

## **JPM 681 Applied Political Research: Intermediate Statistics**

*Summer Semester 2016*

*Institute of Political Studies, Department of International Relations, Prague*

### **PRACTICAL INFORMATION**

Meeting time: Wednesday 11:00 – 12.20 pm  
Meeting room: 4019  
Instructor: Dr. Magda Giurcanu  
Office: 3088  
Email: magda.giurcanu@fsv.cuni.cz  
Office Hours: Wednesday 14:00-16:00 or by appointment  
Course website: <http://dl1.cuni.cz/course/view.php?id=4557>

### **COURSE DESCRIPTION**

This course is designed for advanced undergraduate and graduate students in social science who have a very basic knowledge of statistics. If you have been previously taken an introductory course in statistics, you will have no difficulty following the materials in our course. If you are currently taking a course in research methods, you may have already encountered all the concepts we cover here. I assume, for instance, that you are already familiar with the mean, standard deviation, standard errors, confidence intervals and hypothesis testing, concepts that we will frequently refer to in this course.

This course will be particularly useful to anyone who needs to learn about multiple regression quickly, especially from the perspective of reading and comprehending published research. While you will run regressions during homework, exams, and the research project, I am not expecting you to become statistical researchers. Therefore, we will use as little mathematics as possible, although you may see an occasional equation now and then.

An important part of the course will focus on how to make use of the R statistical package. R has become important in political research because it is free and easily incorporates user-written packages. This program is also widely used in the business community and governmental work, so training as an R user has become a marketable skill. The examples used in this course are however drawn from Political Science, although the techniques used can be applied in other disciplines. In this course, I am assuming no or very little experience with R.

### **COURSE REQUIREMENTS**

This course is an advanced undergraduate/master level and all students are expected to have done the required reading before each seminar and come to class to actively engage in the issues presented. Participating in class also demonstrates that you have done the readings. For these reasons, a grade for classroom participation will be given. One absence during the semester upon prior notification (at least 24 hours ahead) via email is granted.

The seminars will have 2 parts:

- a 30 min review of the issues covered in the homework submitted for the current week
- a 50 min lecture component summarizing the key theoretical and conceptual insights relevant to the topic of the week

## Course materials

We will be reading extensively from the following:

1. Paul Allison. 1999. *Multiple Regression: A Primer*, Pine Forge Press
2. Jeffrey Wooldridge. 2009. *Introductory Econometrics. A Modern Approach*, 4<sup>th</sup> Edition.

For learning R

1. James Monogan III. 2015. *Political Analysis Using R* Springer International Publishing, Switzerland
2. Philip H. Pollock III. 2014. *An R Companion to Political Analysis*. CQ Press

## GRADING POLICIES

Grading is based on:

- A. Classroom participation (10%)—come to class prepared to actively contribute to the discussion
- B. 6 homework sets (40%)—you will have homework every week but you can choose to submit only 6 sets of problems. These will involve both theoretical and computer exercises covering past lectures. It is advised that you work in groups when solving these problems but when you submit these exercises make sure you use your own phrasing rather than simply copying from your colleagues. Basically the more effort you put into these homework sets the better off you'll be at the end of the course.
- C. A replication project that will involve several distinct steps (different deadlines scattered through the syllabus): (30%) (more instructions to follow)
  - a. Find an article to replicate that uses multiple regression and locate the data set that goes with the article. Most authors archive their data and syntax files on the Harvard Dataserve but if the article you decided to work on is not there you have a couple of options—
    - i. Reconstruct the data set following the information in the article—usually very difficult, especially for replication projects
    - ii. Contact the author to see if he/she is interested in sharing the data and the script/do files
    - iii. Or move to find another project.
  - b. Extend the analyses of the article – add your own contribution by including new indicators, recode some variables, or other interventions that you can think of.
  - c. Present the research question, the analysis, and your extension in class during the last 2 weeks of class.
- D. Write-up a final report with the main research question, findings, replication, and extension (to be submitted during the final evaluation period) (20%) Due TBA (final examination period)

**\*\*\*As a starting point in finding 'easy methods papers', you can look at the Political Science papers mentioned on Michael Martinez' page: "Some Political Science Published Works that Use Simple Statistical Procedures"**

