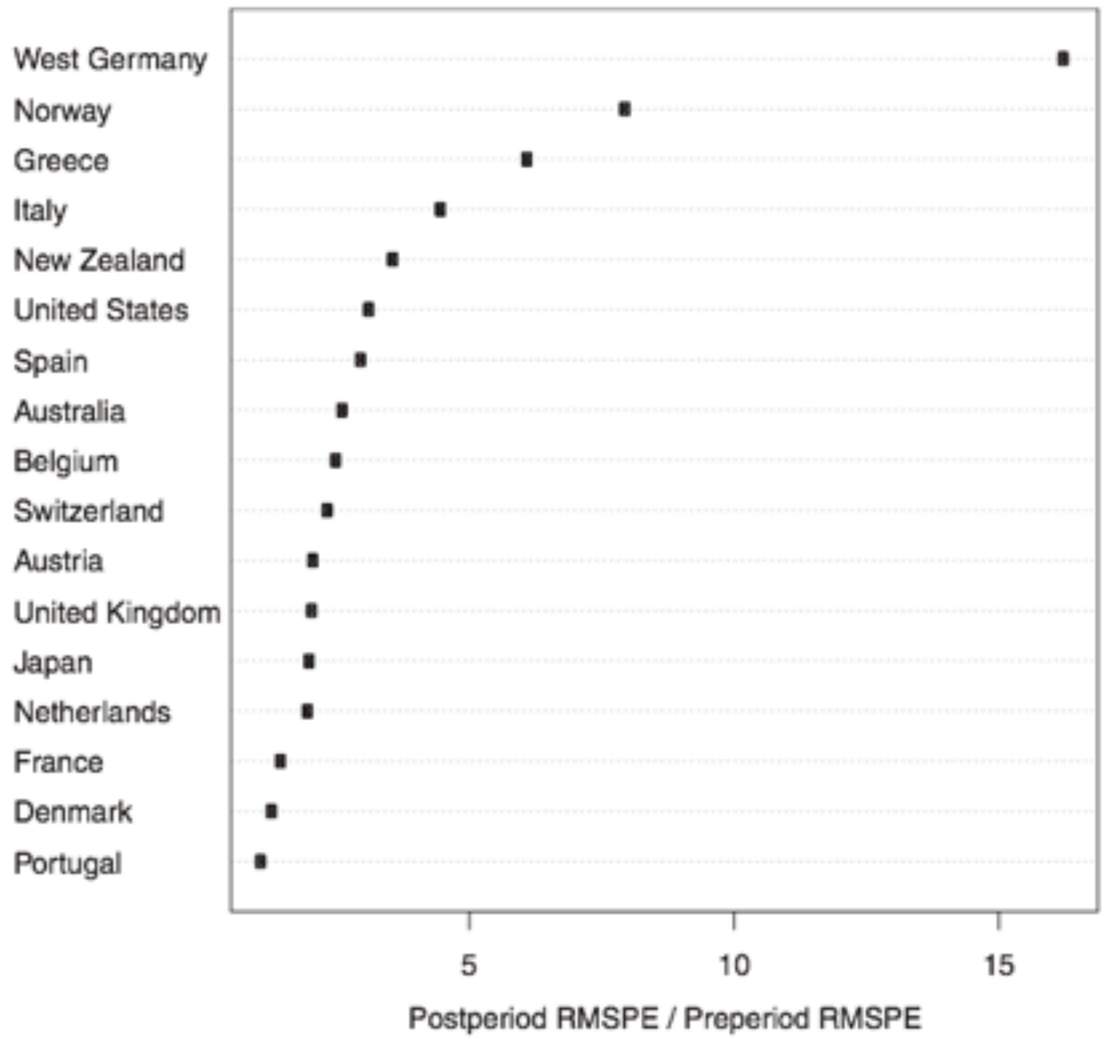
Abadie et al (2010) use randomization inference to characterize uncertainty due to B. They claim not to have uncertainty due to A, because they know the aggregate GDP/cap.

(For more on these issues, see 2015 WP: Abadie, Athey, Imbens, Wooldridge "Finite Population Causal Standard Errors")

Synth: randomization inference (1)

Suppose post-1990 had no effect in any country. Picking a country at random as the “treated” country (placebo), what is the probability of getting a result as large as the one for West Germany?

