Advanced Econometrics: Panel Data

Arthur van Soest, Tilburg University

Email: avas@uvt.nl

Objectives

The aim of the course is to give a broad overview of empirical models used to analyse panel data, with a focus on applications in microeconomics, explaining the agents' economic decisions or outcomes. The main econometric models and methods for panel data will be discussed, and the students will do hands on exercises in order to learn how to apply the techniques in empirical applications such as, possibly, their own research.

Content

Economic agents often make similar decisions repeatedly over time, taking account of changes in prices and other factors driving preferences, opportunities, or expected outcomes. Examples are consumption expenditures, savings and portfolio choice, or investment decisions in housing of private households; labour force participation and labour supply decisions, decisions on smoking, drinking and other kinds of health related behaviour of individuals, etc. Similarly, firms have to decide on hiring and firing, investing in human capital (by training their workers) or physical capital and in research and development, etc.

To analyze how economic agents make these decisions and the consequences of these decisions, empirical researchers make more and more use of panel data techniques. Panel data are data in which decisions of many economic agents are observed several times, usually at different points in time. For example, many socio-economic household surveys have data on wealth and household assets for a number of consecutive years, allowing for an analysis of the determinants of decisions to buy risky assets like stocks and bonds, etc. Panel data techniques are also useful for the analysis of stated preferences or economic experiments where the same subjects answer several questions or make a number of consecutive decisions.

This course presents theory and applications of panel data models, focusing on models that explain decision making by individuals, households, or firms and outcomes such as earnings, health or well-being. Students will learn the basic econometric models and techniques for estimation, testing and model selection. They

will study articles that apply these models and techniques to a variety of micro-economic issues. In addition, they will learn how to apply these models and techniques themselves by doing hands-on exercises using STATA. Applications will come from labour economics, economics of savings and consumer behaviour, development economics, environmental economics, and health economics.

Recommended Reading

- M. Verbeek, A Guide to Modern Econometrics (5th edition), Wiley, Chichester, (2017). Particularly Chapter 10.
- S. Cunningham (2021), Causal Inference – The Mixtape, Yale University Press. Particularly the chapters Panel Data, Difference-in-Differences, Instrumental Variables.
- A.C. Cameron and P.K. Trivedi (2005), Microeconometrics: Methods and Applications, Cambridge University press, New York. Particularly Chapter 23.
- Selected applications in journal articles (to be announced).

More Advanced Reading for Interested Students

- M. Arellano and B. Honoré (2001), Panel data models: some recent developments, in J. Heckman and E. Leamer (eds.), Handbook of Econometrics Volume 5, North-Holland, Amsterdam, pp. 3229-3296.
- J.W. Wooldridge (2010), Econometric Analysis of Cross Section and Panel Data (2nd edition), MIT Press, Cambridge MA.
- D. Arkhangelsky and G. Imbens (2024), Causal models for longitudinal and panel data: a survey, *The Econometrics Journal*, 27(3), C1–C61.

Required Prerequisites

A solid introductory course in econometrics; A solid introductory course in probability theory and statistics; Preferably some basic knowledge of Stata

Schedule

Monday:

11:30 – 12:00 Welcome meeting

12:00 – 13:00 Lunch

13:00 – 14:45 First session (lecture)
15:15 – 17:00 Second session (lecture & intro computer lab)
18:00 – 19:00 Dinner

Tuesday – Thursday:

7:15 - 8:30 Breakfast
8:30 - 10:00 First session (lecture)
10:00 - 10:30 Coffee break
10:30 - 12:00 Second session (lecture)
12:00 - 13:00 Lunch break
13:00 - 14:30 Third session (computer class)
14:30 - 15:00 Coffee break
15:00 - 16:30 Fourth session (computer class & wrap up)
16:30 - 18:00 Free time
18:00 – 19:00 Dinner

Friday:

7:15 - 8:30 Breakfast
8:30 - 10:00 First session (lecture)
10:00 - 10:30 Coffee break
10:30 - 12:00 Second session (lecture)
12:00 - 12:45 Lunch break
12:45 - 14:30 Final session (computer class & wrap up)

Overview

The morning sessions (and the first session on Monday) will mainly be devoted to the econometric models and techniques, with some applications to illustrate these. In the afternoon sessions, students will use Stata to do their own empirical analysis using case studies.

Note: For the computer classes, participants are expected to bring their own laptop with Stata installed. If this is not possible, please indicate this upon registration.

Provisional day-to-day program

Day 1: Brief review of linear regression models for cross-section data (Standard linear model, Generalized linear model, Instrumental variables, 2sls); Weak instruments; Introduction to static linear panel data models with random effects or fixed effects.

