Panel Data Analysis III

Ye Wang
University of North Carolina at Chapel Hill

Linear Methods in Causal Inference
POLI784

Review

- The previous class started from introducing the DID estimator.
- It is motivated by the TWFE model when the data have a DID structure.
- But the estimator only needs the assumption of parallel trends and is robust to heterogeneous treatment effects.
- This is not true when the data have a more complex structure.
- The within estimator may generate negative weights for individualistic treatment effects.
- We can fix the problem by not using untreated observations when fitting the model.
- This leads to the idea of counterfactual estimation.

The classic factor model

- Remember that in the TWFE model, we assume that $h_t(\mathbf{U}_i) = \mu + \alpha_i + \xi_t$.
- Bai (2003) relaxes this assumption to $h_t(\mathbf{U}_i) = \mathbf{f}_t'\lambda_i$.
- Both $\mathbf{f}_t$ and λ_i are r-dimensional vectors.
- The number of parameters is $(N + T) * r \leq N * T$.
- The former is known as factors, and the latter as factor loadings.
- When $\mathbf{f}_t = (1, \xi_t)'$ and $\lambda_i = (\alpha_i, 1)'$, it boils down to the TWFE model.
- Such a model captures the interaction between variables that vary only over time and only across units.
- E.g., country i's endowments in r different resources and the prices of them in year t.

The classic factor model

- Let's omit the covariates, then

$$Y_{it} = \mathbf{f}_t' \lambda_i + \varepsilon_{it},$$

- We can write the model in matrices:

$$\mathbf{Y} = \mathbf{F}\Lambda' + \varepsilon.$$

  where $\mathbf{Y}$ is a $T \times N$ matrix; $\mathbf{F}$ $(T \times r)$ represents factors and Λ $(N \times r)$ represents factor loadings.
- Bai (2003) shows that $\mathbf{F}$ can be found via eigenvalue decomposition of the matrix $\mathbf{YY}'/T$.
- We find a matrix $\hat{\mathbf{F}}$ such that

$$\left(\frac{1}{NT} \sum_{i=1}^{N} \mathbf{Y}_i \mathbf{Y}_i' \right) \hat{\mathbf{F}} = \hat{\mathbf{F}} \mathbf{V}_{NT},$$

  where $\mathbf{V}_{NT}$ is the matrix of eigenvalues.
- Then, we can estimate Λ via OLS, treating each $\hat{\mathbf{f}}_t$ as a known value.

The classic factor model

- Bai (2009) considers the DGP with covariates:

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{F}\Lambda' + \varepsilon.$$

- Knowing β, we can calculate residuals and estimate $\mathbf{F}$ and Λ as before; knowing $\mathbf{F}$ and Λ, we can estimate β via

$$\hat{\beta} = \left(\sum_{i=1}^{N} \mathbf{X}_i' \mathbf{Q}_{\hat{\mathbf{F}}} \mathbf{X}_i \right)^{-1} \left(\sum_{i=1}^{N} \mathbf{X}_i' \mathbf{Q}_{\hat{\mathbf{F}}} \mathbf{Y}_i \right)$$

  where $\mathbf{Q}_{\hat{\mathbf{F}}}$ is the residual-making matrix of $\hat{\mathbf{F}}$.
- We can start from some initial values and iterate this process until convergence.
- Bai shows that $\sqrt{NT}(\hat{\beta} - \beta)$ converges to a normal distribution under regularity conditions.

Synthetic control

- Abadie, Diamond, and Hainmueller (2010) propose the synthetic control (SC) method for case studies.
- It assumes that there are N untreated units and a single treated unit ($i = 1$).
- There are T_0 pre-treatment periods and T_1 post-treatment periods.
- The untreated potential outcome obeys the classic factor model as in Bai (2003) or Bai (2009).
- We are interested in is the treatment effect on unit 1 in any period $t > T_0$:

