Quantitative Methods in Political Science

(a.k.a. Multivariate Analyses)

Fall 2014

Course Details:
Lecture: Th, 8h30-10h00 (A5, B143)
Lab: Th, 10h15-11h45 (A5, C108)

Instructors:

Thomas Gschwend
Office: A 5,6 Room C 219
Office Phone: 181-2413
E-mail: gschwend@uni-mannheim.de
Office Hours: Tue, 13h30-15h00

Lukas Stötzer
Office: A 5,6 Room C 221
Office Phone: 181-2542
Email: lstoetze@mail.uni-mannheim.de
Office Hours: Wed, 15h30-17h00

Course Description:

This course introduces graduate students to quantitative methods in political science. During the first half of the course, we will focus on linear regression models. The topics covered include discussions of the mathematical bases for such models, their estimation and interpretation, model assumptions and techniques for addressing violations of those assumptions, and topics related to model specification and functional forms. During the second half of the course, students will be introduced to likelihood as a theory of inference, including models for binary and count data.

The main goals of this course are to develop sound critical judgment about quantitative studies of political problems, to interpret quantitative analyses in published work, to understand the logic of statistical inference and to recognize and understand basic linear regression models. It provides the skills necessary to conduct your own quantitative analyses and teaches how to do so using R. This class lays the foundation for "Advanced Quantitative Methods" to be taught in spring 2014.

The course language is English. This includes lecture, discussions and assignments.

Recommended for:

Political science graduate students, including M.A. and PhD students (CDSS).

Course Registration:

The course is divided into a lecture-style seminar (Multivariate Analyses) and a computer lab session (Tutorial Multivariate Analyses). During the computer lab session, students

will apply the statistical models introduced in the lecture. Students who wish to take the course should register for "Multivariate Analyses" at the student portal.

Note that this course is highly demanding and entails a substantial work load for students in the form of weekly homework assignments, a mid-term exam and data essays. Students who wish to audit this class should notify the instructors in advance (participation is subject to free room capacity). Please note that only registered students will receive feedback on their written work.

Readings:

We will not use a single textbook for this course. Selected readings are available on the course website. The following books will be used in the course:

Fox, John. 2008. *Applied Regression Analysis and Generalized Linear Models.* 2nd edition. Sage.

King, Gary. 1998. *Unifying Political Methodology.* Ann Arbor: University of Michigan Press.

Wooldridge, Jeffrey. 2009. *Introductory Econometrics: A Modern Approach.* 4th edition. South-Western College Pub.

Matloff, Norman. 2011. *The Art of R Programming.* William Pollock.

Software:

For all calculations, we will primarily support and use the open-source statistical programming language R. It is particularly suited for carrying out state-of-the-art computer-based simulations and data exercises. It also generates really nice publication-quality graphics and runs under a wide array of operating systems. R can be downloaded for free at http://www.r-project.org/. Learning R might seem a bit challenging at first, but you will realize that it is incredibly powerful. A readable introduction is given by Fox, John. 2002. *An R and S-Plus companion to Applied Regression.* Sage.

A descend graphical user interface for R (which we will also use during the lab sessions) is RStudio. In recent years a growing number of features have been added to this graphical user interface, which makes it - especially for Beginners - the preferred choice to learn R. It is cross-platform and open-source. RStudio can be downloaded for free at http://www.rstudio.com/.

The lab sessions will be devoted to learning the various commands in R and apply the statistical models from the lecture to selected political science data sets. The data sets that we will use cover the major fields in political science.

Prerequisites:

There is formally no prerequisite for this course except an open mind and a good command of high school algebra.

Course Requirements:

You will receive a grade for the seminar "Multivariate Analyses" and a pass/fail for the lab session "Tutorial Multivariate Analyses". Grading will be based on the following components:

- **Homework Assignments (1/3)**

  The homework assignments will take the form of problem sets, replications, simulations, or extensions of the analysis in class and the lab. The assignments will be handed out at the end of class and you are expected to hand in the solution at the beginning of the next meeting (unless noted otherwise). Late assignment will not be accepted, unless you provide a doctor's note that shows that you were unable to work for the whole week.

  Homework need to be handed in groups of 2-3 students. Please find one or two coauthors with whom you will work on the weekly homework and indicate the group members on your answer sheets. Group members will receive the same grade. Throughout the years, we learned that the only reasonable way to manage the work-load in the first semester is to work together. Group work saves you time, as not everyone has to type up his or her own answers. Although who is contributing how much to one's homework is completely up to you, you only learn if you try to answer the problem sheets by yourself. From our experience, not getting strongly involved in each of the weekly homework, lowers group work quality and increases the risk of failing both, the mid-term and data analysis project.