**<http://users.clas.ufl.edu/martinez/simplestats.htm>**

For citations when working on the replication project, please use the 'Harvard style' information available at: <http://libweb.anglia.ac.uk/referencing/referencing.htm>

For additional information on citation and referencing please see the short guide for “Acknowledging, Paraphrasing, and Quoting Sources”, available at:  
[http://writing.wisc.edu/Handbook/Acknowledging\\_Sources.pdf](http://writing.wisc.edu/Handbook/Acknowledging_Sources.pdf)

I will apply the following indicative grading scheme:

- 100-90: (A) Mark 1
- 89-70: (B) Mark 2
- 69-50: (C) Mark 3
- < 50: (fail, F) Mark 4

### POLICIES

**Assignments:** All assignments must be submitted on line, **by the deadline**, and handed in to the instructor, at the beginning of the class. Make up exams and late final papers will not be accepted unless there are serious legitimate reasons. Provision of a signed medical note is required, and notice must be given prior to the deadline.

**Academic honesty policies:** The standard plagiarism and academic integrity rules apply, i.e. all the materials you submit in paper or online must be the results of your own individual work. Any signs of plagiarism will be taken very seriously. You do not submit a paper for this course, but make sure that you abide by the academic integrity rules also in the shorter pieces of text you will submit for your homework assignments. Please consult the Faculty policies on plagiarism (see <http://intranet.fsv.cuni.cz/FSVINT-637.html>, only in Czech) or have a look here: [https://writing.wisc.edu/Handbook/Acknowledging\\_Sources.pdf](https://writing.wisc.edu/Handbook/Acknowledging_Sources.pdf)

**Courtesy:** Don't come late and turn off your cell-phones. Any disruptive behavior (reading newspapers or materials related to other courses, talking outside of class discussions) will not be tolerated and you will be asked to leave the class.

### **COURSE SCHEDULE\***

*\*As the semester unfolds we may have to alter the schedule based on how fast or slow we cover the material. If this happens, I will let you know in advance (1 week) what changes need to be incorporated.*

### **WEEK 1 (Feb 17) Introduction to the course and its policies**

Topics:

- Policies and expectations
- What do you expect to get out of this class?
- What type of exposure you had to statistical methods?
- What type of data do we have in political science

## **WEEK 2 (Feb 24) Multiple regression in applied research – an overview**

### Topics:

- What is multiple regression?
- How do I interpret multiple regression results?
- R: How to install R or R Studio (covered via tutorials online)

### Readings:

1. Jeffrey Wooldridge. Introductory Econometrics—Chapter 1
2. Paul Allison, Multiple Regression: A Primer, Chapters 1,2
3. James Monogan III. 2015. *Political Analysis Using R* Springer International Publishing, Switzerland, Chapter 1

## **WEEK 3 (March 2) What can Go Wrong with Multiple Regression?**

### Topics:

- Issues to think about: variables left out, directionality of the effects, measurements of independent variables, biased samples, highly correlated variables
- R: Loading and Manipulating Data in R (covered online via tutorials); in class we will cover issues such as merging data sets, reshaping data.

### Readings:

1. Paul Allison, Multiple Regression: A Primer – Chapter 3
2. James Monogan III. 2015. *Political Analysis Using R* Springer International Publishing, Switzerland—Chapter 2
3. To understand the data set used in the R session take a look at the following: Steven C. Poe and C. Neal Tate 1994. Repression of Human Rights to Personal Integrity in the 1980s: A Global Analysis, *The American Political Science Review*, Vol. 88, No. 4 (Dec., 1994), pp. 853-872.

Homework 1 announced — due Tuesday 5pm the following week via email and a hard copy in class (on Wednesday).

## **WEEK 4 (March 9) The Simple Regression Model 1**

### Topics:

- How Does Bivariate Regression Work?
- What are the properties of OLS on Any Sample Data
- R: Descriptive Statistics—frequency distributions, bar charts, ordinal, nominal, interval data

### Readings:

1. Paul Allison, Multiple Regression: A Primer—Chapter 5
2. Jeffrey Wooldridge. Introductory Econometrics—Chapter 2
3. Phillip Pollock III. 2014—Chapter 2 on Descriptive Statistics

Homework 2 announced — due Tuesday 5pm the following week via email and a hard copy in class (on Wednesday).

