Pols 707—Fall 2006

September 26, 2006

Professor: Paul E. Johnson

Office: Blake Hall 312, 864-9086

http://pj.freefaculty.org/ps707

Office Hours: MW:1:30-4:00 and by appt.

Course Purposes.

- Learn R
- Learn Statistics
- Learn LaTeX document preparation.

The Social Contract

Work hard every week. Be ready at any time to tell the class what you have learned this week.

Every week you owe me either

1) a response to a specific assignment,

or, if I do not make an assignment,

2) a printout, a computer program, or some other "show and tell" item that you can discuss with the class.

Assignments:

Grades will be based on the writing assignments (80%) and the final paper (20%). The final paper will be a "real article". Throughout the semester, there will be several writing assignments that build up to that final paper. The writing assignments, which are graded, will be brief (5-10 page) project write-ups. I hope you will do your best to turn these in when they are due, but I will not penalize your grades unless they are handed in more than one week late.

Avoid plagiarism or the appearance thereof.

About Readings:

John Fox, Applied Regresion Analysis, Generalized Linear Models, and Related Methods, 2ed, (Forthcoming, Sage)

In case you need other insights, you can find them in many many books. I have ordered some at the bookstore, just to make sure you have access.

(K) Peter Kennedy, *A Guide to Econometrics*, 4th ed. This is a "Cliff's notes" for statistics.

(G) William Green, Econometric Analysis, 4th edition. This is the most respected, general purpose econometrics book. It is more difficult than Gujarati.

Other books that I find useful are

(DG) Dadomar N. Gujarati, *Basic Econometrics*, 4th ed. (New York: McGraw Hill, 2003).

(PR) Pindyck and Rubinfeld, *Econometric Models and Economic Forecasts.* Until 2003, this was the main text for this class.

(VR) Venables and Ripley, *Modern Applied Statistics with S+* (I put a photocopy of that on reserve)

and a simpler introduction to regression book:

(CHP) Samprit Chatterjee, Ali Hadi, and Bertram Price*, Regression Analysis by Example, 3rd ed.* (I put a copy of that on reserve too).

(Faraway) *Practical Regression and Anova using R* This is a FREE, ONLINE book from a U. Michigan professor. I have a copy on reserve, you can copy that, or download your own pdf: http://www.stat.lsa.umich.edu/~faraway/book/

(Maindonald) JH Maindonald, *Using R for Data Analysis and Graphics: An Introduction*. This is a FREE, ONLINE book. I have a copy on reserve, you can copy that, or download your own pdf. Find a current link on the R site (see links below)

Scott Long, *Regression Models for Limited and Categorical Dependent Variables. Beverly Hills: Sage.* This is an extremely fine introduction to probit, logit, and count models.

Weekly Schedule

1 Introduction to R.

R has a huge volunteer workforce and a big homepage too: http://www.r-project.org. In that site, look on the left-hand side for the link "Other" under "Documentation." There is a link to "Contributed Documentation." There is also a free step-by-step R tutorial by Mark Myatt that he calls "Open Source Solutions-R". There's Faraway's book, also free. There is an encyclopaedia of R online stuff here: http://www.vanderbilt.edu/quantmetheval/r.htm. I should mention that the VR book, which is on reserve, is the bible for S+/R and you can learn a lot from it too.

2 Learn R; Review of statistics.

Readings:

- W. N. Venables, D. M. Smith and the R Development Core Team, *Introduction to R*, Chapters 1-8. You can buy this, but also it is free with any version of R and it is also on the R website. In R, run help.start() and read it "Introduction to R." There's a copy on the R website: `http://cran.r-project.org/doc/manuals/R-intro.pdf`
- W. J. Owen, *The R Guide* `http://cran.r-project.org/doc/contrib/Owen-TheRGuide.pdf`. This is part of the Contributed Documentation web site.
- Pick one other item listed on the "Contributed Documentation" section of the R website and study it and make a 5 minute report to the class on it. I've not read all of these, but I know many are great. I learned R from the Emmanuel Paradis, *R for Beginners* `http://cran.r-project.org/doc/contrib/Paradis-rdebuts_en.pdf`.
- If you are desperate to buy something, I'd suggest John Verzani, Using R for Introductory Statistics. Chapman, 2005. This is the most polished of the introductory books.

