Introduction

The purpose of this course is to introduce graduate students to the theory, method, and practice of regression analysis and its many variants. Regression analysis is the principle method of empirical inquiry in the social sciences. It is a powerful tool for testing models of political and social phenomena and making inferences about the empirical world. Most other techniques can be understood in terms of how they relate to regression analysis. As such, it is important for advanced students in the social sciences to be well grounded in regression analysis, including its core assumptions, the implications of violations of those assumptions, and corrective actions that can be taken when regression assumptions are not met.

The seminar will proceed as follows. First, we will spend several weeks discussing the basic bivariate and multivariate regression model. Here we will focus on the general logic underlying regression analysis, as well as the assumptions upon which it is built. Second, we will discuss the implications of those assumptions for model estimation and the quality of inferences drawn from regression analyses. In particular, what problems arise when one or more regression assumptions are violated or relaxed? Third, we will spend the bulk of the semester on various techniques used to diagnose and correct for violations of regression assumptions. Finally, as time permits we will explore a range of specialized methods related to regression, including logit and probit models, time-series analysis, non-recursive models, factor analysis, and pooled cross-sectional time-series analysis.

Prerequisites

All students enrolled in this course should have completed (or be in the process of completing) an introductory statistics class such as POLI 7962 (or its equivalent). Any seminar on regression analysis must assume a working knowledge of elementary statistical concepts and techniques. We will conduct a brief review at the beginning of the course, but students should be familiar with such ideas as descriptive statistics, sampling distributions, statistical inference, and hypothesis testing.

Course Requirements and Evaluation

Each student will be evaluated on the basis of the following:

- Midterm Examination: 30%
- Final Examination: 30%
- Research Paper: 40%
1. **Midterm Examination.** A comprehensive midterm examination covering all material from the first half of the course will be required for all students. Information pertaining to the format of the midterm examination will be provided to all students prior to the exam.

2. **Final Examination.** A comprehensive final examination covering all material from the second half of the course will be required for all students. Information pertaining to the format of the final examination will be provided to all students prior to the exam.

3. **Research Paper.** In addition to the midterm and final exams, each student will be required to write a research paper that utilizes one of the statistical techniques discussed during the semester. More detailed information about the paper will be provided later in the semester.

**Required Reading**

There is no required text for this course. However, I will be using the following book as the basis for much of the course:


This book has a list price of $155, and is hence quite expensive. I have checked on Amazon.com and found versions of the same book for about $112 (new) and $111 (used), and there are websites that will sell used copies of the book for less than that. There is also a paperback version of an earlier edition on line that is much less expensive (about $40); I tried to order this version for the class but the book is not available in the United States from the publisher. If you can secure this version of the book online that would probably be fine as well.

This book would be a very useful book to have, but it is not required and there will be no course requirements taken directly from the book. I will leave up to you the decision of whether or not to purchase this book; if you decide to purchase the book you can do so online.

There will be some required and recommended reading during the course, primarily scholarly journal articles or other research papers. These will be made available through Moodle and/or by email.

**Computer Statistics Package**

There are numerous statistics programs that can be used to conduct statistical analysis with a mainframe or personal computer. In this course we will use Stata 12.0, a commonly-used (and easy-to-use) statistics program.

If you are serious about doing quantitative research, I would encourage you to purchase a copy of Stata. Other programs (such as SPSS or SAS) are fine, but Stata is easy to learn and use, and I find that it is much more powerful and flexible for most applications than other programs. This is particularly the case for regression analysis, which will be the focus of this class. Stata will be the only program supported during this class, and computer assignments will be conducted using Stata.

Stata is expensive, but the Stata Corporation provides educational discounts. Stata can be purchased at the following web site:

http://www.stata.com/order/new/edu/gradplans/gp-campus.html

Notice that there are multiple options. First, if you are going to purchase Stata, at the very least you should purchase the Intercooled Stata 12.0 with the one-year license ($98). Second, a mid-range (and recommended) option is Intercooled Stata 12.0 with a perpetual license ($179). Finally, if you (1) will be doing a lot of statistical analysis in your
research in the future, particularly with larger data sets, and (2) can at all possibly afford it, I encourage you to purchase Stata / SE 12.0, which is the most powerful version of Stata but is somewhat pricey ($395). Do not purchase the Small Stata 12.0, which is really for very small data sets and will not accommodate some of the data sets that we will use this semester.

**SPOST and Clarify**

There are two add-ons to Stata that we will use during the course of the semester. Both of these programs are free and can be downloaded and easily integrated into Stata. You should download and install these programs as soon as you have installed Stata on your computer.

