

Time Series Econometrics and Text Mining – Preliminary Syllabus (18.7.2022)

Organizational details

Instructors: *Prof. Dr. Peter Winker, Jenny Bethäuser und Viktoriia Naboka (Gießen)*

Date: September, 12th to 15th, from 9.30 to 17.30

Locations: HS 002, Licher Str. 68, 35394 Giessen

ECTS: 4 - 6 ECTS

Aims

Time series represent an important class of data in different fields of applications including macroeconomics, financial market economics, empirical political science, geography and many others. The statistical modelling of these data differs substantially from the analysis of cross sectional data as the explicit temporal structure has to be taken into account. Otherwise, relevant information content might be lost and the risk of spurious or nonsense regressions arises. We will study principles of univariate and, in particular, multivariate time series analysis. In addition, we consider a new class of data, namely “text”, and how it might be used in econometric analysis, again with a focus on time series.

Active participation in the PhD workshop will result in an overview on some central concepts of time series analysis and text mining including aspects of their practical application. Participants will be able to judge the appropriateness of empirical models applied to time series both in univariate and multivariate settings. They can interpret findings resulting from time series analysis and might conduct own analyses. Furthermore, they learn principles of translating text to data and how to make use of text based indicators in a time series context.

Course Contents

The workshop provides an introduction to the following concepts in time series analysis: ARMA modeling, non-stationarity and cointegration, VAR models: estimation, interpretation and forecasting, SVAR modeling, overview on text based methods, topic modelling and word embeddings, and generating text based indicators. It will also comprise some practical sessions using the econometric software R (for the first part of the course; alternatively, EViews can be used) and Python (for the text-based analysis).

Target group

Doctoral candidates in the MAGKS network with basic knowledge in inferential statistics and basic linear regression analysis as well as experience in working with some statistical software. The course intends to provide an introduction to time series analysis with a focus on multivariate linear models as well as first steps in computer based text analysis.

Course language

English

Registration

Registration is possible through the MAGKS-website or directly with Peter Winker.

Requirement

For active participation in the hands on sessions, you need a Laptop with installed R (e.g. R Studio) and Python. You should have installed already the following packages for R and Python (please note that we might send a further update on packages a few days before the course starts):

R: readxl, vars, ReporteRs or officer, plm, tsDyn, forecast, Metrics

Python: Python \geq 3.6 (e.g. via anaconda <https://www.anaconda.com/distribution/>). Please also have the python packages scikit-learn installed (pip install -U scikit-learn)

Tentative Syllabus

Monday, September 12:

- 9.30 – 11.00 Basics of time series analysis including ARMA modelling
- 11.30 – 13.00 Nonstationarity and Cointegration
- 14.00 – 15.30 Hands on: basic concepts of time series analysis using R
- 16.00 – 17.30 Hands on: testing (non-)stationarity using R

Tuesday, September 13:

- 9.30 – 11.00 Basics of VAR modelling
- 11.30 – 13.00 Interpreting VAR models
- 14.00 – 15.30 Hands on: specification and estimation of VAR models using R
- 16.00 – 17.30 Hands on: (structural) analysis of VAR models using R

Wednesday, September 14:

- 9.30 – 11.00 Bootstrap and confidence bands for VAR models
- 11.30 – 13.00 Hands on: further aspects of VAR modelling using R
- 14.00 – 15.30 Introduction to text mining in economics
- 16.00 – 17.30 Hands on: dealing with textual data using Python

Thursday, September 15:

- 9.30 – 11.00 Topic modelling: LDA and word embeddings
- 11.30 – 13.00 Obtaining text based economic indicators
- 14.00 – 15.30 Hands on: topic modelling using Python
- 16.00 – 17.30 Hands on: generating text based indicators using Python