Ec331: Research in Applied Economics

Spring term, 2020

Panel Data: brief outlines

Remaining structure

Work in Progress Presentations (10%)

Thursday 4pm in S2.84, 15 mins, 8 slides maximum

Wk.5 Chioma, Tom, Michael, Koby (extra hour 5-6: S1.66)

Wk.6 Oliver, Ayomide, Harris, Verona (extra hour 5-6: S2.84)

Wk.7 Alan, Julia, Kam, Will

Resources

Stata search function and pdf documentation

Ec226 Lecture notes

Ec331 Website: Topic handouts

Wooldridge (2009) “Introductory Econometrics: a Modern Approach”
Verbeek (2012) “A Guide to Modern Econometrics”
Cameron and Trivedi (2009) “Microeconometrics Using Stata”
Baum (2006) “An Introduction to Modern Econometrics Using Stata”

Specialist: e.g. Greene and Hensher (2010) “Modeling Ordered Choices”, Ch. 9

Panel Data

Today:

Parallel Worlds: Fixed effects, differences-in-differences, and Panel Data

- Angrist and Pischke Ch. 5

Key to causal inference: control for observed confounding factors.

If important confounders might be unobserved, could try IV, but good instruments hard to find.

Here: discussion of strategies that use data with a time or cohort dimension to control for unobserved but fixed omitted variables by relying on comparisons in levels (while requiring the counterfactual *trend* behaviour of treatment and control groups to be the same)

(1) Individual fixed effects

Y_{it} - outcome i at time t, observed as Y_{0it}, Y_{1it}

D_{it} - treatment

& suppose : $E[Y_{0it} \mid \mathbf{A}_i, \mathbf{X}_{it}, t, D_{it}] = E[Y_{0it} \mid \mathbf{A}_i, \mathbf{X}_{it}, t]$

where $\mathbf{A}_i$ - vector of fixed, unobserved confounders

$\mathbf{X}_{it}$ - vector of observed time-varying covariates

so D_{it} as good as randomly assigned conditional on $\mathbf{A}_i$ & $\mathbf{X}_{it}$

Individual fixed effects

Key to fixed effects estimation:

(i) Assumption unobserved $\mathbf{A}_i$ appears without a time subscript in a linear model

(ii) Assumption that the causal effect is additive and constant

$$\text{(i) } E[Y_{0it} \mid \mathbf{A}_i, \mathbf{X}_{it}, t] = \alpha + \lambda_t + \mathbf{A}_i'\boldsymbol{\gamma} + \mathbf{X}_{it}'\boldsymbol{\beta}$$

$$\text{(ii) } E[Y_{1it} \mid \mathbf{A}_i, \mathbf{X}_{it}, t] = E[Y_{0it} \mid \mathbf{A}_i, \mathbf{X}_{it}, t] + \rho$$

$$\Rightarrow E[Y_{it} \mid \mathbf{A}_i, \mathbf{X}_{it}, t, D_{it}] = \alpha + \lambda_t + \rho D_{it} + \mathbf{A}_i'\boldsymbol{\gamma} + \mathbf{X}_{it}'\boldsymbol{\beta}$$

More restrictive than needed to motivate regression previously; linear, additive functional form necessary to advance consideration of the problem of unobserved confounders using panel data with no instruments.

Individual fixed effects

This implies the fixed effects model:

$$Y_{it} = \alpha_i + \lambda_t + \rho D_{it} + \mathbf{X}_{it}'\boldsymbol{\beta} + \varepsilon_{it}$$

$$\text{where } \varepsilon_{it} \equiv Y_{0it} - E[Y_{0it} \mid \mathbf{A}_i, \mathbf{X}_{it}, t]$$

$$\text{and } \alpha_i \equiv \alpha + \mathbf{A}_i'\boldsymbol{\gamma}$$

Given panel data, the causal effect can be estimated by also considering the fixed effect and the year effect, respectively α_i, λ_t, as parameters to be estimated.

The unobserved time-invariant individual effects are collectively captured as coefficients on dummies for each individual, the year effects as coefficients on time dummies.

Aside: Random Effects

Alternative to the fixed effects specification: random effects model.

