POLS 642 INTERMEDIATE ANALYSIS OF POLITICAL DATA (Spring 2017)

Hye-Sung Kim
Office: ZH 416
Office Hours: TH 2:00 – 3:00 pm
Email: hkim4@niu.edu

Course Information
Classroom: DuSable 246
Class Time: TH 11:00 am -1:40 pm

COURSE DESCRIPTION

This course is an introduction to the theory and application of linear modeling to social science problems. The focus of this class will be to provide you with the theoretical and practical skills necessary to conduct your own empirical research. Topics will include ordinary least squares, hypothesis testing, dealing with violations of the underlying assumptions of multiple regression, instrumental variables estimation, simultaneous equations, time-series econometrics, panel data techniques and causal inference, experiments, quasi-experiments and maximum likelihood estimation.

By the end of the course, students should be able to:

- Identify the most appropriate methodological techniques for analysis given a research question and available data, as well as identify, understand the implications, and offer resolution to various problems encountered during quantitative analysis.
- Conduct data analyses using the methodologies covered in the course. In particular, students should be able to diagnose and test empirical models, and apply the techniques for correcting models that violate statistical assumptions.
- Manage data and conduct analyses using STATA.

REQUIRED TEXTS

- Bailey, M. (2015). *Real Stats: Using Econometrics for Political Science and Public Policy*. Oxford University Press.
- Fox, J. (2015). *Applied regression analysis and generalized linear models*. Sage Publications.

RECOMMENDED TEXTS

- Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach*. Nelson Education.
- Cameron, A. C., & Trivedi, P. K. (2009). *Microeconometrics using stata* (Vol. 5). College Station, TX: Stata press.
- Fox, J. (2015). *Applied regression analysis and generalized linear models*. Sage Publications.
- Kennedy, P. (2008). *A guide to econometrics*. Wiley-Blackwell.
- Greene, W. (2010) *Econometric Analysis.* Pearson.

- Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

STATA

You will need access to STATA. There are 3 ways to obtain it:

1. Purchase it. You can purchase Stata at a discounted price through their "Grad Plan": http://www.stata.com/order/new/edu/gradplans/. Recommended: Stata/IC or Stata/SE 13.

2. Use it in the main SOCQRL (in 222 DuSable; www.socqrl.niu.edu) and in the POLS GA Office located in 476 DuSable. Course fees enable you to use these facilities and there are a number of lab assistants available to answer questions. No one should feel intimidated by these assignments. No prior knowledge of statistics is required, the software is user friendly, and experienced lab assistants will be available (at the SOCQRL) to help with the assignments when needed.

3. To access STATA in the classroom:

1. Click on the Windows Start button.
2. Type "Run" into the search box.
3. The search function will locate the Run command icon. When it does, click on it.
4. When the Run command is visible, type in: \\socqrl\apps
5. If the SOCQRL recognizes your login information, you will gain access to the server with all its applications.
6. Find the STATA version that is appropriate for your computer - 32 byte system so I had to find the 32 byte or 64 byte STATA.
7. Click on the .exe file within STATA

COURE REQUIREMENTS

Exam (30%)

There will be a take-home exam worth 30 % of your course grade handed out on **May 4**. It will be due on **May 7 by 11:59pm**. It will be cumulative and open-book. Unless you receive prior authorization from the instructor in writing, late exams will be penalized two full letter grades for each day-or fraction thereof-that they are late.

Weekly Problem Sets (30%)

You will be assigned 10-12 problem sets, which will count for 30% of the course grade. You are encouraged to work in groups (composed of no more than 3 students). The problem sets handed in on time will be graded on a 3 point scale (check plus, check and check minus). Since a solution set will

be available on the course website as soon as a problem set is handed in, late problem sets will not be accepted (which means you will receive a zero).

Original Research Paper (35 %)

This paper will count for 35 % of the course grade. Conduct empirical research on the question of your choice using a dataset of your choice. A research proposal that includes the dataset, the research question and details about the key variables of interest is due on February 23. The research paper is due on May 4. This paper will include a brief introduction to the research question, a literature review, a description of the data and measures, including descriptive statistics, an explanation of the empirical methodology, results, and the discussion/interpretation of the results. The style of the paper should be similar to that of published journal articles. Unless you receive prior authorization from the instructor in writing, late papers will be penalized two full letter grades for each day---or fraction thereof-that they are late.

