TU-Berlin

Microeconometrics

Tutorial: Introduction

Tomas Repasky

WS 2016/17

Technische Universität Berlin
Institut für Volkswirtschaftslehre und Wirtschaftsrecht
FG Ökonometrie und Wirtschaftsstatistik

Basic Information

- Name: Tomas Repasky
- Office: H 5103
- Office hours: thursdays 11:00-13:00
- Email: *repasky@tu-berlin.de*

... About the Tutorials

- We will meet tuesdays 14:15 - 15:45 in H 0106.
- You will profit the most if you bring your laptop with you.
- Everything will be uploaded to our ISIS-page after the Tutorials, hence, you will be able to rerun everything at home.
- Try to solve the problems in the tutorials and at home, since you will need some skills in R for the assignments.

... About the Concept of the Course

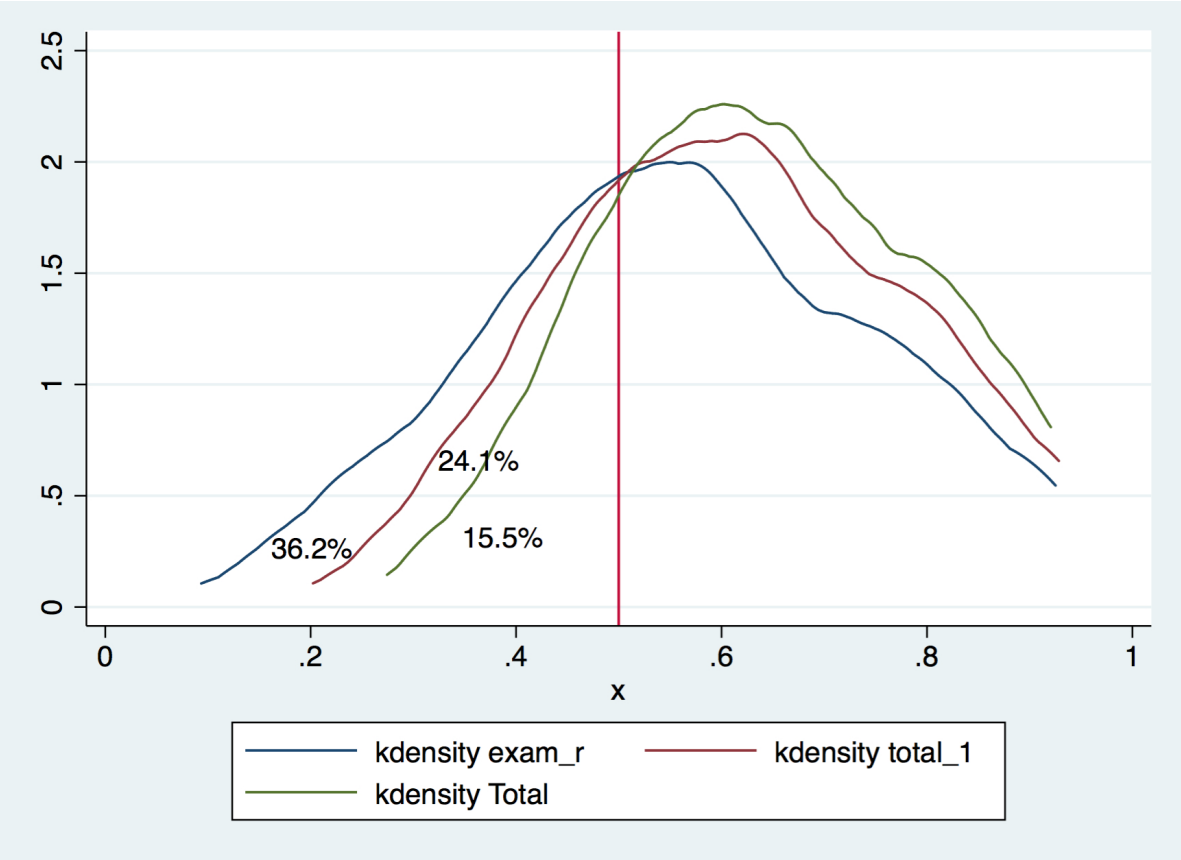
- There will be 2 assignments worth 30% of the points.
- The first assignment will be worth 20% and the second 10% .
- There will be a final exam with the duration of 75 min. worth the rest.
- You will need intermediate proficiency in R for the assignments.
- In the assignments, your programming skills, your understanding of the underlying topics as well as your interpretation skills will be challenged.

