PS 204B QUANTITATIVE METHODS I •
Winter 2005
UCSD

Class: I. Mondays 12:00-13:30, SSB 104 &
        II. Thursdays 12:00-13:30, SSB 104

Lab: Mondays 13:30-15:00, SSB 140/139 Lab

Professor Kristian Skrede Gleditsch, SSB 383, kgleditsch@ucsd.edu
Office hours: Tuesday 9.30-11.30, or by appointment

TA: Idean Salehyan, SSB 328
TA office hours: Thursday 11:00-12:00

This version: 13 January 2005

DESCRIPTION

This course covers basic statistical techniques and their application to political analysis. It introduces students to widely used statistical techniques and data analysis, and provides a basis for learning and understanding more advanced statistical methods.

Students take this course for various reasons. All political scientists need to be good readers of the quantitative literature, almost all will do some quantitative analysis in their own research or dissertation work, and some students will eventually wish to use advanced quantitative methods. This is not a "high end" course by itself, but it should enable all students to undertake serious quantitative analysis. PS204B will also provide the necessary skills for PS273 (Quantitative Methods II) for those that want to go further.

I do not assume any knowledge of math or statistics beyond basic high school algebra (although some willingness to learn along the way will be helpful). In particular, we will not make any use of matrix algebra for this course. Use of calculus will be minimal, primarily to indicate what is going on in some of the proofs. Although the ideas behind some proofs are very important, the specific details will be less essential for many students.

It should be obvious that PS204B caters to a heterogeneous audience. The trick is for everyone to try to get what they need out of the course. I trust that each of you will put in the effort to get what you need to get out of the course. Perhaps a few of you will discover that they can do much more than they thought they could in advance.

ORGANIZATION

This class is based on two lectures as well as a lab session. The lab sessions will be a combination of taught sessions and tutorials, as well as less formal sessions to give you an opportunity to receive help and input on your research project.

• *Disclaimer: The syllabus and course outline is intended to provide an overview over the course. You cannot claim any rights from it. In particular, scheduling and dates may change. Although the syllabus should be a fairly reliable guide for the course, official announcements are always those made in class.*

The class homepage is http://weber.ucsd.edu/~kgledits/ps204b_w2005/ps204b.html. This will contain all the relevant information about course details, so please check often for any changes. Please note that I encourage questions by email, which will allow me to start each session with questions that have arisen since our last meeting.

REQUIRED TEXT

Wooldridge, Jeffrey M. 2003. *Introductory Econometrics: A Modern Approach, Second Edition.* Mason, OH: Thomson-South Western.

Wooldridge will be the key text for the course, and we will follow the text closely. The course will concentrate on the key ideas and their relevance to political analysis. Econometrics is simply the term that economists use for applications of statistical methods to economics, and there is nothing particularly "economic" about these methods themselves. Some topics stressed in econometrics textbooks are more important to us as political scientists than others. The sections of Wooldridge listed in the course outline are primarily meant to serve as a guide for your reading, and some parts will be more relevant than others.

Given the high quality of the text (and the high likelihood of multiple typos when time does not permit careful proofreading) I will not offer course notes in advance for this course. You should feel free, however, to ask for copies of overheads and other material that I present in class.

Additional articles with applications of the course material will be assigned along the way. The first exercises for the course will be based on an article that you should read during the first week of class; Muller, Edward N., and Mitchell A. Seligson. 1987. "Inequality and Insurgency." *American Political Science Review* 87:425-51. This article is available in PDF format through JSTOR (http://www.jstor.org). Please let me know if you should have any problems accessing it.

RECOMMENDED TEXTS

Many of you may find it helpful to look at more than one book to learn this material. You may for example wish to look up particular topics in others source if you find something inadequately explained in Wooldridge. Some alternatives to consider are listed below.

Amemiya, Takeshi. 1993. *Introduction to Statistics and Econometrics.* Cambridge, MA: Harvard University Press. This book provides a very thorough introduction to the foundations of statistical inference and probability theory.

Fox, John. 1997. *Applied Regression Analysis, Linear Models, and Related Methods.* Thousand Oaks, CA: Sage. This is a very good and thoughtful book. It covers most of the topics we will look at in this class and is generally written in an accessible style, but it uses matrix notation a great deal more than Wooldridge.

Gujarati, Damodar. 2003. *Basic Econometrics.* New York: McGraw-Hill. This text is similar to Wooldridge and may appeal to some as an alternative.

Gujarati, Damodar. 2003. *Essentials of Econometrics.* New York: McGraw-Hill. Chapters 1-4 of this book were assigned readings for the old Mathcamp run in the department. They provide a good introduction to statistical inference and probability theory.

Hanushek, Eric A., and John E. Jackson. 1977. *Statistical Methods for Social Scientists.* Orlando, FL: Academic Press. Some people love this book, others dislike it intensely. Note that this book has a discussion-based approach that makes it hard to look up individual concepts without having to read the text extensively.

