MULTIVARIATE ANALYSIS
SYA 5406
Spring 2009

Professor:	Dr. Karin L. Brewster	**Email**:	karin.brewster@fsu.edu
Office:	619 Bellamy	**Office Hours**:	Tu/We, 1:30 – 2:45
Phone:	644-7106		
TA:	Robyn Lewis	**Email:**	rklewis@fsu.edu
Office:	615 Bellamy	**Office Hours:**	Tu/Th, 9:30 – 12:00

Course Description

This course covers the use of multiple regression analysis in the social sciences. We'll begin by reviewing the statistical building blocks of simple linear regression, looking first at the bivariate case and then its extension to the multivariate model. Here, we'll pay close attention to the calculation and interpretation of basic regression statistics. We will then move on to consider the assumptions underlying the use of regression techniques, the estimation problems that result from the violation of these assumptions, and techniques to redress these problems. In addition, we will learn how to use *Stata*, a statistical software package, to prepare secondary data for analysis, estimate regression models, and perform regression diagnostics.

Course Objectives

The overarching objectives of this course are to enable you to become a proficient user of multiple regression and prepare you to move on to the more advanced techniques used in quantitative research in the social and behavioral sciences. Upon successful completion of this course, you will be able to:

1. Estimate and interpret OLS coefficients
2. Evaluate whether regression analysis is appropriate for a given research question and whether the data you intend to use meets the assumptions required for MRC analysis
3. Use Stata to prepare data for analysis, estimate multivariate regression models, and diagnose and resolve regression problems

Prerequisites

Successful completion of SYA 5455, Introduction to Social Statistics. This prerequisite ***may*** be waived if the student achieved a passing grade in a graduate-level introductory statistics course that included coverage of statistical inference and hypothesis testing, and bivariate regression and correlation.

Experience with statistical software. Previous experience with Stata is not required, but you should be conversant with the basics of computer-assisted "data crunching," including basic terminology (e.g., dataset, syntax, variable and value labels, missing values) and operations (e.g., producing tables, reading output).

While a background in calculus or matrix algebra may be helpful, they are not necessary for success in this course.

Required Texts

Cohen, Jacob, Patricia Cohen, Stephen West, Leona Aiken. 2003. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences.* Mahwah, NJ: Lawrence Erlbaum.

Acock, Alan C. 2006. *A Gentle Introduction to Stata.* College Station, TX: Stata Press.

Course Requirements and Grade Calculation

assignments

You will complete ten assignments (5 points each) this semester. These assignments will help you (and me!) evaluate your understanding of the course material, including key concepts, interpretation of regression output, and the use of Stata for data analysis. The assignments are designed to give you experience with multivariate model estimation, including trouble-shooting and diagnostics, and interpretation. Assignments will be posted on the course web site approximately one week before the scheduled due date.

tests

You will take two exams this semester: an in-class mid-term (25 points) and a final exam (25 points). Exams will include multiple choice and short-answer items; schedule and details will be posted on the course web site.

grade calculation

Assignments and mini-project components will be graded as follows:

Check-plus (√+): indicates exceptional work of professional or near-professional quality and counts 5 points

Check (√): indicates adequate understanding of the materials and counts 4.5 points.

Check-minus (√-): indicates deficiencies in your understanding of the materials, which may include serious problems of interpretation or major syntax errors. A check-minus grade is worth 3.75 points.

Course grades will be calculated on a 100-point scale, with each component weighted as described above. Numeric scores will be translated into letter grades using the following rubric:

		89-87	B+	79-77	C+	69-67	D+	59-0	F
100-93	A	86-83	B	76-73	C	66-63	D		
92-90	A-	82-80	B-	72-70	C-	62-60	D-		

Note: I reserve the right to adjust final grades for borderline cases, taking into consideration attendance, class participation (amount and quality), and rate of improvement over the course.