... About the Assignments

- The assignments will be handed in by groups of two.
- Do not cheat - it is more obvious than you think.
- If you cheat in an assignment, not only you will get 0 points but you will not be able to write the final exam at all!
- 1^{st} Assignments is due on the Monday, 19.12.2016 at 20:00.
- You will get the Assignment approximately 2-3 week before the due date.
- It pays off to learn for the class during the semester, since you will be able to use the gained knowledge in the assignments.

Motivation

The assignments will substantially help you to pass the exam and get a better grade!



Probability densities for achieving 0-100 % $x - axis$ in the exam (blue), exam $+1^{st}$ assignment (red) and exam $+$ both assigments (green) after the exam in Ferbruary

If you don't remember most of your Econometrics class

- If you have trouble grasping the basics from your Econometrics class:
  - Please come to the Econometrics tutorials to refresh your knowledge
    - mondays 10-12 in H 1029 or 12-14 in H 1028
    - wednesdays 12-14 in H 2033 or 14-16 in H 1028

Technicalities

1. You can and have to register for the class until 19.12.2016 via QISPOS (or until 30.11.2016 with the Examination Office).
2. After you have handed in your first assignment (on the 19.12.2016) you can't withdraw from the examination anymore (you can still prove that you were sick on the day of the exam of course).
3. Your assignments are part of the exam. That is why you cannot "unregister" once you have handed in one of them.
4. Therefore, if you cheat in your assignment, you will be excluded from the examination with the usual consequences. In sum, cheating in exam and in the assignments is linked with the same consequences.

Technicalities II

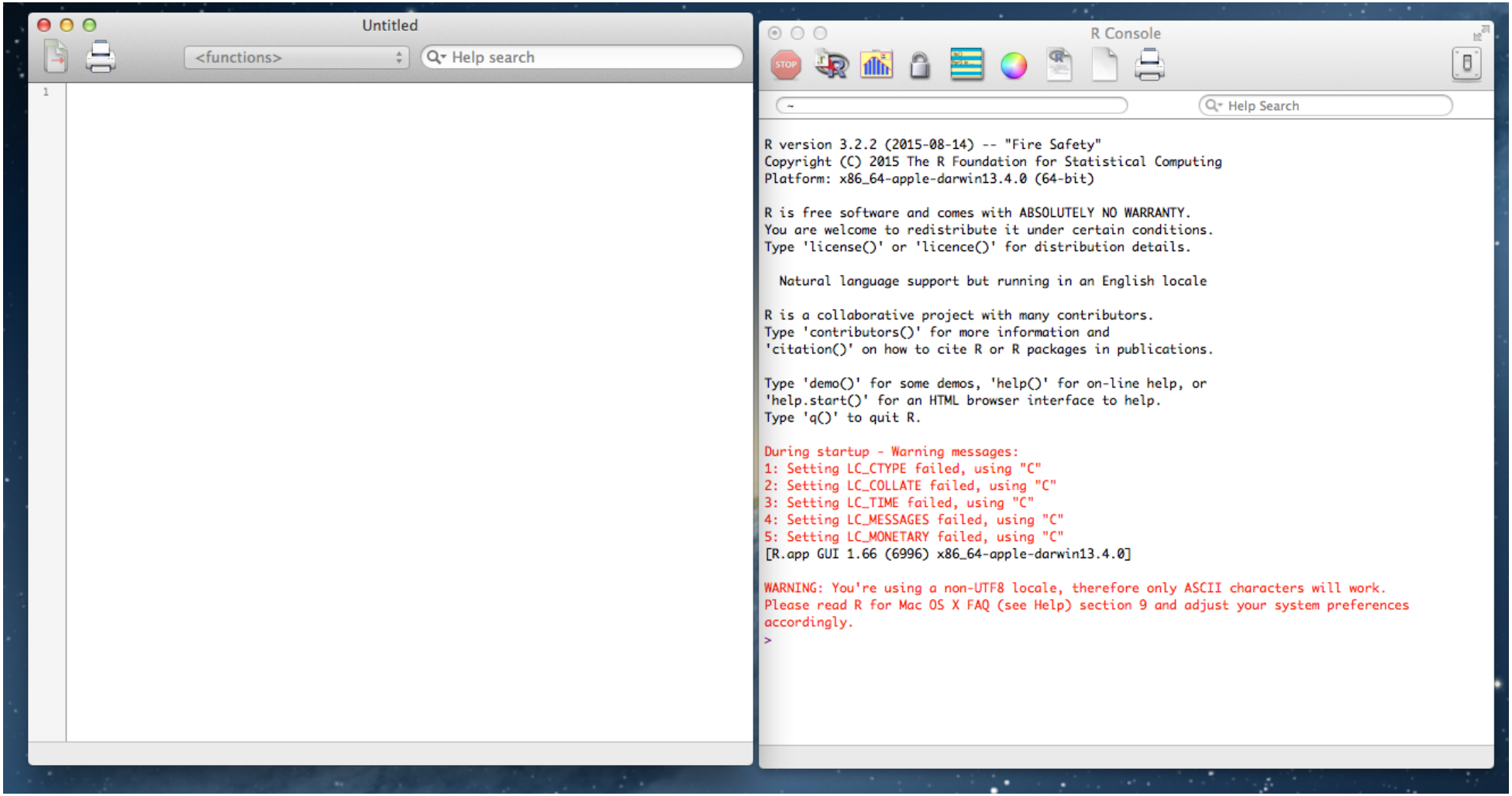
5 As for the assignments, you can use either **STATA** or **R**. In both cases, a well commented script has to be send to me as well.

