

Methods of Empirical Finance

Seminar (UE)

Christoph Huber
University of Innsbruck

Master in Banking and Finance
Winter term 2019/20 (this version: 2019-11-19)

Course outline (Syllabus)

Methods of Empirical Finance

Objective

- this course provides an introduction to basic methodological concepts, methods, and models commonly applied in economics and finance, their weaknesses and strengths, as well as their fields of application
- by focusing on how to choose, apply, and interpret different methods and models, the course provides the essential knowledge to conduct empirical research on your own — especially with respect to data analysis and methodological issues for your master's thesis

Approach

- (**Lecture, VO**) the course tries to present just as much theory as necessary to understand what you are doing and to provide you a sufficient basis to broaden your knowledge in empirical finance, more sophisticated methods and models, and statistical data applications on your own
- (**Seminar, UE**) in a hands-on approach, methodological and theoretical concepts are applied to answer research questions from the field of empirical finance using an appropriate statistical software package

Grading

- grades in the seminar (UE) are based on two assignments to be conducted in pairs; more detailed information on the problem sets will be provided on time
- to pass the seminar with a positive grade, at least 50% of the points in each of the two assignments must be reached: assignments not handed-in in time will result in a negative seminar grade

Grading

- the following grading key will be applied:

Points	Grade
< 50.0%	'deficient' (5)
50.0% – 62.5%	'sufficient' (4)
62.5%– 75.0%	'satisfactory' (3)
75.0%– 87.5%	'good' (2)
> 87.5%	'very good' (1)

- overall course grades equal the ECTS weighted average of the grades in the lecture (VO) and the seminar (UE), i.e. $0.6 \times \text{VO-grade} + 0.4 \times \text{UE-grade}$

Dates

- the seminar will take place at the following times:

Date/Time	Room
Tu 19.11.2019, 10:00-12:45	AR 4 (ZID)
Tu 26.11.2019, 10:00-12:45	AR 4 (ZID)
Tu 03.12.2019, 10:00-12:45	AR 4 (ZID)
Tu 10.12.2019, 10:00-11:45	AR 4 (ZID)
Tu 07.01.2020, 10:00-11:45	AR 4 (ZID)

- the assignments have to be handed in by the following times:

Date/Time	Assignment
Fr 06.12.2019, 23:59	Assignment 1
Fr 17.01.2020, 23:59	Assignment 2

Contact

Christoph Huber, MSc

Department of Banking and Finance
University of Innsbruck
Universitätsstraße 15, 6020 Innsbruck
4th floor, room o.4.08

e-mail: christoph.huber@uibk.ac.at

phone: +43-(0)512-507-73015

chr-huber.com

Introduction

Methods of Empirical Finance

What are we going to cover in the seminar?

- selected topics, which...
 - ... relate to the lecture (VO)
 - ... I think are *interesting* and/or *useful*
(e.g. for future seminar papers, your Master's thesis, etc.)

Software

- you will use statistical software packages
 - in the seminar
 - in the assignments

to **analyze empirical data**

- you are free to use the software you are most familiar with, *as long as* the software package allows for *writing scripts*!
- suitable statistical software packages are, among many others:
R, Stata, Eviews, Matlab, etc.
- however, *I* will mainly use **R**

Documentation

- note that in analyzing empirical data, it is important to **document each step** of the analysis - from loading the data into the statistical software package to transforming the data and creating tables and figures - for replicability purposes
- possible ways to document each step of the analysis are:
 - Notebooks (e.g. *R Markdown*, *Jupyter*, etc.)
 - Annotated code (in the `.R-` or `.do-script` etc.)

Introduction to R

Methods of Empirical Finance

Introduction to R

- R project homepage: <https://www.R-project.org/>
- Open-source software project, GNU General Public License (GPL).
- Comprehensive R Archive Network (CRAN): <https://CRAN.R-project.org>

Installation

- Go to CRAN, pick up the version for your operating system, follow instructions in readme file.
- Microsoft Windows: Download and run setup .exe file.
- Mac OS X: Installer package .pkg for base system and platform-specific GUI, along with additional programming tools (as disk image .dmg files).
- Linux: Pre-packaged binaries for various flavors (.deb or .rpm files), also interfaced in various update managers (*apt*, *yum*, etc.).