DO THIS: About book-keeping:

In your account, create a directory "Manuscripts" and whenever you find a pdf or other useful file worth saving, drop it in there.

Exercises: POLS 707 Workbook sections 1-4.

3 Using R to plot functions and do matrix algebra.

DO THIS

1. Visit http://www.r-project.org. Look at CRAN and survey the Contributed Package list.
2. Visit my tip sheet: http://pj.freefaculty.org/R/Rtips.html.
3. Join the r-help email list. I suggest you do that from a Google mail account, and create a folder/filter to collect those messages.

Exercises: POLS 707 Workbook sections 5-8

3.1 Statistical distributions

These are used to describe i) things in the world and ii) estimators you calculate. Note: All of these have "parameters" that determine their shapes.

Readings:

- In the course web site, look in the Distributions directory. You will find handouts like BetaDistribution or GammaDistribution.
- Law and Kelton, "Ch. 4. Basic Review of Probability and Statistics," Simulation Modeling and Analysis, pp. 137-218.
- "Regress+" a 100 page document with statistical distributions: Search the web for the document called "Compendium". Recently, I found it here:

  `http://www.geocities.com/~mikemclaughlin/math_stat/Dists/Compendium.html`
- There is an "on-line gallery of distributions here:
  http://www.itl.nist.gov/div898/handbook/eda/section3/eda366.htm
- This one often has interesting things: http://www.ruf.rice.edu/~lane/stat_sim.
- See also: Jerry Banks, John Carson, II, Barry Nelson, and David Nichol "Ch. 5, Statistical Models in Simulation" *Discrete-Event System Simulation* pp. 153-203.

3.2 Displaying and Exploring data

Readings: Fox, Chapters 2 & 3

3.3 Vectors, matrices, etc.

Readings:

- Søren Højsgaard, *Linear Algebra in R - A Brief Introduction* `http://genetics.agrsci.dk/~sorenh/misc/Rdocs/LinearAlgebra-inR.pdf`
- My lecture notes are available online, "Vectors". There are linear algebra chapters/appendices in almost all statistics books. See Gujarati's Appendix B, pp. 913-925.

4 Introduction to Regression Models

Readings:

- Fox, Chapters 5 & 6
- My notes, which I will hand out, are available online in a file called "OneInputRegression".

Exercises: POLS 707 Workbook sections 9-12

5 Regression Extensions: Nonlinearity

DO THIS: Everybody needs to bring some sample printout of a regression (along with a scatterplot). Write one good paragraph explaining what the estimated standard errors of the parameters "mean" and what the estimated "Residual standard error" means.

5.1 Dummy Variables

Readings:

- Fox, Chapters 7 "Dummy-Variable Regression"

5.2 Curvy NonLinearity: Parametric Estimation of Intrinsically Linear Models

Readings:

Fox, Chapters 4.1 "Family of Powers and Roots", 4.3 "Transforming Nonlinearity", 12.3 "Nonlinearity" 12.5 "Maximum-likelihood Methods"

Gujarati, Chapter 6. Focus on part 6.4-10. Look for: log-linear log-log models reciprocal models

Here are some examples of intrinsically linear models from social science:

5.2.1 Double log of multiplicative models, aka translog (Remember the B's are elasticities!)

$log(y_i) = \beta_0 + \beta_1 log(x_i) + e_i$

James Morrow, Randolph Siverson, Tressa Tabares, "The Political Determinants of International Trade: The Major Powers, 1907-90" APSR 92: 649-661.