Assumes α_i (the fixed effect) is uncorrelated with the regressors; and so omitting this variable does not induce bias, and it effectively becomes part of the residual.

Most important consequence: residuals for a given person are correlated across periods (with implications for OLS standard errors).

Fixed effect equivalent: deviations from means

If FE seems like a lot of parameters to estimate, not a problem, in practise; individual effects as parameters algebraically equivalent to estimation in deviations from means:

$$\bar{Y}_i = \alpha_i + \bar{\lambda} + \rho \bar{D}_i + \bar{\mathbf{X}}_i' \boldsymbol{\beta} + \bar{\varepsilon}_i \qquad \bar{Y}_i = \frac{1}{T} \sum_{t=1}^{T} Y_{it}$$

& subtracting gives:

$$Y_{it} - \bar{Y} = \lambda_t - \bar{\lambda} + \rho (D_{it} - \bar{D}_i) + (\mathbf{X}_{it} - \bar{\mathbf{X}}_i)' \boldsymbol{\beta} + (\varepsilon_{it} - \bar{\varepsilon}_i)$$

The deviations from means estimator is also known as the "within estimator" and "analysis of covariance", while such estimation is called *absorbing* the fixed effects.

Fixed effect equivalent: deviations from means

Equivalent as, by the regression anatomy formula (see wk. 3 slides), any set of multivariate regression coefficients can be estimated by first regressing the desired set of independent variables on all other included variables, and then regressing the original dependent variable on the resultant residuals. Here, the residuals from a regression on a full set of person-dummies in a person-year panel are deviations from person means.

Alternative: Differencing

One alternative to deviations from means is differencing:

$$\Delta Y_{it} = \Delta\lambda_t + \rho\Delta D_{it} + \Delta\mathbf{X}_{it}'\boldsymbol{\beta} + \Delta\varepsilon_{it}$$

where Δ denotes change from one year to the next

Algebraically the same as deviations from means with two periods, not otherwise.

If more than two periods, homoskedastic and serially uncorrelated errors ensure deviations from means is more efficient.

Susceptibility to attenuation bias

Note: FE estimates are notoriously susceptible to attenuation bias from measurement error.

- economic variables tend to be persistent
- measurement error often changes from year to year

so any observed changes may be mostly noise

i.e. more measurement error in differenced regressors than in levels, which may result in smaller fixed effects estimates.

So if we observe smaller fixed effects estimates than cross section estimates, selection bias in the latter is not the only possible explanation.

Removal of useful information

Variant of measurement error problem:

Differencing and deviations from means estimators typically remove both good and bad variation, the transformations both dealing with some of the potential omitted variables bias but also removing useful information. As such, even small unobserved time-variation in the individual effects could be responsible for substantial bias.

IV methods or external information and adjustment could be used to confront the measurement error problem, but such data rarely available. At a consequence, its important to avoid overly strong claims when interpreting fixed effects estimates.

(2) Differences-in-differences: Pre & Post, Treatment & Control

Fixed effects estimation requires panel data (repeated observations at the individual level), but often the regressor of interest varies only at a more aggregate or group level.

The source of OVB in policy evaluation is then unobserved variables at the group and year level, which may be captured by group-level fixed effects, leading to the differences-in-differences identification strategy (a version of fixed effects estimation using aggregate data).

Differences-in-differences

Heart of DD setup - additive structure for potential outcomes in the no-treatment state, such that the outcome is the sum of a time-invariant group effect and a year effect common across groups (with the former playing the role previously of unobserved individual effect):

$$E[Y_{0ist} \mid s,t] = \gamma_s + \lambda_t$$

where s denotes group

if D_{st} is a treatment dummy,

with assumption: $E[Y_{1ist} - Y_{0ist} \mid s,t] = \delta$

observed outcome:

$$Y_{ist} = \gamma_s + \lambda_t + \delta D_{st} + \varepsilon_{ist}$$

Differences-in-differences

Gives:

$$E[Y_{ist} \mid s = \text{control}, t = \text{post-treatment}] - E[Y_{ist} \mid s = \text{control}, t = \text{pre-treatment}] = \lambda_{post} - \lambda_{pre}$$

$$E[Y_{ist} \mid s = \text{treated}, t = \text{post-treatment}] - E[Y_{ist} \mid s = \text{treated}, t = \text{pre-treatment}] = \lambda_{post} - \lambda_{pre} + \delta$$

