Syllabus for *Quantitative Political Analysis*

David Hendry & Thomas J. Leeper
Department of Political Science and Government
Aarhus University

November 21, 2014

Individuals working in the public and private sectors are often tasked with analyzing quantitative data or making use of analyses performed by others. The purpose of this course is to expand and improve significantly participants' ability to perform quantitative analyses of political science data and, further, to better evaluate the use of quantitative research results. Several advanced statistical and analytic techniques will be introduced and applied to research questions from political science. Participants will leave the course with the ability to better assess published research, better perform their own analyses of quantitative data, and better describe and understand the results of such analyses.

Among others, the course will discuss the following techniques and topics:

- Linear regression analysis
- Interpretation of interactions in regression models
- Regression with binary, categorical, and count outcomes
- Analysis of data gathered over time and across geographies
- Research design for causal inference
- Using Stata for statistical analysis

In addition, topics such as hypothesis generation and research design will be discussed throughout the course. The course does not have a unified theoretical or empirical focus, but we will touch upon and read empirical literature from most areas of political science including comparative politics, public administration, and international relations.

Each week, the course will consist of two sessions:

- Lecture
- Hands-on lab session

1 Objectives

The learning objectives for the course are as follows. By the end of the course, students should be able to:

1. Describe politically relevant research questions and hypotheses
2. Evaluate and deduce observable implications from political science theories
3. Explain statistical procedures and their appropriate usages
4. Apply statistical procedures to relevant research problems
5. Synthesize results from statistical analyses into well-written and well-structured essays
6. Demonstrate how to use Stata for statistical analysis

2 Exam and Weekly Assignments

2.1 Exam (Only for MA/Kandidater)

The final exam is a seven-day written assignment analyzing a topic outlined by the instructors using quantitative data. In addition, the students must write four essays during the course on topics defined by the instructors as a prerequisite for the exam
The essays (3-5 pages) must be written in English and are due via email to the appropriate instructor at 12:00 on the dates listed below. The students will receive feedback on the essays and they must all be approved before taking the final exam.

1. Essay 1 due February 27 to David
2. Essay 2 due March 20 to Thomas
3. Essay 3 due April 10 to David
4. Essay 4 due May 8 to Thomas

2.2 For PhD Students

Students can choose from the following course elements:

1. "Regression with Matrices and MLE" (1 essay, 2 ECTS)
2. Sessions 1–5 (1 essay, 7 ECTS)
3. Sessions 1–5, plus one additional session (2 essays, 9 ECTS)
4. Sessions 1–5, plus two additional sessions (3 essays, 11 ECTS)
5. Sessions 1–5, plus three additional sessions (3 essays, 13 ECTS)
6. Entire course (4 essays, 15 ECTS)

PhD students must notify the course instructors of their choice by the end of Session 5.

3 Reading Material

The assigned material for the course includes about 2400 pages including several textbooks and empirical research articles, all of which are available online or in the printed course packet. All readings should be completed for the day they are described. *There is reading assigned on the first day.* The textbooks for the course are as follows:

– .
– .
– .
– .
– .
– .

Note: The Sønderskov text is also available in English from the publisher.
PhD students are additionally required to obtain:
– .

4 Course Website

All information about the course will be posted on Blackboard. Any changes to the syllabus or additional notes will be made available there.

5 Schedule

The general schedule for the course is as follows. Details on topics covered and the readings for each week are provided on the following pages.

5.1 Introduction and Research Design (Week 6)
5.2 Research Design II (Week 7)
5.3 Ordinary Least Squares Regression (Week 8)
5.4 Ordinary Least Squares Regression II (Week 9)
5.5 Practical Data Issues (Week 10)
5.6 Research Designs for Causal Inference (Week 11)
5.7 Panel Analysis for Continuous Outcomes (Week 12)
5.8 Multi-level Modeling (Week 13)
5.9 Maximum Likelihood Estimation (Week 14)
5.10 Interpretation of GLMs (Week 15)
5.11 GLMs for Ordered, Multinomial, and Count Outcomes (Week 17)
5.12 Survival and Duration Analysis (Week 18)
5.13 Panel Analysis for Discrete Outcomes (Week 19)
5.14 Conclusion and Wrap-up (Week 20)

5.1 Introduction and Research Design (Week 6)

What topics will we cover in this course? How do we think about causality for the purposes of research design? How can experiments help us understand causal relationships?

Instructor: Thomas

Lecture

- Course overview
  - Readings and textbooks
  - Exam
  - Four essays
  - Plan for the course
- Asking good research questions
- Research design
- Philosophies of causality
- Experiments and matching

Readings

– Angrist and Pischke, Ch.2.

– .

– .

See Also:

Lab

- Basics in Stata

Readings for lab

– Sønderskov, Ch. 1–3.