Day 2: Static linear panel data models, continued; Linear probability model; Difference in differences; Fixed effects IV estimation.

Day 3: dynamic linear panel data models; GMM and the Arellano-Bond and Blundell-Bond estimators;

Day 4: Static and dynamic binary choice models with random and fixed effects; Maximum Likelihood and Conditional Maximum Likelihood; the Wooldridge approach.

Day 5: Selected topics in other nonlinear panel data models (e.g. Censored regression models; ordered response models; count data models; mixed logit).

Exam

- The course concludes with a mandatory written exam on September 5, 2025, held at the students' home universities.
- Students may bring at most five double-sided A4 pages of notes to the exam.
- The available time to make the exam is 90 minutes

Applications in Journal Articles

Day 1:

Khandker, S.R. (2005), Microfinance and poverty: evidence using panel data from Bangladesh, *World Bank Economic Review*, 19, 263-286.

Khanna, J., J. Posnett and T. Sandler (1995), Charity donations in the UK: New evidence based on panel data, *Journal of Public Economics*, 56, 256-272.

Powdthavee, N. (2010), How much does money really matter? Estimating the causal effects of income on happiness. *Empirical Economics, 39, 77-92.*

Muto, M. and T. Yamano (2009), The impact of mobile phone coverage expansion on market participation: panel data evidence from Uganda, *World Development*, 37, 1887-1896.

Day 2:

Arellano, M. and S. Bond (1991), Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations, *Review of Economic Studies*, 58, 277-297.

Blundell, R. and S. Bond (1998), Initial conditions and moment restrictions in dynamic panel data models, *Journal of Econometrics*, 87(1), 115-143.

Witt, R., A. Clarke and N. Fielding (1999), Crime and economic activity: a panel data approach, *British Journal of Criminology*, 39(3), 391-400.

Michaud, P.-C., and A. van Soest (2008) Health and wealth of elderly couples: Causality tests using dynamic panel data models, *Journal of Health Economics*, 27, 1312-1325.

Day 3:

Gibbons, R., D. Hedeker, S. Charles and P. Frisch (1994), A random effects probit model for predicting medical malpractice claims, *Journal of the American Statistical Association*, 89(427), 760-767.

Banfi, S., M. Farsi, M. Filippini and M. Jakob (2008), Willingness to pay for energy-saving measures in residential buildings, *Energy Economics*, 30, 503-516.

Heitmueller, A. (2007), The chicken or the egg? Endogeneity in labour market participation of informal carers in England, *Journal of Health Economics,* 26, 536-559.

Arulampalam, W. and S. Bhalotra (2006), Sibling death clustering in India: state dependence *versus* unobserved heterogeneity, *Journal of the Royal Statistical Society*, 169, 828-849.

Poggi, A. (2007), Does persistence of social exclusion exist in Spain? *Journal of Economic Inequality*, 5, 53-72.

Biewen, M. (2009), Measuring state dependence in individual poverty histories when there is feedback to employment status and household composition, *Journal of Applied Econometrics,* 24, 1095-1116.

Day 4:

Vella, F. and M. Verbeek (1999), Two-step estimation of panel data models with censored endogenous variables and selection bias, *Journal of Econometrics,* 90, 239-263.

Grzybowski, L and P. Pereira (2008), The complementarity between calls and messages in mobile telephony, *Information Economics and Policy*, 3(9), 279-287.

Kang, S.J. and M.-J. Lee (2003), Analysis of private transfers with panel fixed-effect censored model estimator, *Economics Letters*, 80, 233-237.

Rettenmaier, A.J. and Z. Wang (2006), Persistence in Medicare reimbursements and personal medical accounts, *Journal of Health Economics,* 25, 39-57.

Day 5:

Euwals, R., A. Eymann and A. Boersch-Supan (1996), Who determines household savings for old age? Evidence from Dutch panel data, *Journal of Economic Psychology*, 25, 195-211.

Groot, W. and H. Maassen-van den Brink (2003), Firm-related training tracks: A random effects ordered probit model, *Economics of Education Review,* 22, 581-589.

Revelt, D. and K. Train (1998), Mixed logit with repeated choices: household choices of appliance efficiency level, *Review of Economics and Statistics*, 80, 647-657.

Hausman, J., B.H. Hall and Z. Griliches (1984), Econometric models for count data with an application to the patents-R&D relationship, *Econometrica*, 52, 909-938.

Lee, K.J. and S. Kobayashi (2001), Proportional treatment effects for count response panel data: effects of binary exercise on health care demand, *Health Economics*, 10, 411-428.