Differences-in-differences

So population difference-in-differences:

$$
\begin{aligned}
&\{E[Y_{ist} \mid s = \text{treated}, t = \text{post-treatment}] \\
&\quad - E[Y_{ist} \mid s = \text{treated}, t = \text{pre-treatment}]\} \\
&-\{E[Y_{ist} \mid s = \text{control}, t = \text{post-treatment}] \\
&\quad - E[Y_{ist} \mid s = \text{control}, t = \text{pre-treatment}]\} \\
&= \delta
\end{aligned}
$$

which is the causal effect of interest, and easily estimated using the sample analog of the population means.

Key Identifying Assumption

Critical identifying assumption: *trends* would be the same in both groups in the absence of treatment, with all differences between treatment and control groups captured by the fixed effect, γ_s .

This common trends assumption can be investigated using data on multiple periods, to see if the year-to-year variation differs substantially between groups for transitions other than treatment itself. Is the control group a good measure of counterfactual outcomes for the treatment group? A graph of outcomes over time may provide suggestive visual evidence.

Regression DD

We can use regression to estimate such models:

$$Y_{ist} = \gamma_s + \lambda_t + \delta D_{st} + \varepsilon_{ist}$$

with Z_s a dummy for treatment group

d_t a time dummy for post-treatment

gives:

$$Y_{ist} = \alpha + \gamma Z_s + \lambda d_t + \delta (Z_s \cdot d_t) + \varepsilon_{ist}$$

with $D_{st} = Z_s \cdot d_t$, $\alpha = \gamma_{untreated} + \lambda_{pre-treatment}$

$\gamma = \gamma_{treated} - \gamma_{untreated}$, $\lambda = \lambda_{post} - \lambda_{pre-treatment}$

δ : causal effect of treatment

Regression DD - advantages

Additional groups or periods can also be added, with a generalisation including a dummy for each group and period.

Regression DD also facilitates the study of policies other than those described by a simple dummy variable, where treatment intensity varies across states and over time.

Finally, also easy to add additional covariates in this framework, modelling counterfactual outcomes as:

$$E[Y_{0ist} \mid s,t,\mathbf{X}_{st}] = \gamma_s + \lambda_t + \mathbf{X}_{st}'\boldsymbol{\beta}$$

Further issues

If the sample includes many years, the regression-DD model can be used to test for "Granger causality", which in this context means checking that past treatment predicts outcomes but future treatment does not. The pattern of lagged effects may be of interest, also.

An alternative check on the DD identification strategy adds group-specific time trends to the list of controls; ideally, the estimated effects of interest are unchanged by such inclusion. Note: this requires at least three periods, and estimation is more robust and convincing when pre-treatment data establishes a clear trend that can be extrapolated to the post-treatment period.

DD designs rely on treatment-control comparison; one potential pitfall is when the composition of treatment and control groups changes as a result of treatment. In this case, we might want to instrument for membership of treatment group.

Fixed Effects Vs Lagged Dependent Variables

For many causal questions, the notion that the most important omitted variables are time invariant is implausible. As such, an alternative estimation strategy may control for past outcomes and dispense with fixed effects.

Applied researchers may face a choice between fixed effects and lagged dependent variables models; one solution is to include both, but the conditions for consistent estimation are much more demanding than for either alone.

As always, it is useful to check the robustness of any findings using alternative identifying assumptions, ideally finding broadly similar results. If a lagged dependent variable is the correct basis for the conditional independence assumption necessary for causal inference, but fixed effects are mistakenly used, estimates of a positive treatment effect will tend to be too large. Similarly, if fixed effects assumptions are correct but lagged outcomes are utilised, estimates will tend to be too small. Thus fixed effects and lagged dependent variables can be thought of as bounding the causal effect of interest.

- see Angrist and Pischke, Ch.5 for more discussion and applied examples