First, Gary King and collaborators have developed a program called Clarify, which is used by researchers to estimate a regression model and then generate predicted values on the dependent variable, all-the-while manipulating the effects of some independent variables and holding the effects of others constant. Clarify is relatively easy to use and can generate results that can be presented in tables or figures. You can obtain a copy of clarify from the following link:

http://gking.harvard.edu/clarify

Once you have downloaded this program as a compressed file, you should place it temporarily in a folder on your hard drive; I would suggest that you create a new folder in your c: directory, such as c:\clarify. Once the zipped file is in this directory, you can extract the files from the compressed file, and then these files can be installed in Stata using the following commands:

```plaintext
net from c:\clarify (or whatever directory you placed the uncompressed files)

net install clarify
```

Second, J. Scott Long is a foremost authority on the use of categorical data analysis in the social sciences. Long has developed an add-on program for Stata, called SPOST, that does many of the same things that Clarify does, as well as some other things that Clarify does not do. We will use SPOST during the course of the semester. In order to download SPOST, type the following command in Stata:

```plaintext
findit spost
```

This will open a viewer screen in Stata with a variety of links. Scroll down and click on the following link:

```plaintext
spost9_ado from http://www.indiana.edu/~jslsoc/stata
```

Note that this link has “ado” in the address. Stata will prompt you to install the files. Click on “Click here to install” and follow the directions. You will now have access to SPOST commands.

**Stata Resources**

Because many of you have not had experience with Stata, there are several Stata resources that I recommend. First, there is also a lengthy (but informative) introduction to Stata that is a PowerPoint presentation converted to a .pdf format:

Christopher Baum, *Introduction to Stata*

This file will be available on the course Moodle page.
Second, I will make available to you a pdf introduction to Stata:

Sven Juul, *Introduction to Stata 8.0*.

Even though this manual is designed for Stata 8.0 rather than Stata 12.0, the differences in the two versions are sufficiently small that Juul’s manual is likely to be helpful to you.

In addition, I would like to direct you to the following web sites that are designed to assist Stata users. These are extremely useful web sites that provide detailed information about Stata commands. You should bookmark these web sites and refer to them often.

The first is a broad-based Stata web site housed at UCLA. One can find a wide range of information about Stata on this web site. I would encourage you to take a look at the links on this web site and familiarize yourself with what this site has to offer:

http://www.ats.ucla.edu/stat/stata/

This is the Stata “starter kit” for new users, also available on the UCLA web site:

http://www.ats.ucla.edu/stat/stata/sk/

Another introductory web site is found on the University of North Carolina web site:

http://www.cpc.unc.edu/research/tools/data_analysis/statatutorial

This UCLA site has information on estimating regression models and other similar models in Stata. This is a very good resource for information about estimating models:

http://www.ats.ucla.edu/stat/stata/webbooks/reg/default.htm

Here is another UCLA site that has “learning modules” for Stata procedures.

http://www.ats.ucla.edu/stat/stata/modules/default.htm

Here is a web page at Princeton that provides a basic introduction to Stata:

http://data.princeton.edu/stata/default.html

Finally, this is a Princeton web site that has a series of downloadable chapters about how to estimate a variety of models in Stata.

http://data.princeton.edu/wws509/stata/

**Calculator**

In addition, many of the statistical problems to be worked out during the semester require the use of a calculator with a square root function. Each student is strongly encouraged either to purchase such a calculator or otherwise to have one available. A simple calculator with a square root function can be purchased for well under $10.
Office Hours

Garand: Wednesday afternoon, 3:00 to 4:30
Other hours by appointment

Academic Misconduct Statement

Academic misconduct is defined by the Code of Student Conduct. You are encouraged to familiarize yourself with the LSU policy on academic misconduct, particularly regarding plagiarism. The LSU Code of Student Conduct can be found on the web site for the LSU Dean of Students:

http://appl003.lsu.edu/lsas/dos.nsf/index

Academic misconduct is a serious violation of university policy, but more importantly it is a significant scholarly violation for political scientists. Plagiarism and other forms of academic misconduct will not be tolerated in this course. Charges of academic misconduct will be turned over to the Dean of Students for appropriate disciplinary action.

I am serious about this.
Class Outline:

The following is a tentative outline of topics for the semester. The instructor reserves the right to make adjustments in the schedule of topics and readings as necessary and with reasonable advance notice.

January 18  Introduction / Statistics review

Recommended:
Agresti and Finlay, Statistical Methods for the Social Sciences, chapters 1-7.
Knoke, Bohrnstedt, and Mee, Statistics for Social Data Analysis, chapters 1-4.

January 25  Statistics review (continued)

February 1  Bivariate regression models

Recommended:
Gujarati and Porter, Basic Econometrics, chapters 1-6.

February 8  Bivariate regression models (continued)

February 15  Multiple regression models

Required:

Recommended:
Gujarati and Porter, Basic Econometrics, chapters 7-9.
Berry and Sanders, Understanding Multivariate Research.
Rubenfeld, Reference Guide on Multiple Regression.
King, "Stochastic Variation: A Comment on Lewis-Beck and Skalaban's 'The R-Square."

