

# Lecture 17

10-25-2021

# What did we cover last time?

Last week we discussed selection on observables

We made a note that selection on observables is often hard to believe

# What are we doing today?

Introducing Panel Data

Introducing Fixed Effects

# What is Panel Data?

Panel Data is data with at least two dimensions.

Panel Data is obtained by observing the same unit over several periods

# Panel Data

country	continent	year	lifeExp	pop	gdpPercap
Afghanistan	Asia	1952	28.801	8425333	779.4453
Afghanistan	Asia	1957	30.332	9240934	820.8530
Afghanistan	Asia	1962	31.997	10267083	853.1007
Afghanistan	Asia	1967	34.020	11537966	836.1971
Afghanistan	Asia	1972	36.088	13079460	739.9811
Afghanistan	Asia	1977	38.438	14880372	786.1134

# Panel Data

A balanced panel data set is a dataset with the same number of time periods for each cross section unit observation.

# The Omitted Variables Problem

Suppose  $\mathbf{X} \equiv (x_1, x_2, \dots, x_k)$  are observable random variables

Let  $c$  be an unobservable random variable.

We are interested in  $E[Y|\mathbf{X}, c]$

# The Omitted Variables Problem

We estimate this model with the regression:

$$E[Y|\mathbf{X}, c] = \beta_0 + \mathbf{X}\beta + c + \epsilon$$

If  $Cov(x_j, c) \neq 0$  for some variable, putting  $c$  in the error term causes serious problems with an omitted variable.



# The Omitted Variables Problem

If we have access to panel data then we have other possibilities to solve this problem.

Suppose we make an assumption that  $c$  is constant, which means it is an unobserved time-constant variable. We call this an *unobserved effect* in panel data analysis

We are now interested in a new CEF

$$E[Y_t | \mathbf{X}_t, c] = \beta_0 + \mathbf{X}_t \beta + c + \epsilon_t$$

# The Unobserved Effects Model

For a randomly drawn cross section unit  $i$ , the UEM is:

$$Y_{it} = \mathbf{X}_{it}\beta + c_i + \epsilon_{it}$$

$c_i$  is often referred to as an individual effect or individual heterogeneity. The errors are called idiosyncratic errors or idiosyncratic disturbances.

# Fixed Effects

A Fixed effects model is a type of UEM. Fixed effects are models for which we are interested in analyzing the impact of variables that vary over time.

If unobserved variables do not change over time, then any change in the dependent variable must be due to influences other than these fixed characteristics (Stock and Watson 2003)

Models with fixed effects tend to make a similar argument for the CIA with the fixed effects.

We've seen fixed effects before when we use regression to analyze a block randomization study.

# Fixed Effects in Randomized Block Designs

We've seen fixed effects before when we estimated block randomized control trials

Each block is a categorical variable for which within each block we have some randomization scheme.

# Fixed Effects of Experiment with estimatr

```
FE_Estimator <- lm_robust(Y ~ D, fixed_effects = ~block)

FE_IPW <- lm_robust(Y~D, weights = 1/D_cond_prob, fixed_effects = ~block)

FE_InvVarWeight <- lm_robust(Y~D, weights = 1/(D_cond_prob*(1-D_cond_prob), fixed_effects = ~block)
```

# Assumptions for Fixed Effects

1. For each unit the model is  $Y_{it} = X_{ij}\beta + c_i + \epsilon_{it}$  where the  $\beta_j$  are parameters to estimate and  $c_i$  is the unobserved effect.
2. We have a random sample from the cross-section
3. Full Rank. Each explanatory variable changes over time for at least some time periods and there is no perfect linear relationships among the explanatory variables
4. For each time period  $E[\epsilon_{it} | \mathbf{X}_i, c_i] = 0$

# Example 1: Eckhouse (2021)

"Metrics Management and Bureaucratic Accountability: Evidence from Policing"

Government depend on agents to enforce decisions. They may build in metrics to assess these. Eckhouse shows how metrics management encourages bureaucrats to prioritize work with rapid measurable results and deprioritize complex and uncertain projects

She focuses on police departments and the use of CompStat

# Example 1: Eckhouse (2021)

Table 2. Unfoundedness and Manipulation in Crime Data

	Founded	Unfounded
Manipulated	Crimes manipulated in ways other than unfounding: downgrading to a less serious charge, suppressing victim reports; reclassifying to a misdemeanor, reducing the value of stolen property	Crimes manipulated by categorizing them as unfounded: social justification available when false reports are believed to be common
Not manipulated	Crime recorded in crime statistics as it was reported to the police	Crime reported that truly did not occur: auto theft reports when the owner



# Example 1: Eckhouse (2021)