**Research Project Step 1 Due in class:** Find an article and have it approved by the instructor

## **WEEK 5 (March 16) The Simple Regression Model 2**

### Topics:

- What are the assumptions of multiple regression
- Expected Values and Variances of the OLS Estimators
- R: More on Graphs and Plots

### Readings:

1. Paul Allison, *Multiple Regression: A Primer*, Chapter 6
2. Jeffrey Wooldridge. *Introductory Econometrics*—Chapter 2 (same chapter)
3. James Monogan III. 2015. *Political Analysis Using R* Springer International Publishing, Switzerland, Chapter 4 Descriptive Stats
4. To understand the data set used in the R session take a look at JEFFREY S. PEAKE and MATTHEW ESHBAUGH-SOHA. 2008. The Agenda-Setting Impact of Major Presidential TV Addresses, *Political Communication* 25, 113-137.

Homework 3 announced — due Tuesday 5pm the following week via email and a hard copy in class (on Wednesday).

## **WEEK 6 (March 23) Multiple Regression Analysis**

### Topics:

- Motivation for Multiple Regression
- Comparison of Simple and Multiple Regression Estimates
- Regression through the Origin
- R: Transforming variables

### Readings:

1. Jeffrey Wooldridge. *Introductory Econometrics*—Chapter 3
2. Phillip Pollock III. 2014—Chapter 3 Transforming Variables and Chapter 4 Making Comparisons

Homework 4 announced — due Tuesday 5pm the following week via email and a hard copy in class (on Wednesday).

**Research Project Step 2 Due in class:** Confirm access the data set. Submit descriptive stats of your variables

## **WEEK 7 (March 30) Multiple Regression Inference**

### Topics:

- Hypothesis Testing
- One-Sided Alternatives, Two-Sided Alternatives
- Confidence Intervals
- Reporting Regression Results
- R: Inferences about Sample Means

### Readings:

1. Jeffrey Wooldridge. *Introductory Econometrics*—Chapter 4
2. Phillip Pollock III. 2014—Chapter 6 Making Inferences about Sample Means

Homework 5 announced — due Tuesday 5pm the following week via email and a hard copy in class (on Wednesday).

## **WEEK 8 (April 6) Further Issues in Multiple Regression 1**

Topics:

- What can be done about multicollinearity
- How to interpret interactions?
- Interactions with Dummy Variables
- R: Correlation and Linear Regression 1

Readings:

1. Paul Allison, *Multiple Regression: A Primer*, Chapters 7 & 8
2. James Monogan III. 2015. *Political Analysis Using R* Springer International Publishing, Switzerland, Chapter 5 Bivariate Associations

Homework 6 announced — due Tuesday 5pm the following week via email and a hard copy in class (on Wednesday).

## **WEEK 9 (April 13) Further Issues in Multiple Regression 2**

Topics:

- How can Multiple Regression Handle Nonlinear Relationships
- How to Choose between non-nested models
- Models with Quadratics
- R: Correlation and Linear Regression 2

Readings:

1. Jeffrey Wooldridge. *Introductory Econometrics*—Chapter 6
2. Phillip Pollock III. 2014—Chapter 8 Correlation and Linear Regression

Homework 7 announced — due Tuesday 5pm the following week via email and a hard copy in class (on Wednesday).

## **WEEK 10 (April 20) More on Data Specification issues**

Topics:

- Outliers, Influential Observations, heteroscedasticity
- R: Regression Diagnostics

Readings:

1. Jeffrey Wooldridge. *Introductory Econometrics*—selected pages from Chapters 8 & 9
2. James Monogan III. 2015. *Political Analysis Using R* Springer International Publishing, Switzerland, Chapter 6 Linear Models and Regression Diagnostics

Homework 8 announced — due Tuesday 5pm the following week via email and a hard copy in class (on Wednesday).

**Research Project Step 3 Due in class: First Attempt to Replicate. Submit Regression Analysis**

## **WEEK 11 (April 27) In-class Working Session**

Topics:

- Finishing up issues not covered and mentioned in the syllabus
- In-class work on student projects (bring laptops)

Research Project Step 4 – in class work on either improving your replication by considering regression diagnostics or improve the project (by adding missing variables, rescaling some variables, removing some outliers, etc).

## **WEEK 12 (May 4) & WEEK 13 (May 11) Conclude and Student Presentations** **Final Project Due –During the Examination Weeks –More Details to Follow.**