5.2 Research Design II (Week 7)

motivating puzzles

Instructor: David

Lecture

- Defining causality in a regression framework
- Model-building
- Reporting regression results

Readings

– AngristPishke2009, Ch.3 (up to p.69).

– .

See Also:

Lab

- Further basics in Stata

Readings for lab

– Sønderskov, Ch. 4–6.

5.3 Ordinary Least Squares Regression (Week 8)

How do we estimate causal effects using regression analysis? How do we interpret linear regression coefficients for different types of variables? What do goodness-of-fit measures tell us?

Instructor: Thomas

Lecture

- OLS method
- Interpretation of coefficients
- Standard errors, t-tests, and p-values
- Goodness-of-fit measures

Readings

– Berry1993, pp.1–67.

– AngristPishke2009, Ch.3 (from p.69).

– CameronTrivedi, .

– .

See Also:

Lab

- OLS in Stata
- Reporting regression results
- Stata factor variables

Readings for lab

– Sønderskov, Ch. 8–9.

5.4 Ordinary Least Squares Regression II (Week 9)

How do state and test hypotheses about heterogeneous effects? How do we interpret those effects in OLS interaction terms? When do we need to estimate alternative standard errors for OLS estimates?

Instructor: David

Lecture

- Effect heterogeneity and interaction terms
- Standard errors
- Heteroskedasticity

Readings

– Angrist and Pischke, 293–315.

– Berry1993, p.67 to end.

– .

See Also:

Lab

- Estimating and interpreting interaction terms
- Heteroskedasticity-consistent standard errors
- Clustered standard errors
- The `margins` command

Readings for lab

– Sønderskov, Ch. 10.

5.5 Practical Data Issues (Week 10)

Instructor: Thomas

Lecture

- Multivariate scaling and reliability
- Variable transformations
- Missing data, case deletion, and imputation

Readings

– Excerpt from .

See Also:

Lab

- Scaling and reliability
- Variable transformations and regression interpretation
- Missing data imputation

Readings for lab

5.6 Research Designs for Causal Inference (Week 11)

Instructor: Thomas

Lecture

- Instrumental Variables
- Regression Discontinuity Designs
- Difference-in-Differences

Readings

– Ch.4,6.

– .

See Also:

Lab

- Regression Discontinuity Design analysis

Readings for lab

5.7 Panel Analysis for Continuous Outcomes (Week 12)

Instructor: David

Lecture

- Difference-in-differences, continued
- First-differences and fixed effects
- Random effects models

Readings

– AngristPischke2009, Ch.5.

– Allison2009, Ch.1,2.

See Also:

Lab

- Data in "wide" and "long" formats
- Panel regression

Readings for lab

– CameronTrivedi, sections.

5.8 Multi-level Modeling (Week 13)

Instructor: David

Lecture

- Multi-level data structures
- Logic of multi-level models
- Hierarchical linear model
- Random effects versus fixed effects

Readings

See Also:

Lab

- Estimation
- Testing for fixed versus random effects
- Interpretation

Readings for lab

5.9 Maximum Likelihood Estimation (Week 14)

Instructor: David

Lecture

- Maximum Likelihood Estimation
- Generalized Linear Models
- Logistic regression

Readings

– Long1997, Ch.1–3.

See Also:

Lab

- Interpretation of logistic regression coefficients
- Interpret Wald tests and likelihood ratio tests

Readings for lab

5.10 Interpretation of GLMs (Week 15)

Instructor: Thomas

Lecture

- Logit versus Probit
- Heterogeneous effects and interaction terms
- Predicted probabilities and marginal effects
- Interpretation in Generalized Linear Models

Readings

See Also:

Lab

- `margins` for generalized linear models
- `marginsplot`

Readings for lab

No class (Week 16)

5.11 GLMs for Ordered, Multinomial, and Count Outcomes (Week 17)

Instructor: David

Lecture

- Ordered logit and probit
- Multinomial logit
- Count outcomes
  - Poisson regression
  - Dispersion and alternative count models

Readings

– Long1997, Ch.5,6,8.

See Also:

Lab

- Estimation of ordered, multinomial, and count models
- Interpretation of results
- Presentation of results
- Testing for overdispersion

Readings for lab

5.12 Survival and Duration Analysis (Week 18)

Instructor: David

Lecture

-

Readings

See Also:

Lab

- Estimation of different models
- Model testing and comparison
- Interpretation

Readings for lab

5.13 Panel Analysis for Discrete Outcomes (Week 19)

Instructor: Thomas

Lecture

- Panel binary outcome models
- Panel count models

Readings

– Allison2009, Ch.3,4.

See Also:

Lab for lab

- Estimating these
- Interpretation

Readings

5.14 Conclusion and Wrap-up (Week 20)

What have we learned? What didn't we learn?

Instructor: David

Instructor: Thomas

Lecture

- Wrap-up
- Course evaluations
- Questions about the exam

Readings

See Also: