Objective of the Course and Requirements

This course is designed to help you read, understand, interpret, use and evaluate quantitative empirical work in social science. To advance that goal, attention is concentrated on one of the main techniques used by social scientists and public policy researchers: regression analysis.
You will learn the assumptions that underlie the appropriate use of regression. You will learn how to perform regressions using STATA, perhaps the most widely used computer program in advanced social science research. Most important of all, you will learn to spot violations of the assumptions that give regression results desirable qualities and how to take the corrective measures necessary to improve your ability to make valid inferences.

While regression is scarcely the only statistical approach used by researchers, you will be well prepared to learn many other techniques as you encounter them when combined with the material you learned in Empirical Analysis I. An especially important issue concerns the role of the analyst’s judgment in drawing inferences from the data and, more broadly, just how “scientific” the whole enterprise of data analysis really is.

This course assumes a background in statistics at the level of PA 5031. The course requirements include three problem sets (45 percent of the course grade), a final exam (40 percent) and oral presentations in teams and class participation (15 percent together). The examination will be closed book.

Each of you will be assigned as a member of a team. These teams will be used for two purposes: to discuss course material and to make a classroom presentation along with an associated write-up (more detailed on how this exercise is to be performed is explained on a separate sheet on Moodle). Work with other members of the team (or anyone else) on problem sets should be confined to considering the basic nature of the problem. You are allowed, and even encouraged, to work together in talking through the theory and basic concepts associated with the problem sets. In other words, you can talk through the meaning of a question or confirm your understanding of why a certain test is used and the like. All collaboration should end before any computer work is done. All statistical work (running regressions, etc.) and interpretation of that work (the actual writing up of the problem set) must be done on your own and presented in your own words. Any violation of this rule will be treated as plagiarism.

All problem sets and class problem write ups must be submitted in hard copy and also uploaded to the appropriate folder on the course website. The problem sets are due at class time; the class problems are due by 9:00am the day of the presentation.

The final examination will be held on Wednesday, March 11 from 9:00-11:00 am in CarlSMgmt L-114.


The optional readings provide more detail or illustration of material in the text. They are entirely optional but will be posted on the website.
January 21, 26, 28  **Introduction, Ordinary Least Squares, the Classical Model and Hypothesis Testing**

**Reading:**
Studenmund, skim Chapters 1, 2, 3 and be sure you are familiar with the material in Chapter 17
Chapter 4: Sections 4.1 and 4.2; 
Chapter 5: Appendix 5.6 pp. 161-163; Chapter 4: Sections 4.3 and 4.4.

**Optional:**


January 28  **Choosing the Independent Variables**

February 2

**Reading:**
Studenmund, Chapter 6.

**Optional:**

February 4  **PROBLEM SET #1 DUE**

February 4, 9  **Choosing the Functional Form**

**Reading:**
Studenmund, Chapter 7.
Feb 9, 11, 16  
**Multicollinearity**

**Reading:**
Studenmund, Chapter 8.

February 18  
**PROBLEM SET #2 DUE**

Feb 18, 23, 25  
**Time Series Data**

**Reading:**
Studenmund, Chapter 9.

Optional:


March 4  
**PROBLEM SET #3 DUE**

March 2, 4, 9  
**The Nature and Consequences of Nonconstant Error variance and Putting it All Together**

**Reading:**
Studenmund, Chapter 10; Chapter 11, pp. 394-398.

Optional:

March 13  Exam