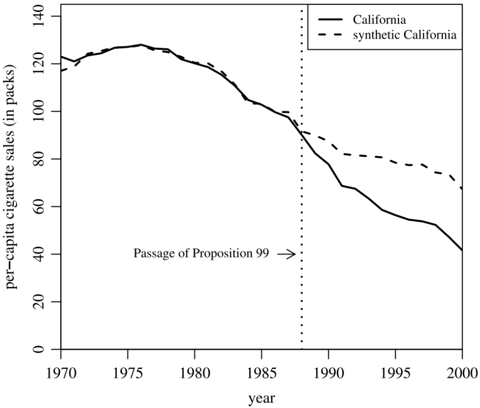
$$\tau_{1t} = Y_{1t}(1) - Y_{1t}(0).$$

Synthetic control

- Abadie, Diamond, and Hainmueller (2010) show that we don't need to estimate factors or factor loadings directly in this case.
- We instead attempt to find a group of weights $\{w_i\}_{i=2}^{N+1}$ such that
$$\sum_{i=2}^{N+1} w_i = 1, \sum_{i=2}^{N+1} w_i Y_{it} = Y_{1t}$$
for any $1 \leq t \leq T_0$.
- We require the treated unit to lie in the "convex hull" of the untreated ones.
- The predicted counterfactual for unit 1 in period $t > T_0$ is
$$\hat{Y}_{1t}(0) = \sum_{i=2}^{N+1} w_i Y_{it}.$$
- Hence, we can estimate τ_{1t} by $\hat{\tau}_{1t} = Y_{1t}(1) - \hat{Y}_{1t}(0)$.

Synthetic control

- Ideally, we would like to see a picture like this:



Synthetic control

- Balancing pre-treatment outcomes can effectively remove interactive fixed effects hence $\hat{\tau}_{1t}$ is consistent for τ_{1t}.
- But $\hat{\tau}_{1t}$ has no asymptotic distribution with only one treated observation.
- We rely on a permutation test for statistical inference.
- It is similar to a placebo test: We replace the treated unit with a randomly selected unit from the control group and estimate the effect.
- This is not Fisher's randomization test since we do not know the treatment assignment algorithm.
- It assumes that each unit has the same probability of being treated.
- There are approaches to construct non-asymptotic confidence intervals but they tend to be wide.

Generalized synthetic control

- In practice, most data include more than one treated unit.
- In theory, we can find a unique set of weights for each treated unit.
- But it could be more straightforward to estimate factors and factor loadings.
- Xu (2017) considers the following model:

$$Y_{it}(0) = \mathbf{f}_t'\lambda_i + \mathbf{X}_{it}'\beta + \varepsilon_{it},$$
$$Y_{it}(1) = Y_{it}(0) + \tau_{it}.$$

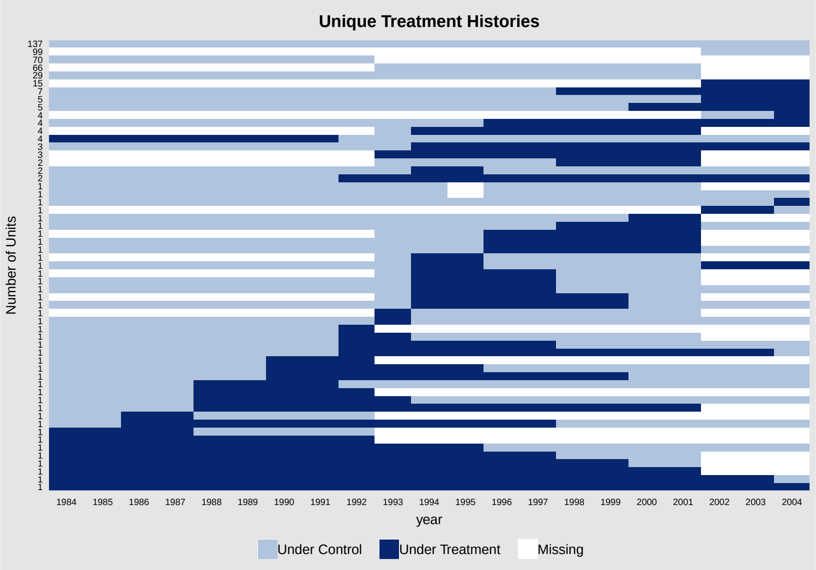
- As in counterfactual estimation, we estimate $\mathbf{f}_t$, λ_i, and β using untreated observations.
- Then, we generate predictions of the counterfactual and estimate the ATT with:

$$\widehat{\tau}_{ATT} = \frac{1}{|\mathcal{M}|} \sum_{(i,t)\in\mathcal{M}} \left(Y_{it} - \hat{Y}_{it}(0)\right).$$