Introduction to R

R as a Calculator

```
1 + 1
```

```
## [1] 2
```

```
2^3
```

```
## [1] 8
```

Mathematical functions: e.g. `log()`, `exp()`, `sin()`, `asin()`, `cos()`, `acos()`, `tan()`, `atan()`, `sign()`, `sqrt()`, `abs()`, `min()`, `max()`, ...

```
log(exp(sin(pi/4)^2) * exp(cos(pi/4)^2))
```

```
## [1] 1
```

Introduction to R

Vector arithmetic

Generation of vectors: e.g., via `c()`:

```
x <- c(1.8, 3.14, 4, 88.169, 13)
length(x)
```

```
## [1] 5
```

Assignment operators: `<-` or `=`

Subsets of vectors:

```
x[c(1, 4)]
```

```
## [1] 1.800 88.169
```

Examples:

```
2 * x + 3
```

```
## [1] 6.600 9.280 11.000 179.338 29.000
```

```
5:1 * x + 1:5
```

```
## [1] 10.000 14.560 15.000 180.338 18.000
```

```
log(x)
```

```
## [1] 0.5877867 1.1442228 1.3862944 4.4792554 2.5649494
```


Introduction to R

Data management

```
mydata <- data.frame(one = 1:10,  
                     two = 11:20,  
                     three = 21:30)
```

```
mydata
```

```
##      one two three  
## 1      1  11    21  
## 2      2  12    22  
## 3      3  13    23  
## 4      4  14    24  
## 5      5  15    25  
## 6      6  16    26  
## 7      7  17    27  
## 8      8  18    28  
## 9      9  19    29  
## 10     10  20    30
```

Data frames: Basic data structure in R.

Select columns:

```
mydata$two
```

```
## [1] 11 12 13 14 15 16 17 18 19 20
```

```
mydata[, "two"]
```

```
## [1] 11 12 13 14 15 16 17 18 19 20
```

```
mydata[, 2]
```

```
## [1] 11 12 13 14 15 16 17 18 19 20
```

Introduction to R

Data management

Import

```
# Plain text  
newdata <- read.table("mydata.txt",  
                      header = TRUE)
```

```
# Excel spreadsheet (.xls, .xlsx)  
library(readxl)  
newdata <- read_excel("mydata.xls")
```

```
# Stata files  
library(foreign)  
newdata <- read.dta("mydata.dta")
```

Export

```
# Plain text  
write.table(mydata, file = "mydata.txt",  
           col.names = TRUE)
```

```
# Stata files  
library(foreign)  
write.dta(mydata, file = "mydata.dta")
```

R format

```
save(mydata, file = "mydata.rda")  
load("mydata.rda")
```

Introduction to R

Data management

Factors:

Categorical information is stored in *factors*, e.g. gender, ethnicity, species, etc.

```
g <- rep(0:1, c(2, 4))  
g <- factor(g, levels = 0:1, labels = c("male", "female"))  
g
```

```
## [1] male   male   female female female female  
## Levels: male female
```

Introduction to R

Data management

Missing values:

Missing values are coded as NA (for "not available"). For many functions you can use the option `na.rm=TRUE` to ignore missing values. E.g.:

```
x <- c(4, 7, 3, 2, NA, 16, NA, 8)
```

x contains two missing values

```
is.na(x)      # shows for each data point whether it is NA
```

```
## [1] FALSE FALSE FALSE FALSE  TRUE FALSE  TRUE FALSE
```

```
sum(is.na(x)) # calculates the sum of all missing values
```

```
## [1] 2
```

```
mean(x)
```

```
## [1] NA
```

```
mean(x, na.rm=TRUE)
```

```
## [1] 6.666667
```

Introduction to R

Packages

Installing and loading packages:

- If connected to the internet, simply type `install.packages("partykit")` for installing *partykit*.
- Additionally for Windows, Mac, RStudio: GUI installer menus.
- Packages are installed in *libraries* (= collections of packages).
- Library paths can be specified (see `?library`).
- Packages are loaded by the command `library()`, e.g., `library("partykit")`.
- `library()` lists all currently installed packages.

CRAN task views: Overview of packages for certain tasks (e.g., environmetrics, psychometrics, time series, . . .).

<https://CRAN.R-project.org/web/views/>

Your turn

- Load the data in file `sp500_data.csv` into your statistical software package
- This file contains data on companies in the *S&P 500 Index* from *Bloomberg*

- Get an overview about the available data
- Prepare *descriptive statistics* about the data at hand
- Try to prepare summary statistics by industry/sector
- Try to prepare figures describing one or more variables in the data set you find interesting

- Make sure to use a *script* and *document* how you solved the task