  Moreover, you are strongly encouraged to seek advice from the instructors during office hours or through the online forum. Homework assignments are graded with a $\sqrt{}+$ (excellent work), $\sqrt{}$ (good work, but some minor mistakes), and $\sqrt{}-$ (significant improvements are necessary).

- **Midterm Exam (1/3)**

  The midterm will be a 90-min closed-book exam that covers the first half of the course materials.

- **Data Analysis Project (1/3)**

  Towards the end of the course, you will be handed out a data analysis project (including a dataset). The project will involve the creative application of the statistical techniques to a substantive problem in political science. Your paper should have 2000 words ($\pm$ 10%, without bibliography). The essay will be marked down if you go below or above the page range. The data essay should emphasize the substantive, statistical, and causal significance of your analysis and the write-up should read very much like the results section of a published article. No collaboration is permitted on the data analysis project. You are welcome to seek advice from the instructors during office hours. Details on the data analysis project will be provided at a later stage.

Other Considerations:

A great website with many R code examples is the UCLA Stat Consulting Site. Another good site that introduces R to SPSS or Stata users is Quick-R. The standard site to search for R on the net is R Seek

Learn to use LaTeX while you can. It is a free typesetting software package and enables you to typeset and print your work at the highest typographical quality, using a prede-fined, professional layout. The main advantages of LaTeX over normal word processors include professionally crafted layouts, support for typesetting of mathematical formulae in a convenient way, a few easy-to-understand commands that specify the logical structure of a document, more complex structures such as footnotes, references, table of contents, bibliographies that can all be generated easily, and free add-on packages for specific tasks (e.g. make a reference list adhere to the exact standards of a scientific journal). A short introduction can be found here. We further intend to offer a short Intro during the first weeks.

Another useful software tool that helps you to type up your answers more easily is *Knitr*. It allows you to integrate R-Code into LaTeX files. Take a look at the homepage, it is very simple and straightforwardly implemented in RStudio. We hope to find a time to provide you with a Intro to *Knitr*.

Course Outline:

Week 1 (4 September 2014): No Class.

Please download R as well as RStudio and install them onto your system.

Week 2 (11 September 2014): Introduction. Visualizing Data.

Fox, John. 2008. *Applied Regression Analysis*. Chapters 1 and 3.

Week 3 (18 September 2014): Fundamentals of Probability.

Wooldridge, Jeffrey. 2009. *Introductory Econometrics*. Appendix B.

Fox, John. 2008. *Applied Regression Analysis*. Appendix D.

Week 4 (25 September 2014): Sampling and Statistical Inference.

Wooldridge, Jeffrey. 2009. *Introductory Econometrics*. Appendix C.

Week 5 (2 October 2014): Linear Regression.

Kennedy, Peter. 2008. *A Guide to Econometrics*. Chapter 3.

Fox, John. 2008. *Applied Regression Analysis*. Chapter 5.

Week 6 (9 October 2014): Linear Regression: Dummies and Interactions.

Fox, John. 2008. *Applied Regression Analysis*. Chapter 7.

Thomas Brambor, William Roberts Clark, and Matt Golder. 2006. "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis* 14: 63-82.

Week 7 (16 October 2014): Linear Regression: Inference & Hypothesis Tests.

Fox, John. 2008. *Applied Regression Analysis*. Chapter 6.

Week 8 (23 October 2014): Linear Regression: Interpreting Substantive Effects via the Simulation Method.

Gary King, Michael Tomz, and Jason Wittenberg. 2000. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation" *American Journal of Political Science* 44(2): 341-355

Week 9 (30 October 2014): Linear Regression: Diagnostics.

Fox, John. 2008. *Applied Regression Analysis*. Chapter 11-13.

Week 10 (6 November 2014): Midterm Exam

Good Luck!

Week 11 (13 November 2014): Non-linear probability models - The likelihood theory of statistical inference.

King, Gary. 1998. *Unifying Political Methodology*. Chapters 1, 2, 4.

Week 12 (20 November 2014): Binary Data

Fox, John. 2008. *Applied Regression Analysis*. Chapter 14.

Week 13 (27 November 2014): Count data

Fox, John. 2008. *Applied Regression Analysis*. Chapter 15.

Week 14 (4 December 2014): Semester Wrap-up