SYLLABUS
PANEL DATA ECONOMETRICS
(PHD)
AUGUST 19-21, 2025

July 25, 2025

1 Administrative Details

Course meets: August 19-21, 2025, Online

Instructor: Yonas Alem (PhD), EfD/Department of Economics, University of Gothenburg; Teaching Assistant, Meseret Abebe (PhD), EfD Global Hub.

E-mail: yonas.alem@gu.se

Website: https://www.gu.se/om-universitetet/hitta-person/yonasalem

2 Course Description

A panel data set contains repeated observations over the same units for several periods. Panel data provides essential advantages in estimating more complex and realistic models than single cross-section or single time-series data. Using panel data, it is possible to address the omission of variables due to the presence of what is called 'unobserved individual heterogeneity'. This course aims to provide a detailed introduction to econometric models often used to analyze panel data in applied economic research for graduate students, interested faculty members, and applied researchers. The course significantly focuses on applications rather than proofs and derivations of the alternative estimators. A graduate-level cross-sectional econometrics course is a prerequisite for this course.

We will begin by discussing the basic linear panel data models used to deal with unobserved individual heterogeneity: fixed effects, FE (through a "within" transformation); random effects, RE (through assuming independent distribution of the unobserved individ-

ual heterogeneity term); and first-differences estimator, FD (through first differencing). We will also discuss the difference-in-differences (DID) estimator, a special first differences estimator, which is one of the most robust quasi-experimental causal inference methods.

Very often, economic research requires controlling for state dependence in a variable. For instance, a household's consumption expenditure in any period is a positive function of its expenditure in the previous period. Such relationships cannot be estimated using the basic linear panel data models, such as FE & RE, because they will result in biased parameter estimates due to correlation between the lagged dependent variable and the error term. One can estimate the model parameters consistently using dynamic panel data models provided that three rounds (or more) of panel data exist. We will discuss how the Anderson-Hsiao two-stage estimator works to estimate such models, as well as the Arellano-Bond Generalized Method of Moments (GMM) estimator, which yields more efficient parameter estimates when more than three rounds of panel data are available.

Next, we will discuss how popular estimators for binary choice models, such as probit and logit, can be extended to a panel data setting using the fixed effects logit and random effects probit estimators. We will also briefly discuss the random-effects tobit model, which is used when the dependent variable is censored. Controlling for a lagged dependent variable in binary models results in a serious econometric issue referred to as the initial conditions problem" due to the correlation between the lagged dependent variable and the unobserved individual heterogeneity term. We will wind up the course by briefly discussing how such a model can be estimated using alternative dynamic probit estimators.

I designed the course for three intensive, full-day sessions. The three days include two lab sessions, during which participants will practice estimating different estimators using Stata. There will also be a session on where to find panel data sets and what kind of research questions they could help answer. The Special Panel on Impact Assessment (SPIA) and the Environment for Development (EfD) generously fund the course. It is free for participants. Interested individuals should register by following the google link below by Friday, July 15, at the end of the day at the latest. If you have any questions, please feel free to approach Dr. Meseret Abebe, the teaching assistant (meseret.birhane.abebe@gu.se).

Google link: https://forms.gle/VcRL2iJGbGVBoQ2QA

3 Course Outline

3.1 Introduction to Panel Data Modeling

3.2 Static Linear Estimators

3.2.1 The Fixed Effects Estimator

3.2.2 The First-difference Estimator

3.2.3 The Random Effects Estimator

3.2.4 Fixed Effects or Random Effects

- Wooldridge (2010): Ch. 10, 11
- Cameron & Trivedi (2009): Ch. 8, 9
- Verbee (2017): Ch.10

3.3 Dynamic Linear Estimators

3.3.1 An Autoregressive Panel Data Model

3.3.2 Dynamic Models with Exogenous Variables

- Bond, S. (2002)
- Cameron & Trivedi (2009): Ch. 9

Lab 1: Estimating a Dynamic Model of Consumption

3.4 Non-linear Panel Data Estimators

3.4.1 Binary Choice Models

3.4.2 The Fixed Effects Logit Model

3.4.3 The Random Effects Probit Model

3.4.4 Tobit Models

- Wooldridge (2010): 15.8
- Cameron & Trivedi (2009):
- Verbeek (2017): Ch.10

3.5 Non-linear Dynamic Models and the Initial Conditions Problem

- Arulampalam & Stewart (2009)

Lab 2: Estimating a Models of Agricultural Technology Adoption

4 Tentative List of References

- Arulampalam, W., & Stewart, M. (2009). Simplified implementation of the Heckman estimator of the dynamic probit model and a comparison with alternative estimators. *Oxford bulletin of economics and statistics*, 71(5), 659-681.
- Bond, S. (2002), Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice, cemmap working paper CWP09/02.
- Cameron, A.C., & Trivedi, P. K. (2009). Microeconometrics Using Stata, Stata Press, College Station, Texas.
- Verbeek, M. (2017). A Guide to Modern Econometrics, fifth edition, John Wiley & Sons, Ltd.
- Wooldridge, Jeffrey M. (2010). Econometric Analysis of Cross Section and Panel Data (2nd ed.). The MIT Press.