Kennedy, Peter. 1992. *A Guide to Econometrics*. Cambridge: MIT Press. This is akin to a "Cliff Notes" book on statistical methods. Many people find this very useful as a reference book.

Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks, CA: Sage. Standard Econometrics textbooks do not offer extensive coverage of regression models with categorical dependent variables, and this book is a very useful supplement to Wooldridge.

Maddala, G. S. 1992. *Introduction to Econometrics.* New York: Macmillan. This text is similar in difficulty to Wooldridge. The author is quite opinionated and has a somewhat quirky style. I like this text a great deal, but other people react less favorably.

Maddala, G. S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press. Very readable reviews of various models for binary and ordinal dependent variables. More advanced than the Long book.

Pindyck, Robert S. and Daniel Rubinfeld. 1981. *Econometric Models and Economic Forecasts.* New York: McGraw-Hill. This is a very good book, similar in level of difficulty to Wooldridge.

SOFTWARE

We will use Stata as the main statistical software for much this course. Stata is relatively easy to learn, and includes many advanced procedures not commonly found in similar statistical packages.

Students are encouraged to purchase a personal copy through the so-called Stata grad plan. I do not have a vested interest in making you buy the software, but it will in all likelihood be helpful for you to be able to work with the material at home. In the long run, most of you will need to own some statistical package, and Stata will be a reasonable option given its price and flexibility and your needs.

For information about the Grad Plan program, see http://www.stata.com/info/order/new/edu/gradplans/gp3-order.html. Faculty and students can purchase this package by calling Stata Corp. at 800-782-8272 (800-STATAPC, ask for Sales). Be sure to identify yourself as a UCSD affiliate to get the special price. After you place your order, Stata will contact our designated representative, Jeff Fritsch, for license and distribution for pickup within a few days at the SSB 301 reception desk.

The best way to learn Stata is to use Stata, which you will do in the exercises for the course. If you have never used *Stata* before, I recommend reading *Getting Started with Stata* (College Station, TX: Stata Press, 2001) at the outset of the course to familiarize yourself with using the software. You should also learn to take advantage of Stata's on-line help system.

GRADING

Your final grade will be determined by four problem sets (60% of the total grade, 15% each) and a final paper (40% of total grade)

The four problem sets will be handed out in class and should be done with Stata. I expect that each student has access to Stata either on personal machines or in the labs. Data will be distributed on the class web site. I expect that all graduate students can download the data from the web page and get it to the appropriate machine. Please contact me if you have any problems.

Each student must also submit an individual research project. This is due during the week of final. I will consult with each student about his or her project. The final paper should be similar to a research paper, but much

heavier on the methods and much lighter on front-end, theory, literature review and substantive conclusions. Some papers may wish to use more advanced methods that we will not be able to cover in this class. If so, you will be expected to read the relevant chapters in Wooldridge or other sources. Many of you, however, will probably find that the material we cover in class will suffice for your paper.

COURSE OUTLINE

1. Monday 3 January

  Topics: Overview and course details
  Math/stat review: random variables, linear functions, descriptive statistics
  Reading: Wooldridge Chapter 1, Apps. A.1-A.4; Muller & Seligson (1987)

2. Thursday 6 January

  Topics: Basic regression ideas, data analysis
  Math/stat review: expectations, probability distribution functions, transformations
  Reading: Wooldridge Chapter 2 (sections 2.1-2.4), Apps. B.1-B.3

3. Monday 10 January

  Topics: Two variable regression, OLS estimation
  Math/stat review: summation operators, variance, covariance
  Reading: Wooldridge Chapter 2 (section 2.5), App. B.4

4. Thursday 13 January

  Topics: Multiple regression; motivation, estimation, and interpretation
  Reading: Wooldridge Chapter 3 (sections 3.1-3.3)

(Monday 17 January is Martin Luther King Day, no formal class)

5. Thursday 20 January

  Topics: Multiple regression; standard errors, normality assumption, misspecification
  Reading: Wooldridge Chapter 3 (sections 3.4-3.5)

  1st problem set is due (in class)

6. Monday 24 January

  Topics: Inference in multiple regression analysis; confidence intervals and hypothesis testing
  Reading: Wooldridge Chapter 4 (sections 4.1-4.3) & McCloskey and Ziliak 1996 (Optional; Cohen 1994)

7. Thursday 27 January

Topics: Multiple regression; goodness of fit tests, testing multiple restrictions
Reading: Wooldridge Chapter 4 (sections 4.4-4.6) (Optional: King 1986; Luskin 1990, and Lewis-Beck and Skalaban 1990 controversy on the R-square)

8. Monday 31 January

Topics: Multiple regression; model selection, prediction, residual analysis, diagnostics
Reading: Wooldridge Chapter 6 (Optional: Bollen 1983, 1988 and Gasiorowski 1988 controversy)

2nd problem set is due (in class)

