Topic 6: Fixed Effects, Difference-in-differences, and Factor Models

ECON 5783 — University of Arkansas

Prof. Kyle Butts

December 2024

Imputation Estimator review

The last set of slides, we introduced an "imputation estimator" for panel data treatment effects:

1. Estimate model for $y_{it}(\infty)$ using observations with $d_{it} = 0$ and get fitted values for full sample, $\hat{y}_{it}(\infty)$
2. Regress $y_{it} - \hat{y}_{it}(\infty)$ on d_{it} or event-study indicatorsto estimate treatment effects
   - $\rightarrow$ Estimating the overall effect, ATT, or dynamic effects of being treated for ℓ periods, ATT^{ℓ} respectively

Imputation Estimator review

The last set of slides, we introduced an "imputation estimator" for panel data treatment effects:

1. Estimate model for $y_{it}(\infty)$ using observations with $d_{it} = 0$ and get fitted values for full sample, $\hat{y}_{it}(\infty)$
2. Regress $y_{it} - \hat{y}_{it}(\infty)$ on d_{it} or event-study indicatorsto estimate treatment effects
   - $\rightarrow$ Estimating the overall effect, ATT, or dynamic effects of being treated for ℓ periods, ATT^{ℓ} respectively

This topic will extend this procedure to "factor models" that will allow more general trending behavior

Factor Models

Factor Model

Untreated potential outcomes are given by a factor model:

$$y_{it}(0) = \sum_{r=1}^{p} f_{t,r} * \gamma_{i,r} + u_{it}$$

- $f_{t,r}$ is the r-th **factor** (macroeconomic shock) at time t.
- $\gamma_{i,r}$ is unit i's **factor loading** (exposure) to the r-th factor.

Factor Model Example

$$y_{it}(0) = \sum_{r=1}^{p} f_{t,r} * \gamma_{i,r} + u_{it}$$

If we are thinking about housing prices, y_{it}:

- γ_i are characteristics of neighborhood / house
- f_t are demand shocks in each period

Factor Model Example

$$y_{it}(0) = \sum_{r=1}^{p} f_{t,r} * \gamma_{i,r} + u_{it}$$

If we are thinking about wages, y_{it}:

- γ_i are worker's latent skills (e.g. computer skills)
- f_t reflect changing firm's sdemand for skills

Factor Model Example

$$y_{it}(0) = \sum_{r=1}^{p} f_{t,r} * \gamma_{i,r} + u_{it}$$

If we are thinking about county employment, y_{it}:

- γ_i are characteristics of a county (e.g. their manufacturing share)
- f_t reflect national shocks to the economy (e.g. the “China shock”)

Two-way Fixed Effect vs. Factor Model

The factor model is a generalization of the TWFE model. If $\boldsymbol{f}_t = (\lambda_t, 1)'$ and $\boldsymbol{\gamma}_i = (1, \mu_i)'$, then our factor model becomes the TWFE model:

$$y_{it}(0) = \boldsymbol{f}_t' \boldsymbol{\gamma}_i + u_{it} = \lambda_t + \mu_i + u_{it}$$

We can add unit and/or time fixed-effects as 'known' factors if we want

Factor Model and Parallel Trends

Say you have a single treatment timing and two periods. Let D_i be out treated group indicator. Then

$$\begin{aligned} \mathbb{E}[\Delta y_i \mid D_i = d] &= \mathbb{E}[y_{i1} - y_{i0} \mid D_i = d] \\ &= \Delta \boldsymbol{f} \, \mathbb{E}[\boldsymbol{\gamma}_i \mid D_i = d] \end{aligned}$$

Under a factor model, the average change in y_{it} for group $D_i = d$ is the change in factor shocks $\boldsymbol{f}$ times the average exposure to those shocks

Factor Model and Parallel Trends

$$\mathbb{E}[\Delta y_i \mid D_i = d] = \Delta \boldsymbol{f}\, \mathbb{E}[\boldsymbol{\gamma}_i \mid D_i = d]$$

Say the treated group has higher exposure to a shock than the control group

- $\implies$ the trends differ by treatment status

Factor Model and Parallel Trends

$$\mathbb{E}[\Delta y_i \mid D_i = d] = \Delta \boldsymbol{f} \, \mathbb{E}[\boldsymbol{\gamma}_i \mid D_i = d]$$

Say the treated group has higher exposure to a shock than the control group

- $\implies$ the trends differ by treatment status

That is, a factor model allows for "non-paralell trends" based on difference in exposures to shocks

Example

$$y_{it}(0) = \sum_{r=1}^{p} f_{t,r} * \gamma_{i,r} + u_{it}$$

Say we are thinking about neighborhood housing prices, y_{nt}. We are interested in some treatment D_n, e.g. access to subways.

- Say γ_n is the walkability of the neighborhood
- f_t are demand shocks for walkable neigbhorhoods

If new subways are built in more walkable neighborhoods, then we do not believe parallel trends hold in this setting