Government 391L
Spring, 2012
Unique number: 39005

Robert C. Luskin
University of Texas at Austin
T 6:30-9:30, GAR 1.134

Office: BAT 3.148, Office phone: 232-7230
Office Hours: Monday 5:00-7:00,
Wednesday 2:30-3:30, & by appointment

Statistical Analysis in Political Science II:
Regression Models

This course is about "regression models," roughly and broadly defined as statistical models to explain some single dependent variable. The domain thus includes nonlinear as well as linear models and models for qualitative as well as quantitative dependent variables. I shall try to please two clienteles: aspiring methodologists and aspiring practitioners. The course outline, below, is ambitious, and we may well not get all the way though it, but I do hope to get at least through Topic XV.

There will be computer-based exercises to provide concrete examples, but the lectures and readings will focus on general questions of modeling, estimation, inference, and interpretation: What sorts of models imply and reflect what sorts of relationships between independent and dependent variables? What assumptions must we make, and what do they mean? How likely are the assumptions to be violated, and with what consequences? How can we tell when violations occur? What alternative assumptions might we make? What quantities should we be focusing on estimating? What estimators provide statistically desirable estimates? Where several different estimators might serve, what are their advantages and disadvantages? What do the estimates tell us, and how certainly?

The lectures and readings will treat these questions practically but abstractly, referring more to *x*'s and *y*'s than to substantive variables. There will be much mathematical notation and mathematically phrased argument and some proof and derivation. The goal is to convey a good, relatively deep understanding of the how's and why's of constructing, estimating, and interpreting the estimates of these models.

To get the taste of actual modeling and analysis, we shall examine some published examples and work through a series of mostly computer-based exercises asking you to write and analyze your own models, rooted in your own substantive interests. I shall help you find datasets, if you don't have any you are already working on or interested in. There will also be an optional term paper, to be centered on a regression model of your devising.

There will be extra sessions, yet to be scheduled, of approximately two hours per week, to review the mathematical and statistical background, go over questions and assignments, and discuss concrete applications.

Prerequisites

The course requires a decent knowledge of descriptive and inferential statistics—as covered in the Government Department's Statistics I or equivalent—and a reasonable facility with ordinary algebra. I encourage you to consult me if you are unsure whether this is an appropriate course for you to take.

Texts

The texts, selected for simplicity without simplism, are:

Damodar N. Gujarati and Dawn C. Porter. 2008. *Basic Econometrics* (5th ed.). New York: McGraw-Hill.

Jan Kmenta. 1997. *Elements of Econometrics* (2nd ed.). Ann Arbor, MI: University of Michigan Press.

A.H. Studenmund. 2010. *Using Econometrics: A Practical Guide* (6th ed.). Reading, PA: Addison-Wesley.

Studenmund is the simplest of these, Kmenta the most sophisticated. You should read Gujarati and Porter and *at least* one of the other two. You should prefer Kmenta to Studenmund, if you can handle it. *Everyone* should read Kmenta's chapters 1-6, under "Statistical Review." I also strongly recommend:

Jeffrey M. Wooldridge. 2008. *Introductory Econometrics: A Modern Approach* (4th ed.). Mason, OH: South-Western College Publishing.

Jeffrey M. Wooldridge. 2010. *Econometric Analysis of Cross Section and Panel Data* (2nd ed.). Cambridge, MA: MIT Press.

the first at approximately the same level as Gujarati and Porter, the second somewhat more advanced.

Topics and Reading Assignments

I Statistical Review

Kmenta, chs. 1-6.
Gujarati & Porter, app. A.

II The Regression Game: Models and Meanings

Gujarati & Porter, chs. 1-2, secs. 7.1 (except the bit about assumptions)-7.3,

and peek ahead to secs. 6.4-6.9, 7.9-7.10.
Kmenta, secs. 7.1, 7.2.
Studenmund, chs. 1, 4, 6-7.
You may also wish to peek ahead to the readings under VII

III Classical Assumptions

Gujarati & Porter, secs. 3.2, 4.2, and 7.1 (the bit about assumptions).
Kmenta, secs. 7.1, 7.2.

