**Panel Data Modeling**

**Georgetown Center for Econometric Practice**

**Michael A. Bailey**

***Course outline***

Panel data sets consist of multiple observations for a specific set of units over time. Such data is incredibly common in economic and policy analysis. Examples range from data on national income since World War 2 or state-level highway deaths from 1980 onward.

Such data presents special challenges as well as opportunities for making better judgements than is possible based on data that is merely cross-sectional (a set of units at one point in time) or time series (many observations over time for a single unit).

This course introduces students to statistical analysis of panel data. We begin by using a panel data example to illustrate how basic OLS can fail when the error term is correlated with the independent variable. We then discuss fixed effects models, demonstrating multiple ways to estimate them and highlighting some pitfalls in implementation of such models. We then apply this logic to difference-in-difference methods that are widely used in policy evaluation. We finish by investigating some of the approaches for dealing with time dependence within the data generating process.

The introductory material will be covered in the morning lectures and then complemented by computer work in the afternoon to consolidate the material covered in the lectures.

In addition to the formal lectures, the course will also introduce participants to the toolkit for applied analysis based on the software package, Stata (www.stata.com).

***Course Objectives***

Participants will understand:

* The fundamental statistical problems that panel data analysis addresses
* How to implement panel data analysis in Stata
* Limits to panel data analysis, including conditions under which panel data models can produce biased estimates.
* How panel data analysis produces substantially different results in real-world cases.

***Instructor***

Professor Michael Bailey will deliver the course. [Professor Bailey](http://faculty.georgetown.edu/baileyma/) received his PhD from Stanford University in 1996 and is the Colonel William J. Walsh Professor in the McCourt School of Public Policy and Department of Government at Georgetown University. He is the author of *Real Stats: Using Econometrics for Political Science and Public Policy* and *Real Econometrics: The Right Tools to Answer Important Questions*, both from Oxford University Press.

***Course schedule***

Day 1 – Thursday, June 9

8:45 – 9:15 Registration and coffee/breakfast

9:15 – 10:10 Introduction to course, including core problems of endogeneity, inference and indentification

10:15 – 11:10 Panel data: the problems with treating it as pooled data and the opportunities for modeling panel data structure

11:15 to 12:10 Fixed effect models, one and two way fixed effect models, their estimation and limits, random effects

1:15 to 3:10 Applications in Stata

3:15 – 4:45 Difference in difference models

Day 2 –Friday, June 10

9:15 – 10:10 Autocorrelation in time series data

10:15 – 11:10 Dynamics in time series data

11:15 to 12:10 Dealing with autocorrelation and dynamics in time series data

1:15 to 3:10 Applications in Stata

3:15 – 4:45 Case studies using panel data