PLSC 504: "Topics in Political Methodology"

Fall 2019

Professor Christopher Zorn
Department of Political Science
Pennsylvania State University
E-mail: zorn@psu.edu
Monday & Wednesday, 9:00 - 10:30 a.m.
Rackley Building, Room 107

Course Description

This is the third (full) course in quantitative methods in Penn State's political science Ph.D. program. The course attempts to do two things. First, it is a brief overview of regression-like statistical methods, with some emphasis on likelihood-based models, including binary logit and probit, multinomial logit and probit, ordered logit and probit, and Poisson and other models for event counts. We will also introduce models for time-series data, survival (time-to-event) data, and panel and time-series cross-sectional data. Second, the course is the "pros-eminar" for the methods field in the department; this means that it is designed to introduce students to topics that they then can learn about in greater depth in other courses. These currently include multivariate models for data reduction and the measurement of latent concepts, network analysis, and approaches for analyzing text.

The models discussed in this course are among the most widely used in the social sciences today. It is not possible to function as an empirical social scientist without at least a passing familiarity with these models; moreover, given the rapid and increasing rate at which more advanced models are being adopted in these fields, these techniques increasingly represent a minimal level of statistical competence necessary to do publishable-quality quantitative work. In other words: knowing these models, and using them appropriately and well, can increase your odds of writing a strong (quantitative) dissertation, landing a job, publishing books and articles, being granted tenure, and generally leading a happier and more fulfilling professional life.

Much of the material in this course is fairly technical. While I have chosen readings that present the models as clearly and with as little jargon as possible, most of the material will still require several readings to fully comprehend. A solid understanding of calculus and linear algebra is required for this class, and the course assumes familiarity with linear regression at the level of PLSC 503 (that is, at the level of Weisberg's *Applied Linear Regression*, Greene's *Econometric Analysis*, or the like). Students are also expected to have at least a nodding acquaintance with probability theory, statistical inference, and simple descriptive statistics, as well as with data management.

This syllabus is designed to provide an overview to the course. Clickable links are printed in Penn State blue.

Course Readings

Recommended Text/Materials

Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage Publications.

AND/OR

Faraway, Julian J. 2006. *Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression*. London: Chapman & Hall.

AND

Box-Steffensmeier, Janet M., and Bradford S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists*. New York: Cambridge University Press.

AND

Everitt, Brian, and Thorsen Hothorn. 2011. *An Introduction to Applied Multivariate Analysis with R*. New York: Springer

The course will also draw on additional readings as necessary, all of which will be available on the course github repository and/or through JSTOR. For the required texts, you should feel free to order them from whatever sources you deem best.

A Few Other Potentially Useful Readings

... can be found in a list, here.

Some Other Useful Resources

The **Political Methodology Section** of the American Political Science Association was created to provide APSA members with an interest in political methodology with a forum in which to meet and discuss ideas. The section publishes a quarterly newsletter (*The Political Methodologist*), a quarterly journal on political methodology (*Political Analysis*), conducts a discussion list on topics relating to political methodology, and maintains an extensive electronic archive of papers, accessible via their homepage.

Also, the **Inter-University Consortium for Political and Social Research** (ICPSR), at the University of Michigan, maintains an extensive archive of data in the social and behavioral sciences. Much of it is accessible via their homepage.

Grading

Grading will be based on ten(ish) more-or-less weekly homework exercises (50 points each) and a final paper/poster presentation (500 points). In most instances, exercises will be due one week from being assigned. Homework exercises will generally involve both simulation-based work and estimation and interpretation of models on real/existing data, using statistical computer software (see below). Feel free to work on the assignments in groups of two or three, but you must write up all assignments individually. Details for the homework assignments and the final project will be announced in class.

Also, note that homework exercises and the final paper/poster should be submitted as electronic files, **in Adobe PDF format**. If you do not know how to create a PDF file, please go learn, now.

Software, Statistical and Otherwise

You are welcome to make use of whatever statistical software you choose to complete the homework exercises, so long as the manner by which your results are generated and conclusions reached are transparent. However,

due to the limits of instructor and TA time and patience, we will support only two software packages. Both are available on the machines in the political science computing labs.

R

R is a statistical environment and high-level programming language for data analysis and display. It is effectively the GNU version of the S language; as such, it is free (both as in speech and as in beer) and open source. The current (late August 2019) version of R is 3.6.1. R is an *object-oriented* language; unlike Stata (and most other statistical packages), it operates by assigning values to objects in the workspace. In the notes, handouts, etc., R commands will be preceded by a caret (">"):

```
> my.results<-lm(Y~X)
```

The Comprehensive R Archive Network (CRAN) is the go-to spot for all things R-related. I cannot begin to list all the R-related resources available on the web; for newbies, however, it might be useful to check out the Introduction to R, this page in getting data into R, and the various R "cheat sheets" here, here, and here. Stata users who are interested in learning R should check out the Moving from Stata to R page at the R Project's wiki.