Brian Pollins, "Does Trade Still Follow the Flag," APSR, 83 (June 1989): 465-480. Note p. 469

Steven Finkel, et al., "Personal Influence, Collective Rationality, and Mass Political Action," APSR 83 (Sept 1989): 885-903. (see p. 895)

5.2.2 Log on the right $y_i = \beta_0 + \beta_1 log(x_i) + e_i$

Aehra F. Arat, "Democracy and Economic Development: Modernization Theory Revisited," Comparative Politics, 21 (October 1988):21-36.

Robert W. Jackman, "On the Relation of Economic Development to Democratic Performance," AJPS, 17 (Aug 1973), 611-621.

5.2.3 Multiplicative Interaction

John R. Hibbing, "The Media's Role in Public Negativity Toward Congress," AJPS 42 (April 1998):475-498.

Diana Evans, "Oil PACs and Aggressive Contribution Strategies," JOP, 50 (November 1988): 1047-1056

Edward Muller and Michell Seligson, "Inequality and Insurgency," APSR, 81 (June 1987): 425-451.

5.2.4 Polynomial model $y_i = \beta_0 + \beta_1 x_i + \beta x_i^2 + e_i$

John Strate, et al., "Life Span, Civic Development and Voting Participation," APSR, 83 (June 1989):443-463. Pretty interesting adaptation of simple nonlinearity. Clearly presented, worth studying

5.2.5 $log(y_i) = \beta_0 + \beta_1 x + e_i$ Log on the left

Robert J. Thornton and Jon T. Innes, "Interpreting Semilogarithmic Regression Coefficients in Labor Research," Journal of Labor Research, 10 (Fall 1989).

Stephen J. Huxley, "Predicting Response Speed in Mail Surveys," JMR (Feb. 1980): 63-68.

5.3 Curvy NonLinearity: Nonparametric (Loess, Splines, "kernel smoothing" etc)

There's a very nice introductory essay linked to the "locfit" package for S/R. On the web, search for an essay by C. Loader called "Locfit: an introduction". Recently I found that at: http://cm.bell-labs.com/cm/ms/departments/sia/doc/meth.html.ht labs.com/stat/doc/locfitscg.ps

5.4 Non-OLS estimating procedures for Models that are Not Intrinsically Linear:

5.4.1 Maximum Likelihood

Gujarati, Appendix 4A, pp. 114-118

In a Linear Model context, ML is equivalent to OLS because the procedure ends up estimating by minimizing a sum of squared errors

$$\sum (y_i - \hat{y}_i)^2$$

ML thus generates the same point estimates of the β vector as OLS. ML has an advantage, however, in that it can be applied to a larger set of theories (nonlinear, nonNormal, etc).

5.4.2 Nonlinear least squares

Look in any stats book you have handy. Generally, the idea is that you can specify any function for the predicted value of y, say $\hat{y} = f(x, b)$. Here, x is a variable across cases and b is a parameter vector. Then the nonlinear sum of squares is

$$\sum (y_i - f(x_i, b))^2$$

Choose $\hat{b}$ to make the sum of squares as small as possible.

5.5 Generalized Additive Models: extending the smoothers into a multivariate framework

We won't be able to do much work on this, but I can give you a 5 minute explanation to set foundations for future work.

6 More about regression estimates

6.1 Multiple Regression background

Readings:

- Review Fox, Ch. 7, "Dummy Variables"
- Fox, Chapters 9
- My notes are online, called "MultipleRegression1".
- See my essays on this web page:

  http://pj.freefaculty.org/ps707/SamplingDistributionEssays
  These essays were prepared for classes in the early 1990s, but I still think they are pretty good! For background, consider "Sampling Distribution of the Normal Mean," and "The Central Limit Theorem With Illustrations." For regression specifics, see:

  "Simulation of Bivariate Regression." That's online in the SamplingDistributionEssays directory.

If you find yourself fumbling about in ignorance, seek out other things to read.