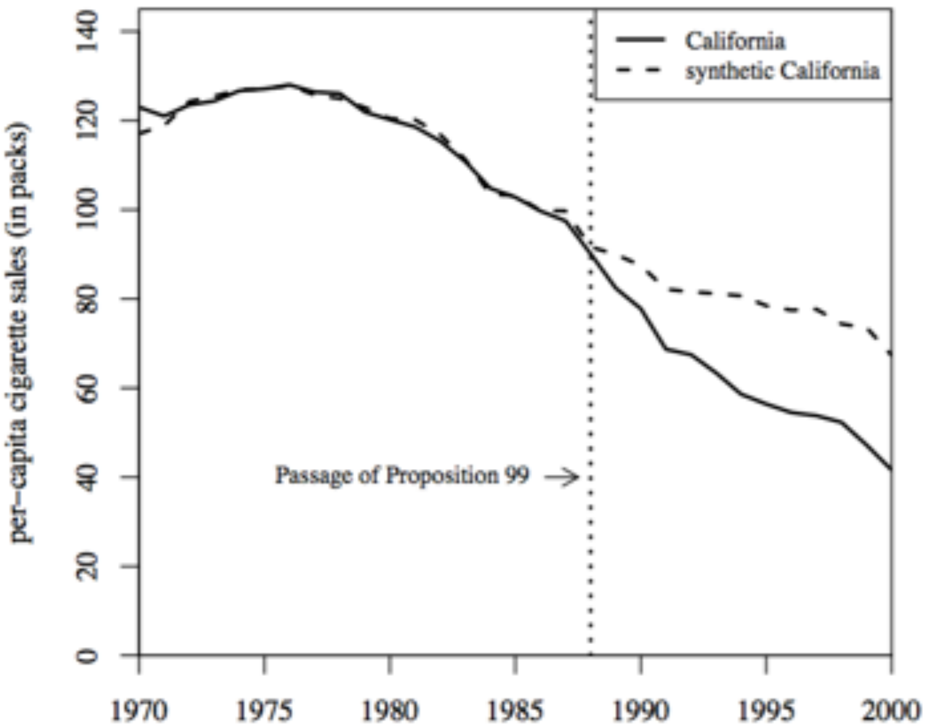
FIGURE 5 Ratio of Postreunification RMSPE to Prereunification RMSPE: West Germany and Control Countries



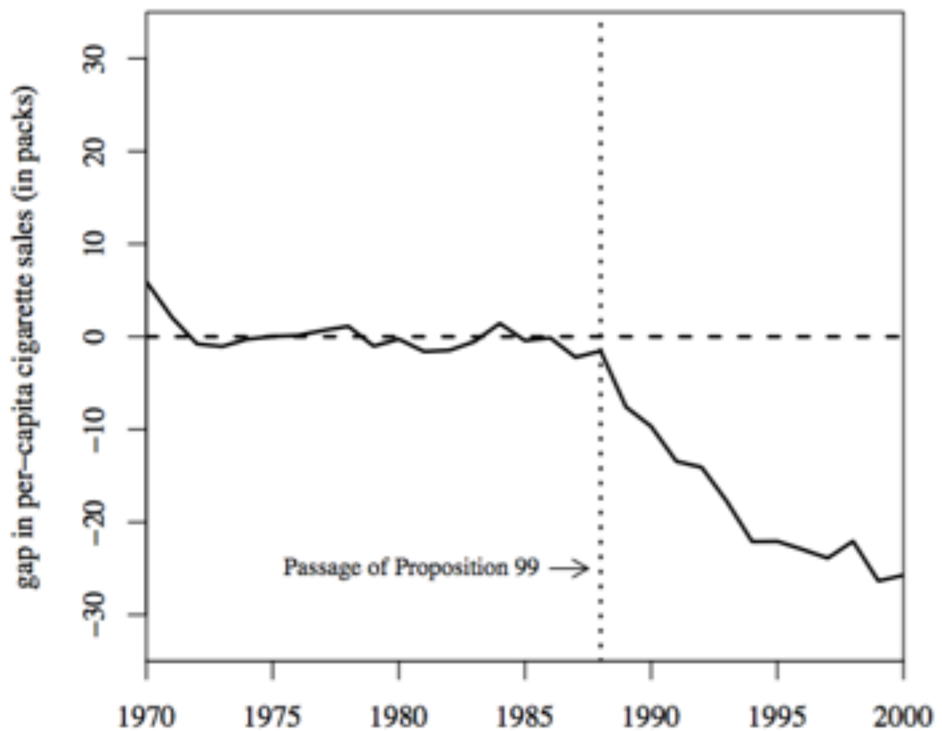
Another synth example: CA's prop 99 and smoking