9. Thursday 3 February

Topics: Dummy independent variables and interaction effects
Reading: Wooldridge Chapter 7 (7.1-7.4, 7.6) & Friedrich 1982

10. Monday 7 February

Topics: Non-constant variances
Reading: Wooldridge Chapter 8 (sections 8.1-8.4) (Optional: Downs and Rocke 1979)

11. Thursday 10 February

Topics: More on specification, measurement error, data analysis
Reading: Wooldridge Chapter 9

12. Monday 14 February

Introduction to problems & extensions: Autocorrelation & pooling observations
Reading: Wooldridge Chapters 10 (10.1-10.3), 12 (12.1-12.3), 13 (13.1-13.4) (Optional: Stimson 1985; Beck and Katz 1995)

14. Thursday 17 February

Topics: Binary dependent variables, the linear probability model
Reading: Wooldridge Chapters 7 (section 7.5), 8 (section 8.5)

15. Monday 21 February

Topics: Limited dependent variables, logit, probit
Reading: Wooldridge Chapter 17 (section 17.1)

(Monday 21 February is President's day, but we will meet this day to make up for classes lost later)

16. Thursday 24 February

Topics: Maximum likelihood estimation, inference in limited dependent variable models
Reading: Wooldridge Chapter 17 (section 17.1), App. C.4

4th problem set is due (in class)

17. Monday 28 February

Lab session only – meet at 12:00 in SSB 139/140 lab
No new reading assigned.

18. Thursday 3 March

Topics: More on limited dependent variables: time dependence, simulation
Reading: Wooldridge Chapter 17 (section 17.1), App. C.4 & Beck et al. 1998; King et al. 2000

No class on Monday 7 March

19. Thursday 10 March

Topics: Instrumental variables and two stage least squares
Reading: Wooldridge Ch. 15 (sections 15.1-15.5) (Optional: Green and Krasno 1988; Jacobson 1990).

A make up class/review session may be scheduled for the week of final exams

Supplementary reading

Beck, Nathaniel, and Jonathan N. Katz. 1995. "What To Do (and Not To Do) with Time-Series Cross-Section Data." *American Political Science Review* 89:634-47.

Beck, Nathaniel, Jonathan N. Katz, and Richard M. Tucker. 1998. "Taking Time Seriously: Time-Series Cross-Section Analysis With a Binary Dependent Variable." *American Journal of Political Science* 42:1260-88.

Bollen, Kenneth A. 1983. "World System Position, Dependency, and Democracy: The Cross-National Evidence." *American Sociological Review* 48:468-79.

Bollen, Kenneth A. 1988. "If You Ignore Outliers, Will They Go Away? Response to Gasiorowski." *Comparative Political Studies* 20:516-22.

Chatterjee, Samprit, and Frederick Wiseman. 1983. "Use of Regression Diagnostics in Political Science." *American Journal of Political Science* 27:601-13.

Cohen, Jacob. 1994. "The Earth is Round (p<.05)." *American Psychologist* 49:997-1003.

Downs, George and David Rocke. 1979. "Interpreting Heteroskedasticity." *American Journal of Political Science* 23:816-28.

Friedrich, Robert J. 1982. "In Defense of Multiplicative Terms in Multiple Regression Equations." *American Journal of Political Science* 26:797-833.

Gasiorowski, Mark J. 1988. "Economic Dependence and Political Democracy: A Cross-National Study." *Comparative Political Studies* 20:489-515.

Green, Donald Philip, and Jonathan S. Krasno. 1988. "Salvation for the Spendthrift Incumbent: Reestimating the Effects of Campaign Spending in House Elections." *American Journal of Political Science* 32:884-907.

Jacobson, Gary C. 1990. "The Effects of Campaign Spending in House Elections: New Evidence for Old Arguments." *American Journal of Political Science* 34:334-62.

King, Gary. 1991 "Stochastic Variation: A Comment on Lewis-Beck and Skalaban's `The R-Square." *Political Analysis* 2:185-200.

King, Gary. 1986. "How Not to lie with Statistics: Avoiding Common Mistakes in Quantitative Political Science." *American Journal of Political Science* 30:666-87.

King, Gary, Jason Wittenberg, and Michael Tomz. 2000. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." *American Journal of Political Science* 44:347-61.

Leamer, Edward E. 1983. "Let's Take the Con out of Econometrics." *American Economic Review* 73:31-43.

Lewis-Beck, Michael S., and Andrew Skalaban. 1990. "The R-Squared: Some Straight Talk." *Political Analysis* 1:153-72.

Luskin, Robert C. 1991. "Abusus Non Tollit Usum: Standardized Coefficients, Correlations, and R2s." *American Journal of Political Science* 35:1032-46.

McCloskey, Deirdre, and Stephen T. Ziliak. 1996. "The Standard Error of Regression." *Journal of Economic Literature* 34:97-114.

Stimson, James A. 1985. "Regression in Time and Space: A Statistical Essay." *American Journal of Political Science* 29:914-47.