**University of Pennsylvania** **Economics 7310, Spring 2023**

# Time-Series and Panel Data Econometrics

**Instructors**: Frank Schorfheide, PCPSE 621
Email: *schorf@upenn.edu*
Office Hours: see CANVAS

Wayne Gao, PCPSE 630
Email: *waynegao@upenn.edu*
Office Hours: see CANVAS

**Teaching Assistant**: Marko Mlikota
Email: *mlikota@sas.upenn.edu*
Office Hours: see CANVAS

**Course Organization**: All course materials as well as announcements are available through CANVAS. The lectures will take place on Tuesdays and Thursdays from 10:15 - 11:45 in PCPSE 202. The TA will offer weekly recitation sessions. Information about the recitations will be posted on CANVAS.

**Course Description**:

This course is a sequel to Econ 7300, which serves as prerequisite. Economic data sets typically have a cross-sectional dimension $N$ and a time dimension $T$. In Econ 7300 you have focused on the analysis of data sets in which the time dimension was $T = 1$. In this course we will consider econometric models for data sets in which the time dimension is $T > 1$. If $N$ is "small" relative to $T$, then these models are typically called time series models. A key feature of time series models is that they try to capture serial dependence in the data. If $T$ is "small" relative to $N$, then the analysis falls under the heading panel data econometrics, which traditionally has been a subfield of microeconometrics. A key feature of panel data models is that they try to capture cross-sectional heterogeneity, and, in some instances, also spatial correlation. Throughout this course, we will consider various $(N, T)$ constellations that you might encounter as economist. The course introduces you to a large set of econometric models that are important for applied work, but it will also set the foundation for theoretical work in econometrics. We study the properties of these models as well as inference methods and computational techniques, including Bayesian techniques.

The first part of the course (until spring break), taught by Frank Schorfheide, focuses on time series analysis. The second part of the course, taught by Wayne Gao, focuses on panel data analysis. Note that you cannot take one of the two parts for credit. To receive credit, you need to take both parts and you will obtain a single grade.

**Prerequisites**: Economics 7300 or equivalent graduate level econometrics.

**Courseware**: You can access the course materials via CANVAS. You can log-in from *http://upenn.instructure.com/*.

**Course Requirements**:

- **Problem Sets**: There will regular problem sets, assigned throughout the semester. The problem sets are designed to give you the opportunity to review and enhance the material learned in class. You are encouraged to form small study groups, however, each member of a study group has to submit his or her own write-up of the solution. These solutions must be submitted on the specified due dates. [20%]
- **Exam 1**: In-class written exam on Thursday, March 2. [40%]
- **Exam 2**: In-class written exam on Tuesday, April 25. [40%]

**Course Readings**: There is no one textbook that exactly matches the material covered in class. However, the new book by Bruce Hansen is a good reference and we recommend that you procure a copy:

Hansen, Bruce E. (2022): *Econometrics*, Princeton University Press, ISBN 9780691235899.

We will make lecture slides/notes available on CANVAS and post further recommended readings for each lecture.

## Course Outline

### Part 1: Time Series Analysis (Frank Schorfheide)

**Univariate Time Series:** $N = 1$**, Large** $T$

Economic time series are characterized by serial dependence and trends. The goal of this part is to introduce you to notions of temporal dependence (and the lack thereof) and to exam how dependence affects the convergence of sample averages. In linear models temporal dependence can be introduced through moving average (MA) errors and an autoregressive (AR) structure, whereby the expected value at time $t$ is a function at the realization in periods $t-1, t-2, \ldots$. We will study the properties of models belonging to the ARMA class. We proceed by introducing trends into the model. Trends come in two forms: deterministic and stochastic. The presence of stochastic trends requires an alternative form of Central Limit Theorems. Finally, we will consider the estimation of AR models. This part will introduce you to Bayesian inference.

- **Dependent Processes:** empirical measures of temporal dependence, covariance stationarity, WLLNs and CLTs for MA processes, strict stationarity, Martingales.
- **Stationary ARMA Processes:** AR models, autocovariance functions, infinite-order MA processes, Wold decomposition, ARMA models.
- **Trend and Difference Stationarity:** trend stationary processes, difference stationary processes, functional central limit theorem and unit root asymptotics.
- **Estimation of AR Models:** likelihood function, frequentist analysis of the conditional MLE, Bayesian analysis.

**Vector Time Series: Fixed $N > 1$, Large $T$**

We now consider models for data sets in which the cross-sectional dimension is greater than $N = 1$, e.g., GDP growth, inflation, and interest rates, instead of just GDP growth. One of the workhorse models for the analysis of multiple time series, or vector time series, is the vector autoregression (VAR). VARs allow us to study the propagation of shocks through the system. We discuss their estimation, and how they can be used to generate forecasts, and study the effects of unanticipated policy changes. We proceed by generalizing the VARs to linear state-space models. The class of state-space models is large and versatile. It includes for instance, VARs with time-varying coefficients, or VARs for observations sampled at different frequencies. The analysis of state-space models requires additional computational techniques.

- **Vector Autoregressions:** theoretical properties, Granger causality, impulse response functions, likelihood function.
- **Bayesian Estimation of VARs:** direct sampling from posterior, Minnesota prior, data-driven hyperparameter selection and shrinkage, forecasting.
- **State-space Models:** specification, filtering, smoothing, data augmentation and Gibbs sampling.

## Part 2: Panel Data Analysis (Wayne Gao)

**Traditional Panel Data Models: Large $N$, Small $T$**

We switch gears and resize our data set, by making the time dimensional relatively short and the cross-sectional dimension large. In order to model panels, one has to allow for some coefficient heterogeneity across units. If the time dimension is small, these coefficients cannot be estimated consistently. One approach is to transform the data (temporal differencing, demeaning) to eliminate the heterogeneous coefficients (e.g., intercepts) and focus on the estimation of homogeneous coefficients. An alternative approach is to use Bayesian/shrinkage techniques to generate estimates of the heterogeneous coefficients that are "good", albeit not consistent.

- **Basics:** Capturing cross-sectional heterogeneity through unit-specific coefficients, incidental parameter bias, fixed versus random effects.
- **Static Panel Data Models:** random effects estimation, fixed effects estimation, correlated random effects estimation.
- **Dynamic Panel Data Analysis:** quasi maximum likelihood estimation and incidental parameter bias, moment-based estimation, correlated random effects estimation.
- **Estimation of Heterogeneous Coefficients and Forecasting with Panel Data Models:** a decision-theoretic approach, empirical Bayes methods, full Bayesian analysis.
- **Miscellaneous Topic:** Errors in variables, unbalanced panels, difference-in-difference estimation.

### Large $N$, Large $T$ Models

In the past three decades, the time dimension of many panel data sets increased substantially such that both dimensions of the data set can be regarded as "large." There are basically two separate strands of literature. The traditional panel literature has used the growth of the time dimension of the panel data sets to develop corrections for the incidental parameter bias problem and to facilitate the estimation of unit-specific coefficients. The VAR literature has expanded the cross-sectional dimension of their models to analyze not just, say, four to six macroeconomic time series, but, say, one- or two-hundred time series.

- **Large $N$, Large $T$ approximations for traditional panel data models.**
- **Factor Models:** principal component analysis, dynamic factor models as state-space models.