Syllabus: Financial Econometrics – Preliminary
NHH
PhD course
Fall 2021
Professor Bernt Arne Ødegaard (BAØ), NHH and UiS. (bernt.a.odegaard@uis.no)
Professor Eric de Bodt (EdB), NHH and University of Lille. (eric.debodt@nhh.edu)

Overview

This course is a first course in empirical finance at the PhD level. The course attempts to lay the groundwork for students who will later do empirical research work. It is therefore a hands on course where the students will have to perform analysis on actual data, and where the examples are chosen to illustrate the typical questions asked in finance research. In the analysis of data we will use the typical computer tools for doing such analysis. For any nontrivial empirical analysis we have to use other tools than Excel and similar spreadsheets. Any empirical researcher have to be familiar with a range of computer tools, and choose the right tool for a given estimation problem. The typical tools are a general tool for numerical linear algebra, such as `Matlab`, and a more dedicated statistics package. We will in the course use the statistics packages `R` and `Stata` for most of the analysis.

Delivery

Bernt Arne Ødegaard's sessions will be delivered as physical lectures in four two-day sessions during the fall. Eric de Bondt's sessions will be delived as weekly video lectures and zoom discussions, with a video lecture posted on monday, starting in october.
The current plan is as follows:
Physical lecture meetings (BAØ)
Session 1 (BAØ) Sep 9/10
Session 2 (BAØ) Sep 26/27
Session 3 (BAØ) Oct 21/22
Session 4 (BAØ) Nov 18/19

Video Schedule (EdB)
Week 40: 4 oct video lecture, 8 oct zoom discussion
Week 41: 11 oct video lecture, 15 oct zoom discussion
Week 42: 25 oct video lecture, 29 oct zoom discussion
Week 44: 1 nov video lecture, 5 nov zoom discussion
Week 45: 8 nov video lecture, 12 nov zoom discussion
Week 47: 22 nov video lecture, 26 nov zoom discussion
Week 48: 29 nov video lecture, 3 dec zoom discussion

As always these days, things may change depending on the Covid-19 situation.
For up-to-date information, and lecture notes, look at the

Course Homepage

`https://ba-odegaard.no/teach/empir_finance_2021/index.html`

Course Contents

- Toolchest for empirical work: Matlab / Julia / R / Stata (BAØ).

- Basic econometrics (BAØ).
  - Linear regressions with applications.
  - Maximum likelihood, example application: Binary Choice regressions
- Investigating the cross section of asset returns (BAØ).
  - Tools to investigate risk-return tradeoff: Black, Jensen, and Scholes (1972) regressions, Fama and MacBeth (1973) analysis, GMM, Principal Components (APT).
  - The Fama-French analysis and current state of the art.
- Event Studies (BAØ).
- Time Series Tools and analysis (BAØ).
  - Univariate time series modelling. VARs. ARCH.
  - Predictability and forecasting.
- Stochastic Discount Factor based asset pricing.
- Data Snooping (BAØ).
- An Introduction to Econometric Programming using `Stata` (EdB).
  - Using `Stata`.
  - Panel data and diff-in-diff estimators.
  - From regression to Lasso and Elastic Nets.

Evaluation
Course evaluation will be based on student handins to empirical problems. In the problems you are typically given a dataset which you need to analyze, and write up your analysis.
You need to do the exercises as you would write the results in an academic paper: Tables summarizing results, detailed descriptions of what is estimated in the table, and a text discussion of what the results mean. In an appendix you should provide the exact estimation in the form of `Matlab/R/Stata` code and output.

Computer Tools
To do empirics one will need to improve the nerd gene a bit. The first challenge for the students in this course is therefore make sure you have available the typical tools. `Stata`, the statistics program, should be available on the schools computers. `R`, a public domain package for doing statistics, is generally installed at most universities. Since it is an open source package, it can be installed on any computer. It is available at `www.r-project.org`. You will also need a matrix handler. Here there are a number of alternatives available. `Matlab` is the alternative for people with no money troubles, it is a commercial program used in the industry. A number of open alternatives to `Matlab` can be used. One is `octave`, available at `www.octave.org`. Another is `scilab`, available at `www.scilab.org`. However, these two programs are being replaced by `Julia`, a more

powerful tool. For the types of examples we do in this course, the syntax of `Julia` is very similar to `Matlab` and `Octave`. `Julia` is also open soource and can be installed anywhere.

Datasets

In the course we will be looking at various examples. A number of typical finance datasets will be referred to. These datasets will both be used in examples in class that you should try to replicate, and in the exercises you should turn in.