Class conduct

attendance

Attendance is required. Please be sure that you sign the attendance sheet each class.

professionalism

I expect professional behavior from all students. Professionalism includes arriving to class

on time and prepared for the day's work, handing in assignments when they are due and in the required format, adhering to the Academic Honor Code (see below), and interacting respectfully with other students and with the professor. Violation of these standards will result in grade penalties or dismissal from the class, at the instructor's discretion.

electronics

Classroom computers should remain off unless you are instructed to use them. You may want to bring a calculator to class. All other electronic equipment should be turned off and stowed away.

late assignments and missed exams

Because you will have at least one week to complete each assignment, late assignments will be accepted ONLY in unusual circumstances (e.g., hospitalization, jury duty), and only with appropriate documentation. If you must miss an exam due to illness or an emergency, please have someone notify me ASAP and be prepared to provide documentation testifying to your inability to attend the test at the scheduled time.

Collaboration and the Academic Honor Code

Much scientific research is collaborative, involving several investigators. My experiences as both a collaborator and lone researcher and as a teacher have led me to believe in the importance of collaborative learning. Therefore, I encourage you to work through the course materials together by exchanging notes, forming study groups, discussing assignment results, and so on. At the same time, your grade in this class will be based on *your* performance.

Here are some guidelines to collaborating without running amok of the FSU Academic Honor Code. First, *to ensure that you develop competency as an independent researcher, I ask you to work on assignments individually*. If you run into a problem *or* a classmate needs help, here are some examples of allowable collaborations:

- "My answer to Question 2 is 3.5. Is this what you got?"
- "I keep getting error messages in Stata. Have you seen these messages before? Do you understand what they mean?"
- "I am not sure whether or not I'm on the right track. Let me explain how I am trying to interpret the output, and can you tell me what you think?"
- "I don't know how I to start on Question 3. Can you tell me which part of the lecture notes and textbook, I should look at?"

The following types of questions are not acceptable – they are violations of the Honor Code:

- "Can you give me the Stata command you used?"
- "I don't have time to explain this to you. Can I just give you the Stata command I used?"
- "Let's work on this assignment together."
- "I forgot to save my Stata output. Can you send me yours?"

While learning can be collaborative, your grade in this class will be based on *your* performance. You are in violation of the code if your work is identical or substantially

identical to a classmate's work or if you are found to be using any materials on an exam other than those that I've supplied. Academic dishonesty **will result in a failing course grade** and **all** students involved will be reported to the Dean of Students. A full explanation of the Honor Code is available on-line at: http://www.fsu.edu/~dof/honorpolicy.htm

Americans with Disabilities Act

Students with disabilities needing academic accommodation should (1) register with and provide documentation to the Student Disability Resource Center; and (2) bring me documentation from the Center indicating the need for and type of accommodation. You should do this during the first week of class.

For information about services available to FSU students with disabilities, contact the Student Disability Resource Center at the Dean of Students Office. They're located at 08 Kellum Hall. Phone them at: (850) 644-9566 (voice) or (850) 644-8504 (TDD), email them at SDRC@admin.fsu.edu or visit their web site at http://www.fsu.edu/~staffair/dean/StudentDisability/

Getting help:

I tend to run class sessions informally, so if you don't understand something I've said, please ask me for clarification. If you have questions or need help completing the assignments, please see me or Robyn. We will be available during posted office hours or you can contact either of us by email to arrange for an appointment.

In doing the assignments, you may feel you don't know where to start. If this is the case, here are some steps to get you going:

1) Re-read the appropriate chapter(s) in the text or Stata book and review your lecture notes carefully, looking for keywords in the assignment.
2) Check your notes against those of another student in the class to be sure you haven't made some important error, or ask a student (or me or Robyn) if you're looking at the right material for the assignment.
3) Sometimes, getting another author's perspective is helpful. See the books I've recommended, below, or check one of the web-based help sites.

If you are coming to Dr. B. or Robyn for help, please help us help you:

- For Stata errors: What is the error message? What strategies have tried to solve the problem? Print the error message and any output and bring it with you.
- Output interpretations: which part of textbook and/or lecture notes did you consult? At what point are you getting bogged down?
- Assignments: What is the goal of the assignment (in your own words)? Bring any completed work with you, as well as your notes.

Recommended resources:

Texts:

Kahane, Leo. 2008. Regression Basics, 2nd edition. Sage.
A non-technical supplement that provides "plain English" explanations of many of the topics in Cohen et al.