Presentation (5 %)

This presentation will count for 5 % of the course grade. On May 4, all students will present their original research papers to the class. Each student will be allotted 12-15 minutes for presenting and each presentation will be followed by 3-5 minutes' worth of questions from the class. Each student will present the following about their assigned paper: the research question, the data used, the empirical methodology, and the findings. The presentation will include an interpretation of the findings and how these results answer the original research question. If your paper is not turned in by the beginning of the day of presentations, you will not be able to present and you will receive a zero for the presentation. No exceptions.

GRADING

Course Grades will be distributed as follows:

96 or above	A
93-95	A-
87-89	B+
83-86	B
80-82	B-
77-79	C+
73-76	C
70-72	C-
67-69	D+
63-66	D
60-62	D-
Below 60	F

ACADEMIC INTEGRITY

All students are responsible for maintaining the highest standards of honesty and integrity in every phase of their academic careers. The penalties for academic dishonesty are severe and ignorance is not an acceptable defense.

COURSE TOPICS

Note: The course syllabus is a general plan for the course; deviations announced to the class by the instructor may be necessary. We may not cover all of these topics. Conversely, time permitting, other topics might be covered in this course.

January 19	Week 1 Introduction and review of simple regression Bailey, Chapter 3 • OLS estimator as random variables • Six assumptions of the classical linear regression model • The Gauss Markov Theorem • Unbiasedness; Precision of Estimates; Consistency • The R-squared goodness of fit measure
January 26	Week 2 Multiple linear regression Bailey, Chapter 5 • OLS and goodness of fit in the K-variable model • The classical assumptions revisited • Endogeneity • Precision and goodness of fit
February 2 and 9	Week 3 – 4 Model specification and functional form Bailey, Chapters 5.5; 6; 7 • Model specification • Functional transformations of dependent and independent variables • Functional form test • Dichotomous (dummy) independent variables • Interaction variables
February 16	Week 5 Interval estimation and hypothesis testing Bailey, Chapter 4 • Common families of statistical distribution • OLS under the normality assumption • Confidence intervals/interval estimation • Testing hypothesis about a single parameter; the t-test and statistical significance • Difference between statistical significance and economic (substantive) significance • Testing hypothesis involving several parameters: F-test

February 23	Week 6 Model specification and omitted variable bias Bailey, Chapters 5.2; 5.5; 14.1 through 14.5 • ***Research proposal due*** • Omitted variable bias • Omitted variable bias tests • Including irrelevant variables • Comparing nested and non-nested models • Different effects across groups
March 2	Week 7 Other data problems Bailey, Chapters 5.3; 14.6; Fox, Chapters 6.4;13 • Multicolinearity • Measurement errors
March 9	Week 8 Nonspherical Errors Wooldridge, Chapter 8; Fox, Chapter 16 • Heterocedasticity • Diagnosis and remedy of heteroscedasticity • Autocorrelation • Diagnosis and remedy of autocorrelation
March 16	Week 9 No Class - Spring Recess
March 23	Week 10 Instrumental Variables Estimation and Two Stage Least Squares Bailey, Chapter 10 • Experiments • Instrumental variables • 2SLS
March 30	Week 11 Time-series econometrics Bailey, Chapter 13 • Modeling autocorrelation • Detecting autocorrelation • Fixing autocorrelation • Dynamic models • Stationarity
April 6	Week 12 Panel-data econometrics Bailey, Chapters 8; 15 • Difference-in-difference • Fixed effects models • Random effects models
April 13	Week 13 Regression Discontinuity Bailey, Chapter 11 • Basic RD Model • More Flexible RD Models • Windows and Bins • Limitation and Diagnostics

April 20	Week 14 Logit and Probit Models for Categorical Response Variables: Models for Dichotomous Data Bailey, Chapter 12; Fox, Chapter 14.1 • Using latent variables to explain observed variables • Probit and logit models • Maximum likelihood estimation • Interpreting probit and logit coefficients
April 27	Week 15 Logit and Probit Models for Categorical Response Variables: Models for Polytomous Data Fox, Chapter 14.2 • Multinomial logit and multinomial probit models • Ordered probit and ordered logit models
May 4	Paper presentations • *Research paper due* • Final exam handed out