Textbooks

There is no single textbook for this course. Campbell (2018), Cochrane (2005) and Ferson (2019) are used as references, and are high-level summaries of both theoretical and empirical asset pricing. Another, although now slightly dated book, is Campbell, Lo, and MacKinlay (1997), which contains a lot of useful information about the classical tests.

Most of the lectures will be accompanied by detailed lecture notes on the course homepage, together with slides etc. Those lecture notes will be relatively self-contained and complete.

In this course we will refer to various econometric, mathematical and computing topics that may or may not be known to you from before. If these are completely new areas for you, you may want some guidance and references. I'm listing a reference or two to various textbook sources I find useful.

- Econometrics. One standard text for PhD level studies is Greene (2018), which contains most of what one need to check up on, but it not an introductory text. If you have had no econometrics before, I view Stock and Watson (2019) as a good introduction for finance students. For part of the material Angrist and Pischke (2008) is used as a reference.

- Computing. Most of the computer tools we work with have loads of information easily accessible on the web. But sometime it is helpful with more of a gentle overview and introduction to the tools in paper format. Let me mention a couple of possibilities for the various tools.

  - `R`: Adler (2010) is a nice overview of the usage of `R`.
  - `Octave`: Hansen (2011) is a similar book for `Octave`.

These books are available from amazon and similar online retailers.

Detailed overview of the course

Detailed teaching overview, grouped by teacher.

Taught by *Bernt Arne Ødegaard*

- Part I: **Toolchest**
  - Introduction.
  - Data gathering.
  - Introducing the computing tools.
  - Linear Algebra.
    - Using a matrix program for Linear Algebra.
    - Some financial applications of matrix algebra.

  - The geometric interpretation of least squares.
  - `R` - the statistics tool
  - The value of a good picture.

- Part II: **Basic econometric analysis**, finance applications.
  - Linear Regressions
    * Financial application: Performance of equity portfolios.
    * Dummies in Regressions.
  - Maximum Likelihood
    * Maximum Likelihood estimation
    * Discrete Choice as an alternative to regressions.

- Part III: **Crossectional Asset Pricing**

  This lecture uses the problem of explaining the *cross-section* of asset returns, i.e. the relation between risk and return, what determines the riskiness of a single asset, as the unifying theme. We start with the classical tests of CAPM, show methodological innovations ending up in GMM estimation in the cross section, introduce the Fama French factors, and end up with the current "state of the art" in crossectional modelling.

  - Violations of OLS Assumptions - GLS, HAC corrections
  - CAPM testing (classical): Black et al. (1972), Fama and MacBeth (1973)
  - The APT, looking for additional explanatory "factors." Factor Analysis. Principal Components.
  - Introduction to GMM. Chaussé (2010), Jagannathan, Skoulakis, and Wang (2002)
  - CAPM tests in a GMM setting: MacKinlay and Richardson (1991)
  - Gibbons-Ross-Shanken statistics.
  - The Fama and French sequence of "factors", from three to five and beyond.
  - Current Status of crossectional estimation.

- Part IV: **Event studies**
  - Event Studies in Economics and Finance. (MacKinlay, 1997).

- Part V: **Time Series**
  - What is special about time series?
  - Univariate time series modelling
  - ADL
  - VAR
  - Volatility modelling - ARCH

- Part VI: **Predictability and Market Efficiency**
  - Predictability
    - Trading strategies
    - Useful tool: Factor-mimicking portfolios.
    - Forecasting
  - Market Efficiency
    - How to measure market efficiency
    - Application: #`fincap`
- Part VII: **Data Snooping etc**. Recent trend in finance research: Asking existential questions. Data snooping, p-hacking, replicability.
  - Mclean and Pontiff (2016).
  - Campbell Harveys presidential adress.
  - The factor zoo.

Taught by *Eric de Bondt*:

An Introduction to Econometric Programming using Stata

Lecture 1 : Anatomy of a research project in Stata

• Writing Stata code • Stata coding guidelines • Optimization: • Compiled commands • Implicit loops • Structuring Stata code for research projects • Data management • Descriptive statistics • Models estimation • Tables/figures generation • The first seven : cls/clear/pause on/set more off/set linesize/cd/version • Structuring the code: program/args • My best friends: display/pause/log/timer/trace

