Augmented Panel Data Models with Staggered Adoption

Eli Ben-Michael, Avi Feller, and Jesse Rothstein

Berkeley UNIVERSITY OF CALIFORNIA

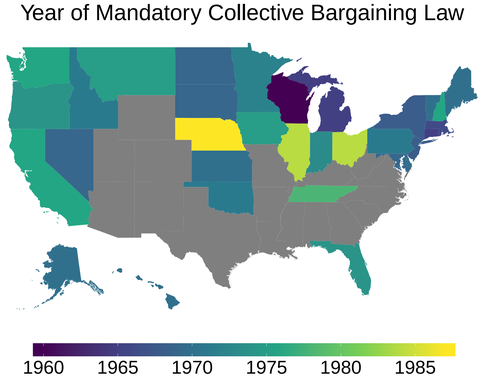
Summary

- Units often adopt policies over time
- Differences in Differences (DiD) is a workhorse, but parallel trends isn't always reasonable
- Synthetic Control Method (SCM) is more flexible, but designed for a single treatment time
- We extend SCM to staggered adoption setting, partially pool across times, and combine with DiD

Mandatory Collective Bargaining

Do public sector unions influence public spending?
Data: Mandatory collective bargaining laws [1]

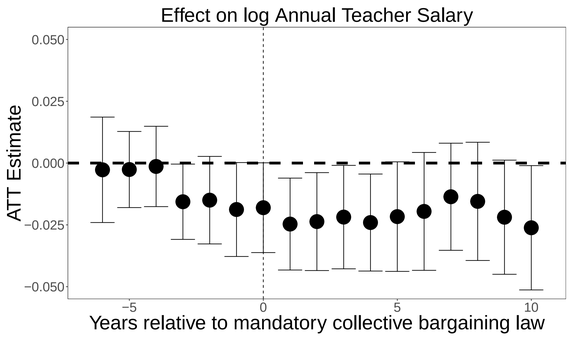
- Study period: 1959 to 1997
- 33 states mandate bargaining with public unions
- 10 states allow but do not require, 7 prohibit
- Evaluate effect on average teacher salaries



Poor pre-treatment fit

⇓

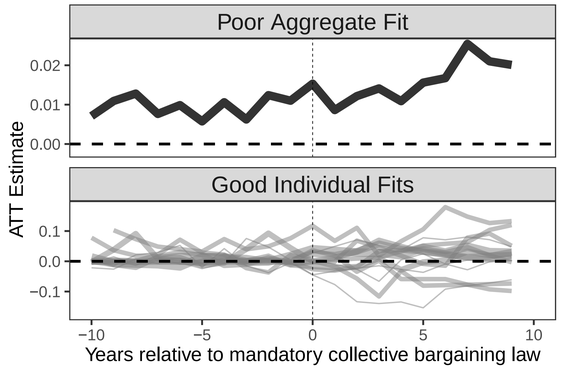
Non-parallel trends!



No parallel trends? Don't give up!

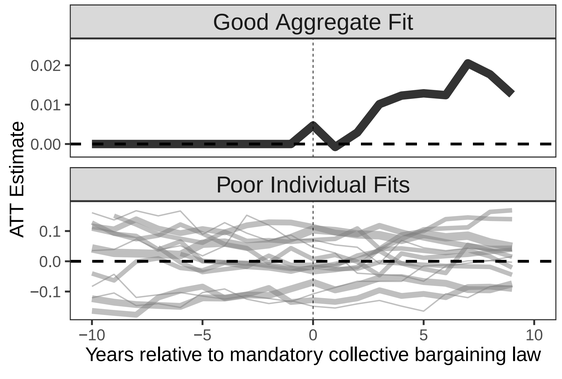
Separate SCM

Fit many separate synthetic controls



Aggregated SCM

Fit a single aggregated synthetic control

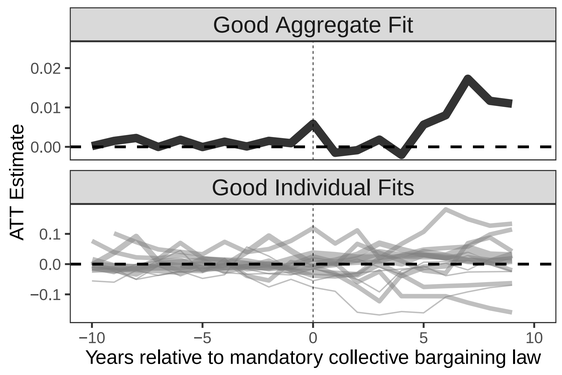


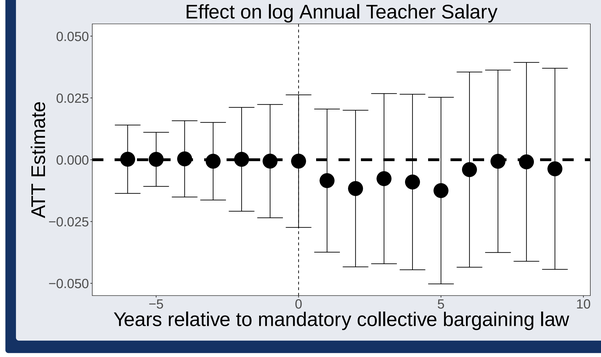
Partially Pooled SCM

Solution: Combine objectives

$$\text{Aggregate Imbalance} + \sum_j \text{Individual Imbalance}_j$$

- Move continuously between both solutions
- Trades off individual fit and aggregate fit
- Dual view as partial pooling across adoption times





Combining Weighting and DiD

Augment w/pre-period average: weighted DiD [2]

$$\frac{1}{T_j}\sum_{t'=1}^{T_j}\left[\frac{1}{N_{1j}}\sum_{W_i=j}(Y_{it}-Y_{it'})-\sum_{W_i=0}\hat{\gamma}_{ij}(Y_{it}-Y_{it'})\right]$$

- Adjust for non-parallel trends
- Balance *residuals* after unit/time fixed effects
- Add fixed effects and weighted residuals
- Key assumption: conditional parallel trends [3, 4]

Interpretation: Generalized Propensity Score

- Lagrangian dual ⇒ implicit p-score estimate
- Selection on lagged outcomes

$$Y_{i(T_j-K):T_j} = (Y_{i(T_j-K)}, \ldots, Y_{iT_j})$$

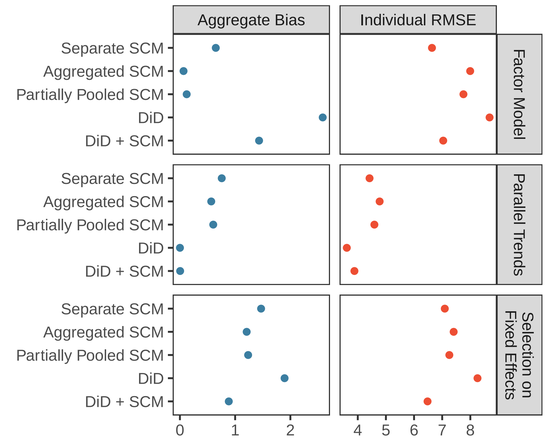
- Separate parameters β_j for treatment time T_j

$$\log\frac{P(W_i=j \mid Y_{i(T_j-K):T_j})}{P(W_i=0 \mid Y_{i(T_j-K):T_j})} = \alpha_j + \sum_{\ell=1}^{K}\beta_{\ell j}Y_{i(T_j-\ell)}$$

- Hierarchical propensity score model
- Shrinkage to a global model w/params μ_β

$$\beta_{j\ell} \sim N(\mu_{\beta\ell}, \sigma_\beta^2)$$
$$\mu_{\beta\ell} \sim N(0, \sigma_\mu^2).$$

Data Calibrated Simulation Study



References

[1] A. S. Paglayan, "Public-sector unions and the size of government," *American Journal of Political Science*, vol. 63, no. 1, pp. 21–36, 2019.

[2] E. Ben-Michael, A. Feller, and J. Rothstein, "The Augmented Synthetic Control Method." 2018.

[3] A. Abadie, "Semiparametric difference-in-differences estimators," *The Review of Economic Studies*, vol. 72, no. 1, pp. 1–19, 2005.

[4] C. Hazlett and Y. Xu, "Trajectory balancing: A general reweighting approach to causal inference with time-series cross-sectional data," 2018.

[5] A. Abadie, A. Diamond, and J. Hainmueller, "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," *Journal of the American Statistical Association*, vol. 105, no. 490, pp. 493–505, 2010.