6 The 1^{st} Exam will take place on 28.2.2016, 12:00-13:15 (75 min.)

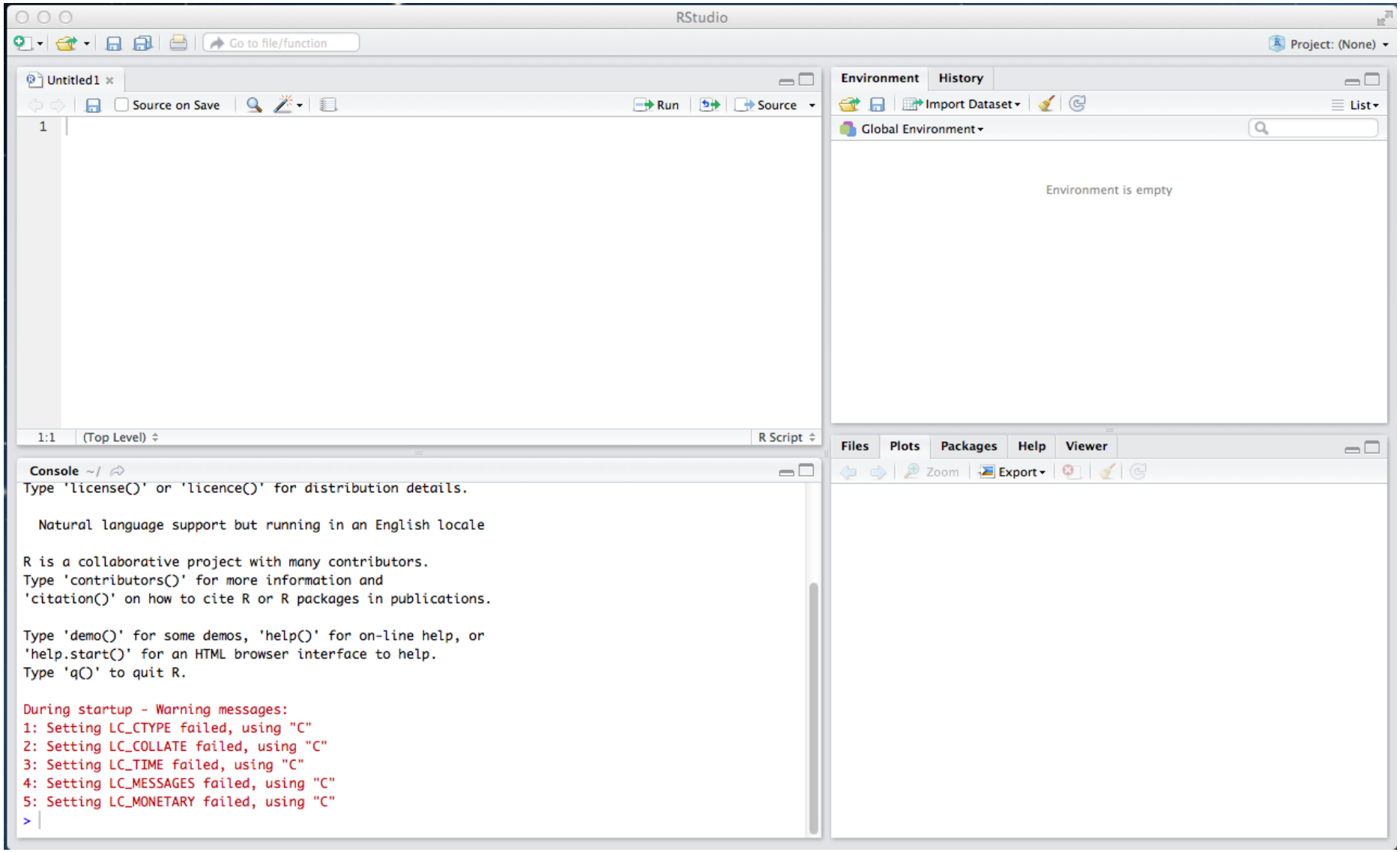
7 The make-up Exam will take place presumably in the 1^{st} week of April. Only people that are going to be sick on the 28.2. or who will fail the 1^{st} Exam will be eligible for it.