Assignment : Restructuring an existing code

Lecture 2: The Building Blocks

• directories: cd/pwd • data types: byte/int/long/float/double/str/strL • missing data: missing() / mvencode + mvdecode • handling errors: capture / r() • controlling the display: quietly • local / global macro (literal, expression, nested, evaluation, ++) • macro extended functions • scalars/matrices • namespaces • command classes: r_class/e_class/c_class/s_class • stored results • return/ereturn • preserve/restore • implicit loops: by & statsby prefix, _n & _N constants • explicit loops: foreach/for values • Loops on macros, variables, etc • Loops on observations (x[i])

Assignment: Testing multiple model specifications

Lecture 3 : More Building Blocks

• varlist (wildcards * and -) / numlist • creating & updating variables: generate / replace / egen / egenmore • selecting observations • keep / drop • if/in/cond()/inlist()/inrange() • data conversion functions and commands: • real()/string()/recode()/autocode()/egen cut()/egen group() • tostring/destring/recode/encode/decode/recode • factor variables / leads and lags (calendar dates) versus x[_n-1] • managing dates • Date data types • Date formats : %td/%tc/... • From date literal to date value : td()/tw()/... • Date conversion functions: date()/mdy()/... • memory management: compress/memory/task manager

Assignment: Computing Hoberg and Phillips 10 nearest neighbors average similarity score

Lecture 4 : Data management • importing data: infile/insheet/import excel • augmenting data: append/merge • restructuring dta: • reshape • fillin/cross/joinby • xpose / stack / separate • testing: assert/duplicates/isid/unique/codebook

Assignment : Importing SDC data and merging with CRSP/Compustat data

Lecture 5 : More on programming

• Tables: estimates/esttab/estadd/tabout • Figures: • Graph types • Essential twoway options • Stata matrix : svmat/mkmat/matrix list & operators • postfile • frame

Assignment: Bootstrapping two stage estimators

Lecture 6 : Panel data and diff-in-diff estimators

• Panel data management • xtset • xtdescribe • xtsum / xttab • xtdata • xtline • Panel data regression • xtreg / areg • reghdfe • xtivreg

Assignment: Trade flows and cross-border M&A activity – testing for a causal relation

Lecture 7 : From regression to Lasso and Elastic Nets – a first step towards machine learning

• Stepwise regression • Lasso for prediction on model selection

Assignment: Agnostic fundamental analysis works, Bartram and Grinblatt (2018)

References

Joseph Adler. *R in a nutshell*. O'Reilly, 2010.

Joshua D Angrist and Jörn-Steffen Pischke. *Mostly Harmless Econometrics*. Princeton University Press, 2008.

Söhnke M. Bartram and Mark Grinblatt. Agnostic fundamental analysis works. *Journal of Financial Economics*, 128(1):125–147, 2018. doi: 10.1016/j.jfineco.2016.11.008.

Fisher Black, Michael Jensen, and Myron Scholes. The capital asset pricing model, some empirical tests. In Michael C Jensen, editor, *Studies in the theory of capital markets*. Preager, 1972.

John Y Campbell. *Financial Decisions and Markets. A Course in Asset Pricing*. Princeton University Press, 2018.

John Y Campbell, Andrew W Lo, and A Craig MacKinlay. *The econometrics of financial markets*. Princeton University Press, 1997.

Pierre Chaussé. Computing generalized method of moments and generalized empirical likelihood with r. *Journal of Statistical Software*, 34(11):1–35, 2010.

John Cochrane. *Asset Pricing*. Princeton University Press, revised edition, 2005.

Eugene F Fama and J MacBeth. Risk, return and equilibrium, empirical tests. *Journal of Political Economy*, 81:607–636, 1973.

Wayne Ferson. *Empirical Asset Pricing. Models and Methods*. MIT Press, 2019.

William H Greene. *Econometric Analysis*. Pearson, eight edition, 2018.

Jesper Schmidt Hansen. *GNU Octave, Beginner's Guide*. PACKT Publishing, 2011.

Ravi Jagannathan, Georgios Skoulakis, and Zhenyu Wang. Generalized method of moments: Applications in finance. *Journal of Business & Economic Statistics*, 20(4), October 2002.

A Craig MacKinlay. Event studies in economics and finance. *Journal of Economic Literature*, XXXV: 13–39, March 1997.

A Craig MacKinlay and Matthew P Richardson. Using generalized method of moments to test mean-variance efficiency. *Journal of Finance*, 46:511–27, 1991.

R David Mclean and Jeffrey Pontiff. Does academic research destroy stock return predictability? *Journal of Finance*, LXXI(1):5–30, 2016.

James H Stock and Mark W Watson. *Introduction to Econometrics*. Pearson, 4th edition, 2019.