For a simple introduction, consult Chatterjee, et al, Ch. 2

If Chatterjee is too hard, consult the "little green book" by Michael S. Lewis-Beck, *Data Analysis: An Introduction, "Ch. 6: Simple Regression"*

I make the undergrads in POLS 306 read this: Bowen & Weisberg, *Introduction to Data Analysis*, Ch. 10 "Statistical Inference" because it is very clear.

Faraway gives the matrix algebra treatment of multiple regresion, Ch. 2&3

6.2 Show and Tell

DO THIS: Everybody has to pick an article from a modern political science journal. The dependent variable should be a continuous variable and the model should be estimated by OLS or ML (or some similar method). Bring a "summary sheet" showing the estimated model; make copies for each member class. A summary sheet has the article's title, a succinct description of the model's variables, and a table presenting one of the models estimated. Here are some articles that I'm pretty sure you can handle. If you find some other article, let me know ahead of time so I can review it.

Timothy B. Krebs, "The Determinants of Candidates' Vote Share and the Advantages of Incumbency in City Council Elections," AJPS 42 (July, 1998): 921-935

Sally Coleman Selden, Jeffrey Brudney, J Edward Kellough, "Bureaucracy as a Representative Institution" AJPS 42 (July 1998): 717-744.

Robert Erikson, "Economic Conditions and the Presidential Vote," APSR, 83 (June 1989): 568-573.

7 Multiple Regression

7.1 Practicum: Dummy Variables

DO THIS: Estimate a multiple regression model in which some of the input variables are dummy variables and some are not, and make sure to include at least one interaction term. Prepare a nice looking table summarizing the estimates and write one paragraph about the estimates. Make a graph to illustrate a significant point you want to convey.

Examples

Wendy Rahn, John Aldrich, and Eugene Borgida, Individual and Contextual Variations in Political Candidate Appraisal. APSR, (March 1994) 88:193-199

M. Lewis-Beck and J. Alford, "Can Government Regulate Safety...?" APSR, 1980, pp. 745-756. That's a piecewise linear model

7.2 Diagnostics

Reading:

Fox, Chapter 11 "Unusual and Influential Data".

8 Choosing Variables

8.1 Required Reading on Multicollinearity:

Reading:

Fox, Chapter 12, "Collinearity and its Proported Remedies"

My notes are called "MultipleRegression2-mc"

If that is not understandable to you, consult :

Kennedy, Ch. 11

Chatterjee, Ch. 9, 10.

Greene discusses MC on p. 255-259. I never want to forget his advice, "Suggested 'remedies' for multicollinearity might well amount to attempts to force the theory on the data."(p. 259)

Or Faraway, Chapter 9. That goes into "principal components" and "ridge regression," two approaches to multicollinearity that I've not used, but admire and respect and want to learn.

8.2 R-Square and Standardized Coefficients (betas)

Reading:

IMPORTANT Gary King, "How Not To Lie With Statistics," AJPS, Aug 1986

Robert Luskin, "R-Square Encore," in Political Methodologist, Sp. 1991, pp. 21-23.

My notes are called "MultipleRegression3-betas"

9 Heteroskedasticity and Autoregression.

Reading:

Fox, Ch. 12. Chapter 12.2 is about heteroskedasticity and "weighted least squares"

Fox, Ch. 16, About correcting for Auto-Correlation and "generalized least squares"

Other good treatents you might consult:

Pindyck and Rubinfeld, Ch. 6

Simpler survey is here Kennedy, Ch. 8, 9

Pretty good treatment here Chatterjee, Ch. 7, 8

Very advanced treatment in Greene, Ch. 11-13

Here we discuss of two common statistical problems. If the error term in regression does not follow the assumptions of OLS, corrections need to be made. I try to look at this as a modelling opportunity rather than a smudge on OLS models. In the case of heteroskedstasticity, we say "give the observations that have error terms with higher variance less weight." This class will focus on detection and treatment of AR(1) processes in practice– Cochrane-Orcutt procedure and variants of it.