State	Weight
Colorado	0.164
Connecticut	0.069
Montana	0.199
Nevada	0.234
Utah	0.334

Cigarette sales in CA and synthetic CA

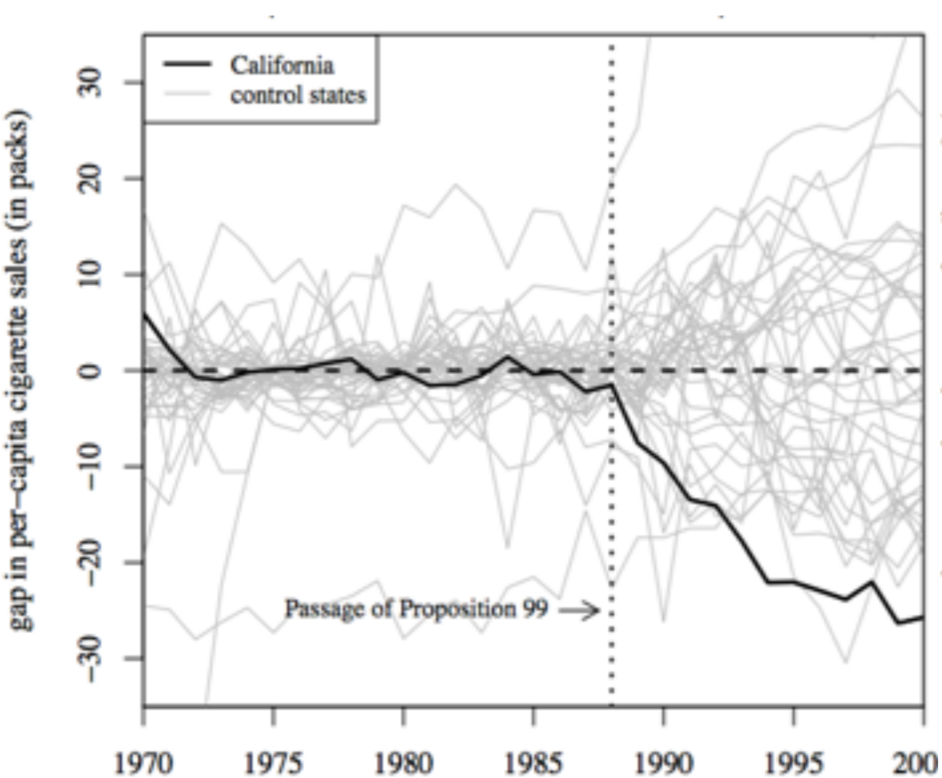


Gap

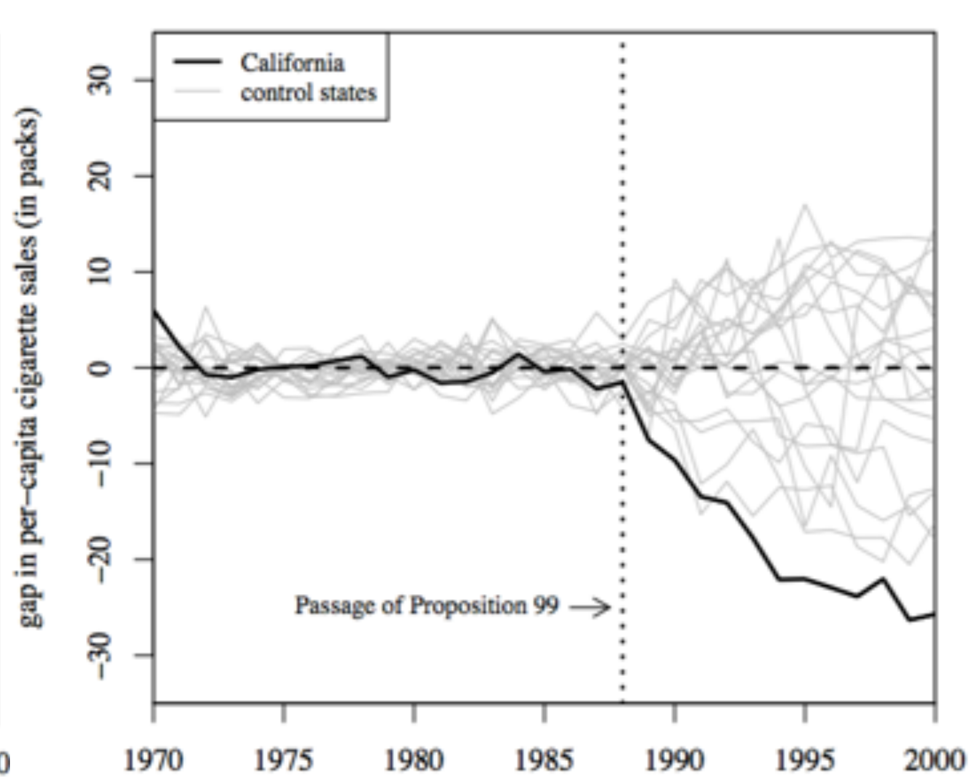


Synth: randomization inference (2)

"Effect" of post-1988 for CA and 38 other states



"Effect" of post-1988 for CA and all other states with at least as good of a pre-"treatment" prediction



Generalizing synth method

Generalization 1: Why restrict weighting approach to the case where there is only 1 treated unit? Couldn't the parallel trends assumption of a DID be made more believable by matching or weighting?

See

- Abadie (2005) "Semiparametric Differences-in-Differences Estimators": matching before DID
- Jens Hainmueller (2011) "Entropy balancing for causal effects": weighting as a generalization of matching $\rightarrow$ build control group whose pre-treatment outcomes (and variance of pretreatment outcomes) match the treated group

Generalizing synth method

Generalization 2: Why restrict weighting approach to the case where treatment is applied only at one point in time? Couldn't the CIA of a panel-DID be made more believable by matching or weighting?

See

- Xu (2015) working paper "Generalized Synthetic Control Method for Causal Inference with Time-Series Cross-Sectional Data"

Synth: further reading

- Abadie et al (2015) AJPS, "Comparative Politics and the Synthetic Control Method"
- Xu (2015) working paper "Generalized Synthetic Control Method for Causal Inference with Time-Series Cross-Sectional Data"
- on inference issues: Abadie et al (2014) working paper "Finite Population Causal Standard Errors"

Random effects vs. fixed effects

Recall motivation for fixed effects

General:

Task: Causal inference in grouped data.

Problem: Group-specific unobservable characteristics affect (are related to) both treatment and potential outcomes.

Solution: Include fixed effects (or dummy variables) for each group.

Conditional Independence Assumption: Within groups, treatment unrelated to potential outcomes.

One-way example:

- **Task:** What is effect of defendant's race on sentencing decision? (Judges hear multiple cases.)
- **Problem:** Judges who are more likely to hear cases against black defendants may be systematically more lenient or strict.
- **Solution:** Judge fixed effects.
- **CIA**: Cases against black and white defendants heard by the same judge are comparable.

Or Levitt on campaign spending (candidate-pair fixed effects), or Ansell on house prices and preferences (year and city fixed effects, via first differences)

"Between" and "within" estimation

Recall "deviation from means" notation from one-way fixed effects: $\tilde{y}_{\mathbf{i}}=y_{\mathbf{i}} - \overline{y}_{\mathbf{j}(i)}$, $\tilde{x}=x_{\mathbf{i}} - \overline{x}_{\mathbf{j}(i)}$, etc.