Fox, John. 2008. Applied Regression Analysis and Generalized Linear Models. Sage.
Similar coverage to Cohen et al. with mathematical proofs and explanations. Assumes understanding of matrix algebra and calculus.

Richard A. Berk. 2003. Regression Analysis: A Constructive Critique. Sage.
A non-mathematical discussion of the benefits and limitations of OLS regression with some consideration of how OLS can be extended to other models.

Sage Monographs: "Quantitative Applications in the Social Sciences" (a.k.a., "greenies"):
Each volume covers a specific topic. The following are most relevant to our course:

Lewis-Beck, Michael. 1980. "Applied Regression: An Introduction." No. 22

Berry, William, and Stanley Feldman. 1985. "Multiple Regression in Practice." No. 50

Hardy, Melissa. 1993. "Regression with Dummy Variables." No. 93

Jaccard, James, Robert Turrisi, and Choi Wan. 1990. "Interaction Effects in Multiple Regression." No. 72.

Stata:

Built-in "Help" function in Stata. *From the top menu, click "Help" or, if you know the command name but not the syntax, type 'help command-name' in the command window.*

Stata Base Reference Manual (4 volumes).
Available at the Sociology computer lab. These manuals provide mathematical background as well as subcommands and options for each procedure. More in-depth than the built-in help.

Hamilton, Lawrence C. 2006. *Statistics with Stata*. Duxbury.
Available through the Stata bookstore or you may be able to borrow one from a peer – many students from prior cohorts have copies. I recommend (strongly) that you invest in a copy of your own if your academic and/or career plans include quantitative research with Stata.

Web-based:

Two invaluable resource sites:

- FAQ at stata.com: http://www.stata.com/support/faqs/.
- UCLA, Advanced Technology Services: http://www.ats.ucla.edu/stat/stata/.
  Includes FAQ, "getting started" movies, links to various textbooks, and more!

Course Outline and Readings

Preliminaries
Review statistical building blocks of MRC; getting started in Stata

Reading:
MRC Chapter 1
GIS Chapters 1 – 3 (skim 2)

Topic I **Bivariate Regression and Correlation**
Estimating and interpreting linear relationships between two variables

Reading:
MRC Chapters 1, 2
GIS Chapters 4 – 5, 8

Topic II **Multiple Linear Regression**
Logic of control; redundancy, suppression, and indirect effects; assessing model fit

Reading:
MRC Chapter 3
GIS Chapter 10, parts 1, 2, 3, 4

Topic III. **Regression Diagnostics, part 1**
Regression assumptions; detecting and remedying violations

Reading:
MRC Chapter 4
GIS Chapter 10, parts 5, 6, 7

Topic IV. **Analytic Strategies**
Simultaneous vs. stepwise estimation; testing nested models

Reading:
MRC Chapter 5
GIS Chapter 10, continued

Topic V. **Nonlinearity and Nonadditivity**
Power polynomials, logarithms, first versus higher-order effects, centered predictors

Reading:
MRC Chapters 6, 7
GIS Chapter 10, part 10

Topic VI. **Regression with Categorical Predictors**
Dummy coding, effects coding, nominal x nominal interactions, nominal x quantitative interactions

Reading:
MRC Chapters 8, 9
GIS Chapter 10, part 9

Topic VII. **Regression Diagnostics, part 2**
Identifying and remedying outliers; measuring and remedying multicollinearity

Reading:
MRC Chapters 10, 11

Topic VII. **Regression with Categorical Dependent Variables**
Logistic regression (binomial and multinomial)

Reading:
MRC Chapters 13
GIS Chapter 11

Key dates

Test 1: **2-26-09,** 8:00-9:15 AM
Topics I – IV, MRC Chapters 1-5

Test 2: **4-30-09**, 3:00 to 5:00 PM
Topics V – VII, MRC Chapters 6-11, 13

Assignment due dates:

Assignment 1: 1-20
Assignment 2: 1-29
Assignment 3: 2-10
Assignment 4: 2-19
Assignment 5: 3-3
Assignment 6: 3-17
Assignment 7: 3-26
Assignment 8: 4-7
Assignment 9: 4-16
Assignment 10: 4-27

Syllabus Change Policy

This syllabus is a guide for the course and is ***subject to change with advance notice***. Changes will be announced in class and on the course web site.