February 22  Multiple regression models (continued)

Required:
Garand, "Income Inequality, Party Polarization, and Roll-Call Voting in the U.S. Senate," Journal of Politics (October 2010: 1109-28).
### Class Outline (continued):

<table>
<thead>
<tr>
<th>Month</th>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>February</td>
<td>29</td>
<td>Multiple regression models (continued)</td>
</tr>
<tr>
<td>March</td>
<td>7</td>
<td>Regression assumptions</td>
</tr>
<tr>
<td>March</td>
<td>14</td>
<td>Regression assumptions (continued)</td>
</tr>
<tr>
<td>March</td>
<td>21</td>
<td>Regression assumptions (continued)</td>
</tr>
<tr>
<td>March</td>
<td>28</td>
<td>Regression assumptions (continued)</td>
</tr>
<tr>
<td>April</td>
<td>4</td>
<td>Discrete dependent variables</td>
</tr>
<tr>
<td>April</td>
<td>11</td>
<td>Spring Break (no class meeting)</td>
</tr>
<tr>
<td>April</td>
<td>18</td>
<td>Discrete dependent variables (continued)</td>
</tr>
<tr>
<td>April</td>
<td>25</td>
<td>Discrete dependent variables (continued)</td>
</tr>
<tr>
<td>May</td>
<td>2</td>
<td>Time-series analysis / analysis of panel data</td>
</tr>
</tbody>
</table>

**Recommended:**


Berry, *Understanding Regression Assumptions*.

Berry and Feldman, *Multiple Regression in Practice*.

Fox, *Regression Diagnostics*.


Suggestions for Further Reading

In addition, the following are useful reference or supplemental readings. These works should be available in the Middleton Library, through J-STOR, or for purchase from Amazon.com or some other online book seller. I have placed an asterisk (*) next to works that I find particularly useful.

General

* Alan Agresti and Barbara Finlay, *Statistical Methods for the Social Sciences*.
* Eric Hanushek and John Jackson, *Statistical Methods for Social Scientists*.
  Gary King, *Unifying Political Methodology: The Likelihood Theory of Statistical Inference*.
  Lawrence Mohr, *Understanding Significance Testing*.

Regression Analysis

  William Berry and Mitchell Sanders, *Understanding Multivariate Research*.
* William Berry and Stanley Feldman, *Multiple Regression in Practice*.
  Christopher Achen, *Interpreting and Using Regression*.
  William Berry, *Understanding Regression Assumptions*.
  John Fox, *Regression Diagnostics*.
  Melissa Hardy, *Regression with Dummy Variables*.
  James Jaccard, Robert Rurrsli, and Choi Wan, *Interaction Effects in Multiple Regression*.
* Jan Kmenta, *Elements of Econometrics*.
* Andrew Gelman and Jennifer Hill, *Data Analysis Using Regression and Multilevel/Hierarchical Models*.

Logit and Probit Models

  Maddala, *Limited Department and Qualitative Variables in Econometrics*.
Time-Series and Pooled Models

Box and Jenkins, *Time Series Analysis.*

Paul Allison, *Event History Analysis.*
About the Instructor

James C. Garand (Ph.D., University of Kentucky, 1984) is the Emogene Pliner Distinguished Professor of Political Science and the R. Downs Poindexter Professor of Political Science at Louisiana State University. In addition to holding named professorships in the Department of Political Science, he serves on the faculty of the Manship School of Mass Communication and is a Senior Fellow at the Reilly Center for Media and Public Affairs.


Professor Garand received the 2006 LSU Distinguished Research Master Award in recognition of outstanding faculty accomplishments in research and scholarship. In 2009 he was recognized as an LSU “Rainmaker,” an award given by the LSU Office of Research and Economic Development (ORED) for national and international recognition “for innovative research and creative scholarship.” He was President of the Southern Political Science Association in 2004, and he is also former president of the State Politics Section of the American Political Science Association. He served as Vice-President and Program Chair in 2001 for the Southern Political Science Association. Professor Garand is former editor of the *American Politics Quarterly*, one of the leading subfield journals in American politics. He currently serves on the editorial boards of the *American Politics Research*, *Journal of Political Marketing*, and *Ralph Bunche Journal of Public Affairs*, and he is a former member of the editorial boards of the *American Journal of Political Science*, *Journal of Politics*, *State Politics and Policy*, and *Legislative Studies Quarterly*.

Professor Garand has received numerous faculty awards. In 1997 Professor Garand received the LSU Alumni Association Distinguished Faculty Award in recognition of sustained excellence in teaching, research, and service. In 2001 he received the LSU Foundation Distinguished Faculty Award in recognition of his excellence in graduate teaching, and he is the 1990 recipient of the university-wide Student Government Association Teaching Excellence Award for undergraduate teaching. He is also a recipient of the Alpha Lambda Delta Freshman Honor Society certificate of recognition for superior instruction of freshman students during the Fall 2000 semester.