IV Estimating the Coefficients: Ordinary Least Squares, Maximum Likelihood

Gujarati & Porter, secs. 3.1, 3.3-3.4, 4.1, 4.3-4.4, 7.4-7.7, pp. 217-19, apps. 3A, 4A.
Kmenta, sec. 7.3.
Studenmund, chs. 2-3.

V Forecasting, Fit, and Inference

Gujarati & Porter, secs. 3.5, 7.5, 5.2-5.10, 8.1-8.8
Kmenta, sec. 7.4.
Studenmund, ch. 5, 15.

VI Matters of Interpretation

Gujarati & Porter, sec. 6.1-6.3, 7.2-7.3, 7.6, app 6A.2.
Gary King. 1986. How Not to Lie with Statistics: Avoiding Common Mistakes in Quantitative Political Science. *American Journal of Political Science*, 35: 1032-46.
Robert C. Luskin. 1991. Abusus Non Tollit Usum: Standardized Coefficients, Correlations, and R^2s. *American Journal of Political Science*, 35: 1032-46.
Studenmund, ch. 11.

VII Some Political Science Applications

Robert W. Jackman. 1987. Political Institutions and Voter Turnout in the Industrial Democracies. *American Political Science Review*, 81: 405-23.

Clive Bean and Anthony Mughan. 1989. Leadership Effects in Parliamentary Elections in Australia and Britain. *American Political Science Review*, 83: 1165-80.

VIII Trouble in Regression City: Heteroskedasticity and Autoregression (Introducing Generalized Least Squares)

Gujarati & Porter, chs. 11-12
Kmenta, ch. 8, sec. 9.2.
Studenmund, chs. 9-10.

IX More Trouble: Collinearity and Related Ills

Gujarati & Porter, ch. 10.
Kmenta, sec. 10.3.
Studenmund, ch. 8.

X Still More Trouble: Correlations of Regressor and Disturbance

Gujarati & Porter, secs. 7.7, 13.2-13.5.
Kmenta, sec. 10.4.
Studenmund, review ch. 6; see also ch. 14.

XI Diagnostics, Specification Tests, and Specification Searching

Gujarati & Porter, secs. 5.12, 13.1, 13.6-13.10.

XII Linear Regression Encore (in Matrix Notation)

Gujarati & Porter, app B.
Kmenta, secs. 10.1, 10.2, 12.1.

XIII Nonlinear Models

Gujarati & Porter, secs. 6.3-6.9, 7.10-7.11.
Kmenta, sec. 11.3.
Studenmund, review ch. 7.

XIV Models with Discrete Regressors (including ANOVA and ANCOVA)

Gujarati & Porter, ch. 9.
Kmenta, pp. 460-73.

XV Models for Discrete or Limited Dependent Variables (Introducing Logit, Probit, Tobit, and Poisson Regression)

Gujarati & Porter, ch. 15.
Kmenta, sec. 11.5.
Studenmund, ch. 13.

XVI Some Models for Time-Series: Lagged Variables and Distributed Lags

Gujarati & Porter, ch. 17.
Kmenta, sec. 11.4.

XVII Pooled Time-Series of Cross-Sections and Seemingly Unrelated Regressions

Gujarati & Porter, ch. 16.
Kmenta, secs. 12.2-12.3.

XVIII More Political Science Applications

Jan E. Leighley and Jonathan Nagler. 1992. Socioeconomic Bias in Turnout 1964-1988: The Voters Remain the Same. *American Political Science Review*, 86:725-36.

Benjamin Radcliff. 1992. The Welfare State, Turnout, and the Economy: A Comparative Analysis. *American Political Science Review*, 86: 444-54.

Assignments and Grading

There will be two exams, a series of exercises, and, optionally, a term paper. The exams will be in-class and closed-book. The exercises will be a mix of pen-and-paper and computer-based, the former to help cement the math, the latter, calling for you to write and analyze models of your own, to provide a taste of actual modeling. The optional term paper should propose, justify, detail, estimate, and interpret the results from a regression model of your devising, using data of your choosing. It may build on but must go well beyond the exercises.

For students *not* writing a term paper, each exam will count for 30% of the course grade, and the exercises collectively for 40% (thus each exercise for $2/5n$, where n is the number of exercises). For students writing a term paper, the exams, exercises, and term paper will each count for one-fourth (thus each exercise for $1/4n$).