1. EXAMPLE: WLS to adjust for "group means" type dependent variable data.
   C. Neal Tate and Pany Sittiwong, "Decision Making in the Canadian Supreme Court: Extending the Personal Attributes Model Across Nations," JOP, 51 (November 1989): 900-916 (esp. p. 908, fn. 7)
2. EXAMPLE: AR(1)
   AR(1) in practice: The Cochrane-Orcutt procedure.

   Michael S. Lewis-Beck, "Economic Conditions and Executive Popularity: The French Experience," AJPS 24 (May 1980): 306-323.

10 Qualitative Variables I:

There are many great things to read, and there is so little time.

Readings:

Fox, Ch. 14, “Logit and Probit Models”

Long: Ch. 3, Binary Outcomes.

Other useful treatments:

Gujarati, Chapter 15, pp. 580-616

I still love this chapter: PR, Ch. 11.1 "Models of Qualititative Choice", pp. 298-318.

And I think this is prett good too: Aldrich and Nelson, *Linear Probability, Logit and Probit Models*, pp. ch 1.0-1.3 (pp. 9-22), 1.5 (pp. 24-30). CH. 2.0-2.2 (pp. 30-35).

Maybe this will help: Kennedy, Ch. 15.

Or this: Chatterjee, Ch. 12.

Or: Hanushek and Jackson, Chpt. 7 "Models With Discrete Dependent Variables." in Statistical Models in the Social Sciences.

Or: Gary King, Unifying Political Methodology: The Likelihood Theory of Statistical Inference (Cambridge: Cambridge U. Press, 1989).

11 More general development of Logit and Probit models.

11.1 Maximum likelihood (individual level) approach

Readings:

Abdelmonem Afifi, Virginia Clark, and Susanne May, Chapter 12, "Logistic Regression," in *Computer-Aided Multivariate Analysis,* 4ed (Chapman and Hall, 2004) , pp. 281-329.

Long, Ch. 4

Aldrich and Nelson, *Linear Probability, Logit and Probit Models*, Ch. 2.3, pp. 35-37, Ch. 3, pp. 48-66

Others:

To more deeply understand why these models work, one should look into the theory of Maximum Likelihood estimation. Chapter 10 in PR does a pretty good job. Other descriptions of ML can be found in many texts, including Maddala's Econometrics, 1977; Bornstadt and Knoke, Statistics For Social Data Analysts, 1987; King's Unifying Political Methodology.

When in doubt on the technical issues, I consult a great text by Ed Greenberg and Charles Webster, Advanced Econometrics: A Bridge to the Literature, New York: Wiley, 1983.

11.2 Grouped Data (so-called minimum chi square methods)

Aldrich and Nelson, *Linear Probability, Logit and Probit Models*, Ch. 4.0-4.1

Others:

Charyl L. Maranto, "Corporate Characteristics and Union Organizing," Industrial Relations, 27 (Fall, 1988). A stock application of minimum chi square logit analysis to grouped data.

11.3 Compare a & b.

Required Reading: David Flath and E.W. Leonard, "A Comparison of Two Logit Models in the Analysis of Qualitative Marketing Data," JMR 16 (Nov. 1979), pp. 533-538.

11.4 Multicategory dependent variables (ordered and unordered).

Long Ch. 5- 6

PR, Ch. 11.2

Aldrich and Nelson, *Linear Probability, Logit and Probit Models*, 1.4, 2.4, 4.2

Others:

An authoritative source on all variations of the qualitative variables problem is G.S. Maddala, Limited-dependent and qualitiative variables in econometrics. (Cambridge U. Press, 1984)

12 Applications & Interpretation of qualitative models.

What kinds of diagnostic information do these models provide? How are the results interpreted?

DO THIS: Everybody has to bring printout of a logit model that we can discuss. Make copies of any necessary plots and results for everybody.

12.1 How to interpret these models:

Readings:

Long covers this in his chapters. Look that over.

Gary King, Unifying Political Methodology. Ch. 5.1-5.2,pp. 98-110 (Photocopy)

See also: Aldrich and Nelson, *Linear Probability, Logit and Probit Models*,2.5-2.5.2 pp. 40-44.

