ECON 690: Selected Topics in Panel Data Econometrics

Professor Mohitosh Kejriwal

Spring 2021

Lectures: Tuesdays and Thursdays, 1:10-2:40pm, online via Zoom.

Office Hours: On Zoom by appointment.

Email: mkejriwa@purdue.edu

Grader: Xuewen Yu (yu656@purdue.edu)

Course Overview: This course covers selected topics in the econometrics of panel data with an emphasis on developing a clear understanding of the methodological issues involved as well as identifying a set of substantive empirical issues where such methods can be applied. Each topic will begin with a survey of the relevant state of the art econometric methods followed by a discussion of their applicability to analyzing various economic questions. The course will be primarily based on a discussion of journal articles and working papers. Familiarity with a matrix oriented programming language is essential. Considerable emphasis will be placed on the replication of simulation and empirical results from published papers.

Prerequisites: ECON 671 & 672 or equivalent/Consent of Instructor.

Course Format and Grading: There will be **no exams**. The evaluation for the course will be based on three assignments (20% each), two class presentations (20%) and a literature review (20%) on a topic selected from a list I will provide. The assignments will include replication of results from simulation experiments based on artificially generated data as well as empirical results based on real data. For most of the assignments, you will need to write your own program codes so I encourage you to work on them in teams (of at most two). One of the two presentations will be based on a published/working paper (Week 5, February 16-18) while the other will be based on your literature review (Week 8, March 9-11).

Course Website: All material related to the course, including recordings of the lectures, will be available through Brightspace at https://purdue.brightspace.com/. You will need to log in with your Purdue username and password.

Textbooks: There is no required textbook for the course. Books that may serve as useful references are: (1) *Time Series and Panel Data Econometrics,* Oxford University Press (by M. Hashem Pesaran), (2) *Econometric Analysis of Cross Section and Panel Data*, Second Edition, MIT Press (by J.M. Wooldridge), (3) *Panel Data Econometrics*, Oxford University Press (by M. Arellano), (4) "*Econometrics,*" (by Bruce Hansen), available at https://www.ssc.wisc.edu/~bhansen/econometrics/Econometrics.pdf

Emergency: In the event of a major campus emergency, course requirements, deadlines and grading percentages are subject to changes that may be necessitated by a revised semester calendar or other circumstances.

Course Description

- An inevitable fact as we move forward in time and as information technology improves is that data will be available for many more series and over an increasingly long span. While the availability of more data provides the opportunity to understand economic phenomena and anomalies better, researchers can also suffer from an information overload without some way to organize the data into an easy to interpret manner. The course is primarily concerned with the problem of cross section dependence in the analysis of large panel data models and their relevance in applied econometric research. Cross section dependence can arise due to spatial effects, or due to unobserved (or unobservable) common factors. We will attempt to understand the role of factor models as a tool for dimension reduction in the analysis of large-dimensional data as well as a device for modeling cross-sectional correlation in panel data analysis when the cross-sectional and time series dimensions are comparable in magnitude. Certain miscellaneous topics in the traditional panel data context such as dynamic panels, clustered standard errors, measurement error in panels, and models with random coefficients will also be covered.

- *Topics (Tentative List)*

1. Dynamic panel data models
2. Measurement Error in panel data models
3. Identification issues in factor models
4. Determination of the number of factors
5. Estimation of the factors
6. Estimation of panel regression models with a factor error structure
7. Inference when estimated factors are used as regressors
8. Aggregation in Large Panels
9. Testing for cross-sectional dependence
10. Clustered standard errors

- *Primary Reading List*

1. Bai, J. and Ng, S. (2008): "Large dimensional factor analysis," *Foundations and Trends in Econometrics* 3, 89-163.

2. Bai, J. (2003): "Inferential theory for factor models of large dimensions," *Econometrica* 71, 135-172.

3. Bai, J. and Ng, S. (2002): "Determining the number of factors in approximate factor models," *Econometrica* 70, 191-221.

4. Bai, J. and Ng, S. (2006): "Confidence intervals for diffusion index forecasts and inference with factor-augmented regressions," *Econometrica* 74, 1133-1150.

5. Bai, J. and Ng, S. (2007): "Determining the number of primitive shocks in factor models," *Journal of Business and Economic Statistics* 25, 52-60.

6. Bai, J. (2009): "Panel data models with interactive fixed effects," *Econometrica* 77, 1229-1279.

7. Boivin, J. and Ng, S. (2006): "Are more data always better for factor analysis?," *Journal of Econometrics* 132, 169-194.

8. Forni, M., Hallin, M., Lippi, M. and Reichlin, L. (2000): "The generalized dynamic factor model: identification and estimation," *Review of Economics and Statistics* 82, 540-554.

9. Pesaran, M.H. (2006): "Estimation and inference in large heterogeneous panels with a multifactor error structure," *Econometrica* 74, 967-1012.

10. Sarafidis, V. and Wansbeek, T. (2012): "Cross-sectional dependence in panel data analysis," *Econometric Reviews* 31, 483-531.

11. Wooldridge, J.M. (2010): "Econometric Analysis of Cross Section and Panel Data", Chapter 11.