

# Introduction to Panel Data Analysis

Dr Anja Neundorf  
University of Nottingham

## Short bio – Anja Neundorf

Dr. Anja Neundorf is an Associate Professor in Politics and Research Methods at the University of Nottingham. She previously held a Post-doctoral Prize Research Fellowship at Nuffield College, Oxford and received her PhD from the University of Essex. Her research interests lie at the intersection of political behavior, research methods, and comparative politics. Her research has been published in the Journal of Politics, British Journal of Political Science, Public Opinion Quarterly, and Social Forces.

## Course description

This is an introduction to longitudinal data analysis on an applied level using Stata. The focus of the course is on data management and analyzing micro panel data.

The module will begin by discussing the advantages (and limitations) of panel data, and will show how to handle and describe a panel dataset. We will then cover linear regression techniques: fixed and random effects models as well as hybrid models. Following each lecture, participants will work through practical examples in the computer lab using the Stata statistical package using e.g. the British Household Panel Survey.

The focus of the module is applied, but some math will be used to formalize theoretical concepts. Note that the module does not cover specific techniques for macro panels (e.g. data on countries over time) or panels with small numbers of cross-sectional units but many time points. The module does not cover survival (event history or duration) or time-series analysis.

## Software

- Stata - version 14

## Prerequisites

Essential requirements for the module are:

- (a) Final year undergraduate level knowledge of linear regression methods (OLS regression) and some familiarity with issues like sample selection and endogeneity.
- (b) Intermediate level proficiency in Stata: familiarity with basic commands and experience of writing Stata do files.

## Schedule

July 3, 2017

Time	Basics - Longitudinal Data
14:00-15:45	Lecture: <ul style="list-style-type: none"> <li>• What are longitudinal data (examples)? □</li> <li>• Why use panel data? Advantages and challenges. □</li> <li>• Patterns of observations in panel data (non-response and attrition) □</li> <li>• Describing panel data</li> </ul>
15:45-16:15	Break
16:15-18:00	Lab: <ul style="list-style-type: none"> <li>• Data management: Building a working file using the British Household Panel Survey □</li> <li>• Handling panel data in Stata - some basic commands.               <ul style="list-style-type: none"> <li>○ Learn to use the commands <code>merge</code> and <code>append</code> to compile panel data □</li> <li>○ Learn the difference between wide and long panel format. Use command <code>reshape</code>. □</li> </ul> </li> </ul>
	<ul style="list-style-type: none"> <li>• Learn different ways of describing panel data.</li> </ul>

□

July 4, 2017

Time	Modeling panel data - Introduction □
14:00-15:45	Lecture: <ul style="list-style-type: none"> <li>• Modeling panel data - overview □</li> <li>• Challenges of panel data modeling □</li> <li>• Statistical dependencies:               <ul style="list-style-type: none"> <li>○ Serial-correlation in panel data</li> <li>○ Unobservables □</li> <li>○ Measurement error bias</li> </ul> </li> <li>• Pooled Ordinary Least Squares and robust standard errors □ □</li> </ul>
15:45-16:15	Break
16:15-18:00	Lab: <ul style="list-style-type: none"> <li>• Exploring serial correlations of continuous and categorical variables □</li> <li>• De-meaning □</li> <li>• Comparing the estimates and serial correlation of □               <ul style="list-style-type: none"> <li>○ Pooled OLS □</li> <li>○ Pooled OLS using time demeaned data □</li> </ul> </li> <li>• Exploring the associated factors of the unit-specific heterogeneity <math>u_i</math></li> </ul>

**July 5, 2017**

<b>Time</b>	<b>Modeling continuous-level dependent variables - Levels</b> □
14:00-15:45	Lecture: <ul style="list-style-type: none"><li>• Fixed effects regression □<ul style="list-style-type: none"><li>◦ Least Squares Dummy Variables (LSDV)</li><li>◦ Within regression (time demeaning) □</li></ul></li><li>• Random effects regression □</li><li>• Hybrid models □</li><li>• Test statistics □</li></ul>
15:45-16:15	Break
16:15-18:00	Lab: <ul style="list-style-type: none"><li>• Comparing the properties of estimates using several panel regression models. □</li><li>• Using different tests statistics. □</li></ul>

## **References**

Course largely based on:

Andreß, Hans-Jürgen, Golsch, Katrin, and Alexander W. Schmidt. 2013. *Applied Panel Data Analysis for Economic and Social Surveys*. Springer.

Longhi, Simonetta and Alita Nandi. 2015. *A Practical Guide to Using Panel Data*. Sage.

## **Short bio**

Dr. Anja Neundorf is an Associate Professor in Politics and Research Methods at the University of Nottingham. She previously held a Post-doctoral Prize Research Fellowship at Nuffield College, Oxford and received her PhD from the University of Essex. Her research interests lie at the intersection of political behavior, research methods, and comparative politics. Her research has been published in the *Journal of Politics*, *British Journal of Political Science*, *Public Opinion Quarterly*, and *Social Forces*.