12.2 Examples: LOGISTIC regression.

DO THIS: Dig one of these out (JSTOR?) and examine the way they present the parameter estimates and discuss them. Make a "summary sheet" to bring to class.

Lori Hausegger and Lawrence Baum, "Inviting Congressional Action: A Study of Supreme Court Motivation in Statutory Interpretation," AJPS 43 (January, 1999): 162-185

Lonna R. Atkeson and Randall W. Partin, "Economic and Referendum Voting: A Comparison of Gubernatorial and Senatorial Elections," APSR, 89 (March 1995): 99-106

Larry Bartels, "Candidate Choice and the Dynamics of the Presidential Nominating Process," AJPS, Feb. 1987, excerpt pp. 1-18. See p. 16 and the use of an interaction term in a logit model. Ask yourself, why does he use ordinary regression sometimes and logit in others.

Donald Kinder, et al, "Economics and Politics in the 1984 American Presidential Election," AJPS (May 89): 491-515.

R. Robert Huckfeldt, "Political Loyalties and Social Class Ties," AJPS, vol. 28, May 1984, pp. 399-417.

Jack Wright, "PACs, Contributions, and Roll Calls: An Organizational Perspective," APSR, (June 1985) 79: 400-414.

John Zipp, "Perceived Representativeness and Voting: An Assessment of the impact of 'choices' vs. 'echoes'," APSR, 79 (March 1985): 50-62.

12.3 Examples: PROBIT regression (look at any of these):

Carole Kennedy Cahney and Grace Hall Saltzstein, "Democratic Control and Bureaucratic Responsiveness: The Police and Domestic Violence," AJPS 42 (July, 1998): 745-768.

Dean Lacy and Philip Paolino, "Downsian Voting and the Separation of Powers," AJPS 42 (October 1998): 1180-1199

Timothy Johnson and Andrew Martin, "The Public's Conditional Response to Supreme Court Decisions," APSR 92 (June 1988): 299-309. (uses LR tests!)

Richard L. Hall and Robert P. VanHouWeling, "Avarice and Ambition in Congress" APSR 89: 121-136.

Charles H. Franklin and Liane Kosaki, "Republican Schoolmaster: The U.S. Supreme Court, Public Opinion, and Abortion," APSR 83 (Sept. 1989): 751-771. (Photocopy) Includes a test of a "cross equation constraint." Very clear explanation of the McKelvey-Zaviona model

John Aldrich, John Sullivan, Eugene Borgida, "Foreign Affairs and Issue Voting: Do Presidential Candidates 'Waltz Before a Blind Audience," APSR, 83 (March 1989): 124-141

Robert Luskin, John McIver, and Edward Carmines, "Issues and the Transmission of Partisanship," AJPS (May 1989) 33: 440-458

Charles Ostrom, Jr. and Brian Job. 1986. The President and the Political Use of Force. APSR 80: 541-566. Pretty Good description of the model.

Jeffrey Segal, "Senate Confirmation of Supreme Court Justicies:..." Journal of Politics, v. 49, Nov. 1987, pp. 998-1016.

Paul Abramson, et al. "Progressive Ambition among United States Senators: 1972- 1988," JOP, v. 49, Feb. 1987, pp. 3-35.

Paul Brace, "Progressive Ambition in the House: A Probabilistic Approach," JOP, v. 46, May 1984, pp. 556-546.

Eric Uslaner and M. Conway. 1985. The Responsive Congressional Electorate: Watergate, the Economy, and Vote Choice in 1974. APSR, 79: 788-803.

12.4 (OPTIONAL TOPIC): Multi-category unordered response models. (multinomial logit)

Patrick Sellers, "Strategy and Background in Congressional Campaigns," APSR 92 (March, 1998): 159-171

13 Count Models

Reading:

Fox, Ch. 15, "Generalized Linear Models"

For your information: There is a very fine book dedicated to a survey of models for count data. Any count-model-users should check it out:

Cameron, A. Colin, and Pravin K. Trivedi. 1998. *Regression Analysis of Count Data.* New York: Cambridge University Press.