Consider three regressions:

Pooled: $y_{\mathbf{i}} = \beta_0 + \beta_{\mathbf{1}}\, d_{\mathbf{i}} + \beta_{\mathbf{2}}\, x_{\mathbf{i}} + \varepsilon_{\mathbf{i}}$

Within: $\tilde{y}_{\mathbf{i}} = \beta_0 + \beta_{\mathbf{1}}\, \tilde{d}_{\mathbf{i}} + \beta_{\mathbf{2}}\, \tilde{x}_{\mathbf{i}} + \tilde{\varepsilon}_{\mathbf{i}}$

Between: $\overline{y}_{\mathbf{i}} = \beta_0 + \beta_{\mathbf{1}}\, \overline{d}_{\mathbf{i}} + \beta_{\mathbf{2}}\, \overline{x}_{\mathbf{i}} + \overline{\varepsilon}_{\mathbf{i}}$

Key points:

- We showed in Week 2 that LSDV/fixed effects is equivalent to the **within** regression.
- If there are *group-specific unobservable confounders,*
    - **within/FE** is attractive, because these drop out in deviation from means
    - **between** and **pooled** will give wrong answer, because treated and control units not comparable across groups
- If there **aren't any** group-specific unobservable confounders, the **within** regression is throwing out useful information (i.e. the information in the **between** regression)

Random effects: intuition

Pooled: $y_{\mathbf{i}} = \beta_0 + \beta_{\mathbf{1}} d_{\mathbf{i}} + \beta_{\mathbf{2}} x_{\mathbf{i}} + \varepsilon_{\mathbf{i}}$

Within: $\tilde{y}_{\mathbf{i}} = \beta_0 + \beta_{\mathbf{1}} \tilde{d}_{\mathbf{i}} + \beta_{\mathbf{2}} \tilde{x}_{\mathbf{i}} + \tilde{\varepsilon}_{\mathbf{i}}$

Between: $\overline{y}_{\mathbf{i}} = \beta_0 + \beta_{\mathbf{1}} \overline{d}_{\mathbf{i}} + \beta_{\mathbf{2}} \overline{x}_{\mathbf{i}} + \overline{\varepsilon}_{\mathbf{i}}$

Suppose there were **no** group-specific unobservable confounders.

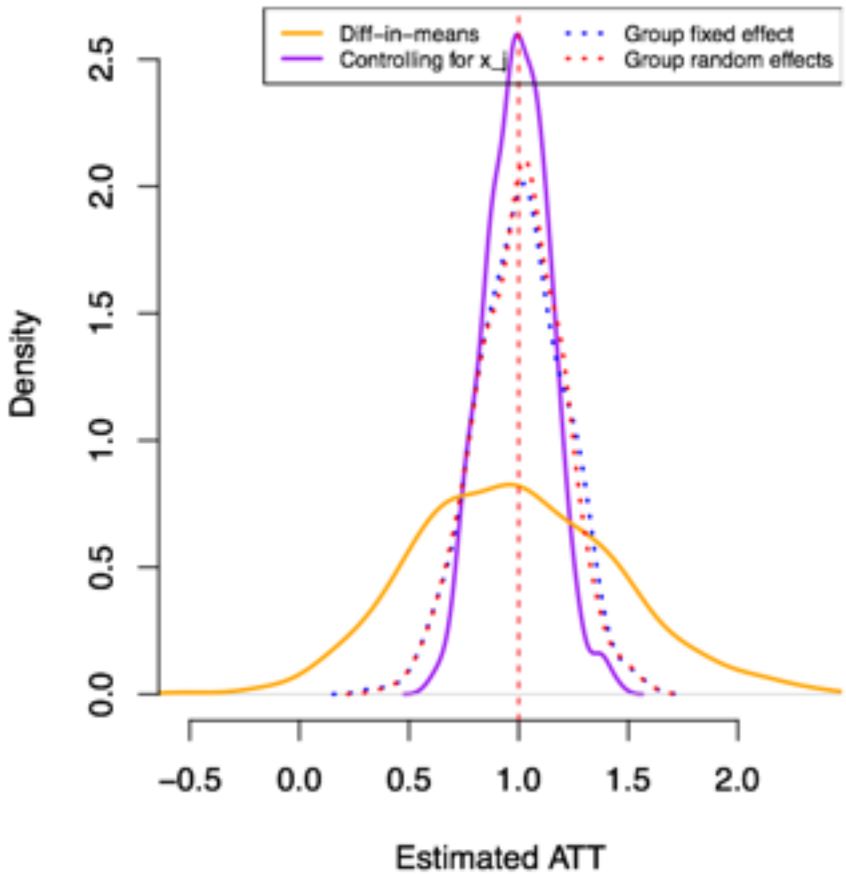
All models would give you more or less "right" answer, though weighted differently across groups (which matters if effect size varies across groups, group sizes differ).