Before we can start

- Both R and R-Studio are free-ware, so you can download them for free.
- Download it from *https://cran.r-project.org*.
- If you wish, download R-Studio from *https://www.rstudio.com/products/rstudio/download/*.
- R-Studio is a graphical overlay (user interface) that makes R look and feel more user-friendly. You don't have to use it.
- The code will always look the same. Whether you use it is up to you.



Screenshot of R: new script and R console on the left and on the right, respectively



Screenshot of R-Studio

How does it work, in principle?

- In R everything is an object. Even the dataset is only a Matrix (or data-frame), which you need to define. *Upside:* You can work with several datasets at once.
- There are fewer built-in commands in comparison to STATA. Most commands are based on packages. *Downside:* It is less user friendly than STATA and you are confronted with suboptimal help-files.
- For help use google whenever possible.

Bottom line:

You need to find out what you want to do, google how to do it, download the respective package, use it's help-file to find out how it works.

Basic Commands: Help

- If you know the command you are looking for type *?* and the word to be searched for
- By typing *help.search()* - you can specify your search further...
- Personally, I find the built-in help-files in R abysmal. Don't despair and use google whenever stumbling upon a problem.
- Alternatively you can type in *help.start()* and you can "click" your way to the desired package or *help(package=foreign)*.
- E.g. you want to load STATA file and you don't know how. First google: "how to read .dta file in R". You will find out that the package "foreign" might be helpful with your task. Then you can google "help file to foreign".
- **Anyway, using google is a dominant strategy!**

Basic Commands I

- E.g.: to clean (nearly) everything you can type *remove(list=ls())*
- You see, in R you have a command (remove) which is a function of arguments (list), which needs to be specified.
- As with STATA, not all arguments have to be specified. When some arguments are missing, R will assume their default values. You have to check the respective help file to find out more.
- As mentioned, everything in R are just "objects".
- Start R and try 5+6 ; a=5+6 ; b=c(3,4,5,6); d=seq(from=1,to=9,by=1) ; d1=seq(9) ; d2=seq(from=1,to=9) What do you see?

Basic Commands II

- The most common objects are: Scalars, Vectors, Matrices, and Strings.
- Try *is.vector(x= object)* or *is.matrix(x= object)* to find out about these objects.
- You can force R to change the format with *as.matrix(object)* or doing the following: *dim(d2)<-c(3,3)*.

Basic Commands III

Functions:

Sometimes you want to find out the mean, standard deviation or a sum of a vector. You will need specific functions for each task. *mean(x=object), sd(object), sum(object)*

Example:

Try the following and find out what it does:

```
x=rnorm(100, mean=10, sd=3.4)
x<-as.matrix(x)
colnames(x)<-"income"
x<-data.frame(x)
```

Basic Commands IV: Loading data set

- We have to create an object that contains the data set we want to analyze.
- This specific object is called data.frame (:= a Matrix that contains the names of variables).
- Since we have the STATA .dta file, we need to be able to read it first.
- Install and load the respective package *install.packages("foreing"), library(foreign)*

How to:

```
data<-data.frame(read.dta(file="nlswork_adj.dta")
```

Homework

- Read and solve all the tasks in the "A (very) short introduction to R" that I have uploaded to ISIS.
- Send me your solution on Monday before the Tutorials latest.