13.1 Poisson Regression

Reading:

Gujarati, Ch 15.12, pp. 620-622.

Gary King. 1988. Statistical Models for Political Science Event Counts: Bias in Conventional Procedures and Evidence for The Exponential Poisson Regression Model. American Journal of Political Science, 32(3):838-863.

13.2 Negative Binomial and other extensions

Reading:

Long (There's a chapter on Poisson, Negative Binomial, and "zero inflated" models).

See this WebPage:

Hun Myoung Park, "Regression Models for Event Count Data Using SAS, STATA, and LIMDEP" http://www.indiana.edu/~statmath/stat/all/count/count.html

14 Multi-equation systems.

What's Wrong with OLS? Simple "walk-through" of 2SLS and instrumental variables, kinds of multi-equation systems.

Reading:

I'm afraid none of the stats books are very clear and helpful on this. Fox avoids it entirely. There are chapters in most econometrics book, including

Gujarati, Ch. 18-20.

PR, Ch 12, pp. 287-305

Kennedy, Ch. 10

Dig around until you find something that is remotely understandable to you.

Examples:

Alan Gerber, "Estimating the Effect of Campaign Spending on Senate Election Outcomes Using Instrumental Variables," APSR 92 (June 1998): 401-411

Richard B. Freeman and Morris M. Kleiner, "Employer Behavior in the Face of Union Organizing Drives," Industrial and Labor Relations Review 43 (April 1990): 351-365.

Class time is limited, but your future work should probably touch on these two additional methods.

1. There are systems methods for models with qualititative endogenous variables. Examples of TWO STAGE PROBIT:

Gregory Caldeira and John Wright, "Lobbying for Justice" AJPS 42 (April 1988): 499-523

Mark Peffley, et al. Economic Conditions and Party Competance: Processes of Belief Revision. JOP, Feb. 1987, pp. 100-122.

Charles Franklin and John Jackson. The Dynamics of Party Identification. 1983. APSR, 77: 957-973.

Charles Franklin and John Jackson. Article in Weisberg's Political Science: The Science of Politics (1986).

Harold Clarke and Marianne Stewart, "Dealignment of Degree: Partisan Change in Britain, 1974-83," JOP, vol. 46, pp. 689-719

15 Missing Values

Readings:

Fox, Ch. 20

This is pretty complicated, and so is this famous article:

Gary King, James Honaker, Anne Joseph and Kenneth Scheve. "Analyzing Incomplete Political Science Data," *American Political Science Review*, Vol. 95, No. 1, (March, 2001): Pp. 49–69. Online: http://gking.harvard.edu/files/evil.pdf

If you want something readable, perhaps you should look here:

Schafer, J. and Olsen, M. (1997), Multiple imputation for multivariate missing-data problems: a data analyst's perspective, *Multivariate Behavioural Research*, Vol. 33, pp. 545-571. also available from http://www.stat.psu.edu/~jls. Schafer offers packages for S/R.

There is a "little green book" by Paul Allison that extends on some of his observations in this argicle:

Allison, P. (2000). Multiple imputation for missing data: A cautionary tale. *Sociological methods and research*, 28 (3), 301-309. `http://www.ssc.upenn.edu/~allison/MultInt99.pdf`.

16 Pooled Cross-Sectional Time Series (Panel Data)

Reading:

Gujarati, Ch. 16

I have prepared many handouts on this topic. A key starting point is "Generalized Least Squares." After that, you can move on to consider the special variations to deal with CX-TS data.

This article brought the methodological issues (actually, the entire field of "political methodology," to the forefront).

Neil Beck and Jonathan N. Katz. 1995. What to Do (and What Not to Do) with Time-Series Cross-Section Data. *American Political Science Review*, Vol. 89, No. 3 (September): 634-647.