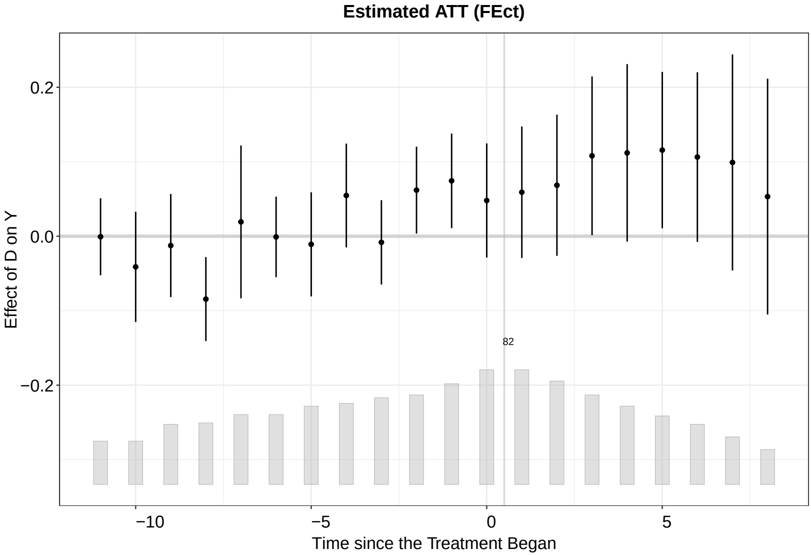
Factor models: application

- Let's examine the study in Anzia and Berry (2011).
- The argument is that female legislators face more structural barriers than their male counterparts.
- Once elected, they will have better performance.
- The data include 700 districts in the US over 21 years.
- $D_{it} = 1$ if there is a female legislator from district i in year t.
- The outcome is measured by the amount of federal spending a legislator can secure for their district in year t.

