# Week 1 Panel Data

Nick Graetz[1]

[1] University of Pennsylvania, Population Studies Center

9/4/2020

# Today

1. Quick overview of class
2. R Basics
3. Latex: R Markdown / Overleaf
4. R Training (into office hours)

# What are we doing?

- We know what panel data are: repeated observations over a unit (for us, an individual)

# What are we doing?

- We know what panel data are: repeated observations over a unit (for us, an individual)
- What are we trying to do?

# What are we doing?

- We know what panel data are: repeated observations over a unit (for us, an individual)
- What are we trying to do?
- Predictive inference (in- or out-of-sample) vs. causal inference

# Predictive inference

- Predictive inference is used to describe or forecast

# Predictive inference

- Predictive inference is used to describe or forecast
- Perhaps we want to infer something in-sample about some underlying population parameter, but we have to deal with unit heterogeneity (i.e. correlated errors)

# Predictive inference

- Predictive inference is used to describe or forecast
- Perhaps we want to infer something in-sample about some underlying population parameter, but we have to deal with unit heterogeneity (i.e. correlated errors)
- Perhaps we want to forecast out-of-sample (or in-sample with tiny samples), and need to leverage these more complex correlation structures (space, time, etc.)

# Predictive inference

- Predictive inference is used to describe or forecast
- Perhaps we want to infer something in-sample about some underlying population parameter, but we have to deal with unit heterogeneity (i.e. correlated errors)
- Perhaps we want to forecast out-of-sample (or in-sample with tiny samples), and need to leverage these more complex correlation structures (space, time, etc.)
- In our class, we will be using panel data analysis largely to describe, test group differences, etc.

# Causal inference

- Lots of panel data techniques assume that you've randomized something.

# Causal inference

- Lots of panel data techniques assume that you've randomized something.
- In sociology and demography, people are almost always using observational data.

# Causal inference

- Anytime we start introducing independent variables to explain differences (“links”, “determinants”,“associations”) we are implicitly relying on lots of causal assumptions.

# Causal inference

- Anytime we start introducing independent variables to explain differences ("links", "determinants","associations") we are implicitly relying on lots of causal assumptions.
- For example, why are there group differences in some parameter?

# Causal inference

- Things like fixed effects "help" in this case, in the sense that you're conditioning on unobserved time-invariant confounders, but there are a lot of other considerations.

# Causal inference

- Things like fixed effects "help" in this case, in the sense that you're conditioning on unobserved time-invariant confounders, but there are a lot of other considerations.
- Matthay et al (2020): Alternative causal inference methods in population health research: Evaluating tradeoffs and triangulating evidence