Random effects model can be seen as attempt to efficiently pool information from **within** and **between** regression**.** In fact, estimate will be a weighted average of the estimates given by **within** and **between**.

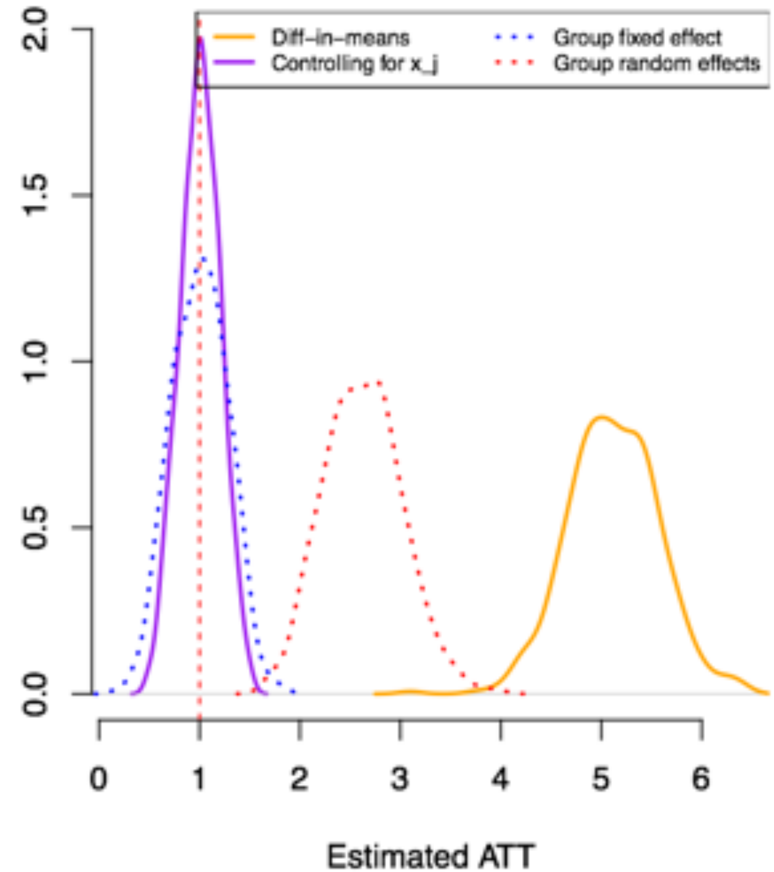
But: the **whole motivation** for panel data in causal inference is group-specific unobservable confounders. In presence of such confounders, **between** (but not **within**) is biased!

Random effects applied to week 2 simulations

Random assignment of treatment: FE and RE both unbiased; RE slightly more efficient



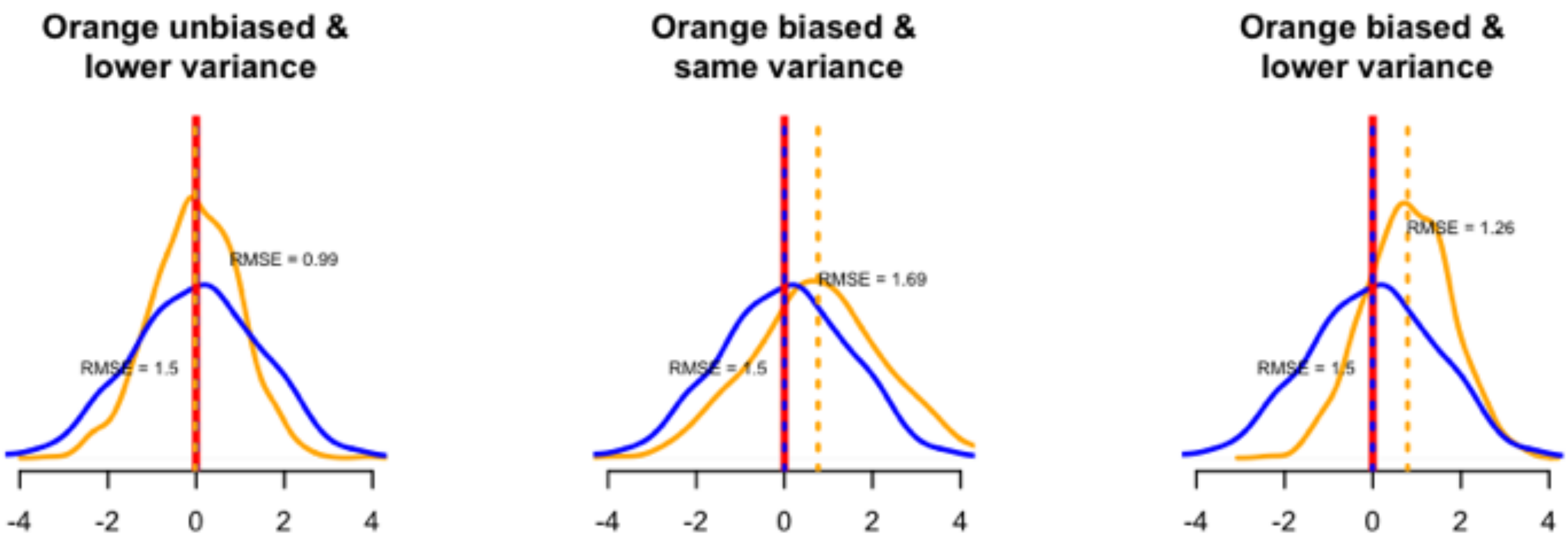
Assignment of treatment based on group-specific covariate: FE unbiased; RE estimates partway between Diff-in-means and truth



