Econometrics

Difference-in-Differences

Paul P. Momtaz

The Anderson School
UCLA

Motivation

Discussion of triple diff approach in
Dissanaike Drobetz Momtaz Rocholl (2018)

Introduction

Difference - in - differences estimator:

$$\hat{\beta}_{DD} = [\overline{y}_{ta} - \overline{y}_{tb}] - [\overline{y}_{ca} - \overline{y}_{cb}]$$

Or in a regression:

$$y_i = \beta_0 + \beta_1(\text{treat}) + \beta_2(\text{after}) + \beta_{DD}(\text{treat} \times \text{after}) + u$$

Advantages: Controls for pre-existing differences between controlled and treated; compares changes in treatment group to those in control group.

Parallel Trend Assumption

Parallel Trend Assumption: Heteroskedasticity adjusted s.e. necessary unless homoskedastic small sample cluster s.e. use additional random effects $\varepsilon_{ig} = V_g + \eta_{ig}$, then correlation $\rho_e = \frac{\sigma_v^2}{\sigma_v^2 + \sigma_\eta^2}$

Moulton Factor

Moulton Factor:

$$\sqrt{\frac{v(\hat{\beta}_1)}{V_c(\hat{\beta}_1)}} = \sqrt{1 + (n-1)\rho_c}$$

where

$$V(\hat{\beta}_1) \text{ correct sampling variance}$$
$$V_c(\hat{\beta}_1) \text{ conventional OLS sampling variance}$$

- Tells us by how much overestimate precision ignoring intra-clam correlation
- Conventional s.e. become increasingly misleading as ρ_0 and η increase
- Formula above is special case regressor fixed within groups and group sizes fixed.

Moulton Factor

General Case

$$\frac{V(\hat{\beta}_1)}{V_c(\hat{\beta}_1)} = 1 + \left[\frac{V(n_g)}{\overline{n}} + \overline{n} - 1\right] \rho_x \rho_\varepsilon$$

$$\rho_x = \frac{\sum_g \sum_j \sum_{i \neq j} (x_{ig} - \overline{x})(x_{jg} - \overline{x})}{V(x_{ig}) \sum_g n_g (n_g - 1)}$$

$\Rightarrow$ Worry most when regressors fixed within groups

Practical Solutions

- Adjust OLS s.e.
- Cluster s.e.
- Use group averages instead of micro data
- Block bootstrap

Diff-in-diff and Panel Data

Serial correlation in panels and difference - in - differences: $\hat{\beta}_{DD}$ inconsistent

- Violation of PTA! Need multiple periods and/or groups.
- Pass clustering back one level light, but reduces the number of clusters.