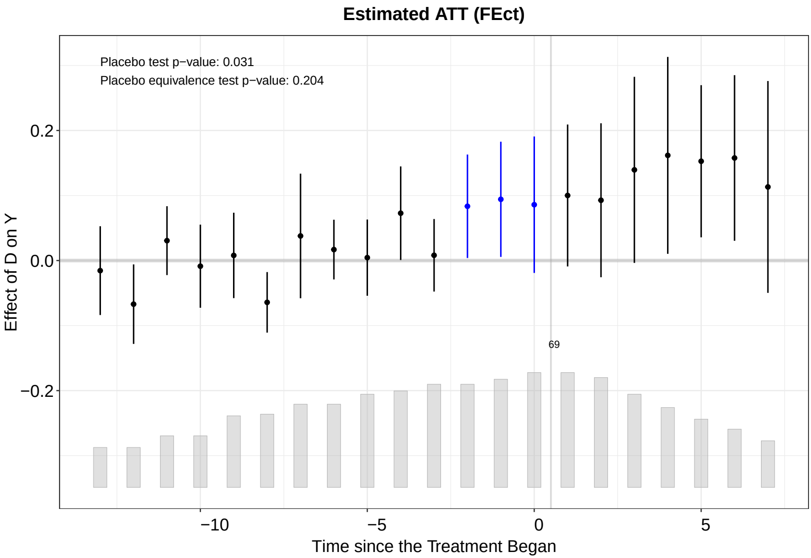
Factor models: application



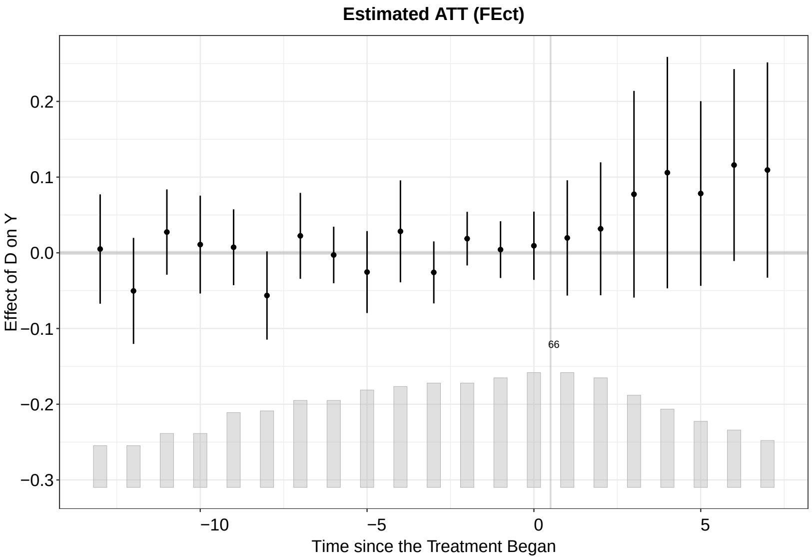
Factor models: application



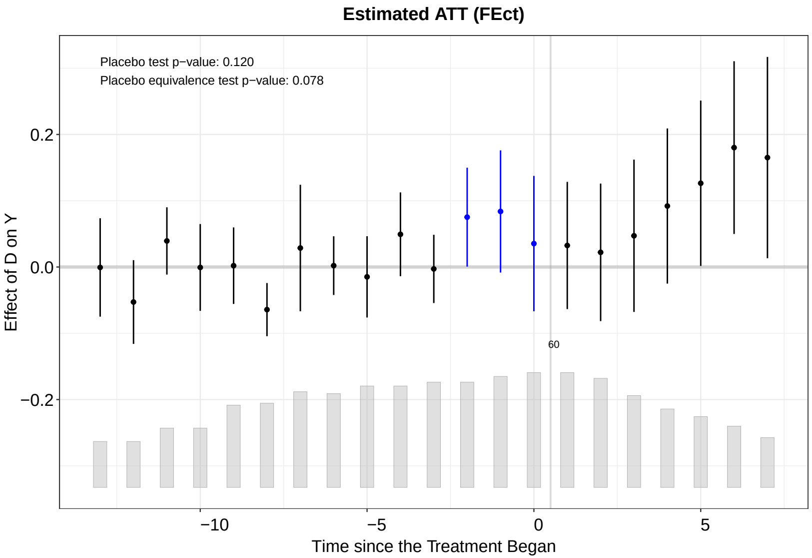
Factor models: application



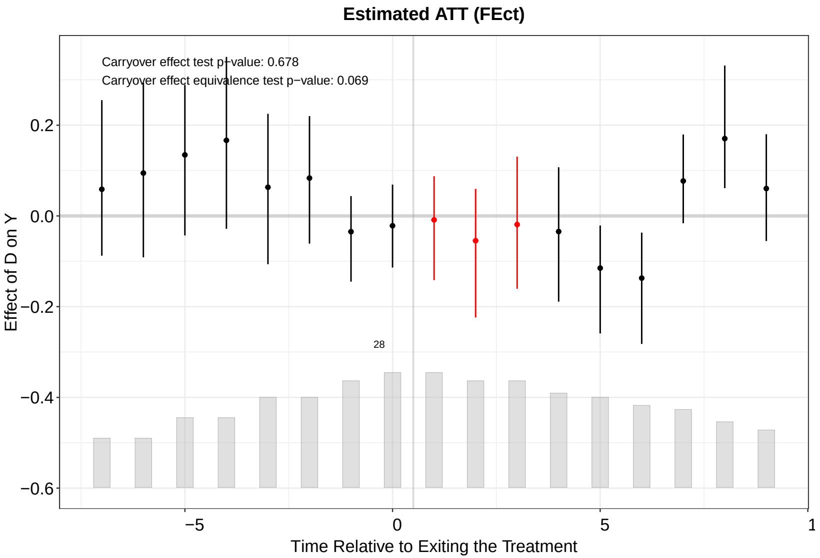
Factor models: application



Factor models: application



Factor models: application



References I

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association* 105 (490): 493–505.

Anzia, Sarah F, and Christopher R Berry. 2011. "The Jackie (and Jill) Robinson Effect: Why Do Congresswomen Outperform Congressmen?" *American Journal of Political Science* 55 (3): 478–93.

Bai, Jushan. 2003. "Inferential Theory for Factor Models of Large Dimensions." *Econometrica* 71 (1): 135–71.

———. 2009. "Panel Data Models with Interactive Fixed Effects." *Econometrica* 77 (4): 1229–79.

Xu, Yiqing. 2017. "Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models." *Political Analysis* 25 (1): 57–76.