Clark and Linzer (2014), "Should I use fixed or random effects?"

RE vs FE as example of **bias-variance tradeoff:** If there **are** unit-specific unobserved confounding variables, RE is biased. But RE is also more efficient. So the average RE estimate may be closer to the truth!

$$\text{Root Mean-Squared Error (RMSE)} = \sqrt{E[(\text{estimate - truth})^2]}$$



Clark and Linzer (2014) continued

Typical recommendation is to use Hausman Test:

- check whether coefficients differ between RE and FE
- if they differ, use FE, because FE is unbiased

Clark & Linzer point out that from RMSE perspective, RE may still be better even if biased.

In Clark & Linzer's simulations,

- FE dominates RE (on a RMSE basis) when treatment is not "sluggish" (i.e. it varies substantially within groups)
- RE is better when treatment is sluggish and not so correlated with group effects

Still: remember that purpose of FE was to **rule out** group-specific confounding variables.

Multilevel/hierarchical modeling (briefly)

“Explaining fixed effects”

Recall LSDV version of fixed effects: a separate intercept for each group in the dataset.

Consider doing a second regression:

- Unit of analysis: groups from first regression
- DV: the intercept estimates (i.e. the fixed effects)
- Independent variables: (fixed, i.e. time-invariant) group characteristics.

For example:

- Levitt on repeat challengers: explain Dem vote share across pairs of candidates as function of education, age, gender, IQ, height, etc of candidates, state/region where they run, etc.
- Ansell on house prices: explain support for redistribution across cities as function of region, age of city, industrial composition, ethnic composition, etc.

"Explaining fixed effects"

More formally:

Regression 1 (level 1): $y_i = \alpha_{j(i)} + \tau x_i + \varepsilon_i$

Regression 2 (level 2): $\alpha_j = \beta_0 + \beta_1 v_j + \beta_2 z_j + \varepsilon_j$

Think of multilevel/hierarchical models as estimating these two regressions in a single model.

Key points:

- Regression 1 (fixed effects) is enough if we think of the variation across groups as a nuisance in our attempts to estimate τ
- Unlike Regression 1, Regression 2 tells us something about the "effect" of attributes that are fixed within groups
- If there are unobserved group-level confounders (motivation for fixed effects regression), Regression 2 (and possibly Regression 1) will be biased

Further reading

- Bell, Andrew and Kelvyn Jones (2015), "Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data" *Political Science Research and Methods*. **Key points:** context should be modeled, not controlled for; we can allow "between" and "within" effects to differ, though few do.
- Plumper, Thomas and Vera Troeger (2007), "Efficient Estimation of Time-Invariant and Rarely Changing Variables in Finite Sample Panel Analyses with Unit Fixed Effects" *Political Analysis*.
- Gelman and Hill (2006) *Data Analysis Using Regression and Multilevel/Hierarchical Models.*