All in-class examples, code, graphics, and so forth will use R.

Stata

At the present time (but not for long), Stata is probably the most widely-used statistical package in the social sciences. It is a powerful tool for data management, analysis, and display, and boasts some of the best manuals and on-line help of any existing software package. Stata is commercial software; the current version of Stata is 16.0, but previous versions (back to v. 12, at least) can also be used for the class. In the rare instance when they appear in the class notes, handouts, etc., Stata commands will appear in a fixed-width font and will be preceded by a period (."):

```
. regress Y X
```

Stata newbies may want to check out:

Getting Started with Stata for Windows, Release 15. 2017. College Station, TX: Stata Press.

and/or Stata's dedicated "new users" page:

https://www.stata.com/links/resources-for-learning-stata/.

Beyond this, the Stata homepage is a valuable resource for questions about the Stata statistical software. There are a number of useful Stata references on the web, including Scott Long's page at IU and an excellent Stata "help page" sponsored by UCLA.

Other Considerations

In no particular order:

- Your instructor does not have a formally-stated preference for either Stata or R. My recommendation would be to learn to use both, as each has its strengths and weaknesses. Stata has a far flatter learning

curve than R, which means students tend to gravitate toward it given a choice. But R is far more flexible and powerful, and will likely be more useful to you in the long run.

- Learn to use LaTeX, now, while you have the time. You will be glad you did.
- If you insist on using Microsoft Word (or any other WYSIWYG program) for writing assignments, papers, etc., **do not under any circumstances cut and paste graphs from Stata and R into those programs**. Save whatever figures you want to use as `.pdf`, `.png`, `.tif`, or `.jpg` files, and import them into the software.

Obligatory Statement on Academic Integrity

Academic integrity is the pursuit of scholarly activity in an open, honest and responsible manner. Academic integrity is a basic guiding principle for all academic activity at The Pennsylvania State University, and all members of the University community are expected to act in accordance with this principle. Consistent with this expectation, the University's Code of Conduct states that all students should act with personal integrity, respect other students' dignity, rights and property, and help create and maintain an environment in which all can succeed through the fruits of their efforts.

Academic integrity includes a commitment by all members of the University community not to engage in or tolerate acts of falsification, misrepresentation or deception. Such acts of dishonesty violate the fundamental ethical principles of the University community and compromise the worth of work completed by others.

In cases of any violation of academic integrity it is the policy of the Department of Political Science to follow procedures established by the College of the Liberal Arts. More information on academic integrity and procedures followed for violation can be found here.

Obligatory Statement on Accommodations for Disabilities

Penn State welcomes students with disabilities into the University's educational programs. Every Penn State campus has an office for students with disabilities. Student Disability Resources (SDR) website provides contact information for every Penn State campus (`http://equity.psu.edu/sdr/disability-coordinator`). For further information, please visit the Student Disability Resources website (`http://equity.psu.edu/sdr/`).

In order to receive consideration for reasonable accommodations, you must contact the appropriate disability services office at the campus where you are officially enrolled, participate in an intake interview, and provide documentation: See documentation guidelines at `http://equity.psu.edu/sdr/guidelines`. If the documentation supports your request for reasonable accommodations, your campus disability services office will provide you with an accommodation letter. Please share this letter with your instructors and discuss the accommodations with them as early as possible. You must follow this process for every semester that you request accommodations.

Obligatory Statement on Counseling and Psychological Services

Many students at Penn State face personal challenges or have psychological needs that may interfere with their academic progress, social development, or emotional wellbeing. The university offers a variety of confidential services to help you through difficult times, including individual and group counseling, crisis intervention, consultations, online chats, and mental health screenings. These services are provided by staff who welcome all

students and embrace a philosophy respectful of clients' cultural and religious backgrounds, and sensitive to differences in race, ability, gender identity and sexual orientation.

Counseling and Psychological Services at University Park (CAPS)
(http://studentaffairs.psu.edu/counseling/): 814-863-0395

Counseling and Psychological Services at Commonwealth Campuses
(http://senate.psu.edu/faculty/counseling-services-at-commonwealth-campuses/)

Penn State Crisis Line (24 hours / 7 days/week): 877-229-6400. Crisis Text Line (24 hours / 7 days/week): Text LIONS to 741741.

Obligatory Statement on Educational Equity and Reporting Bias

Penn State takes great pride to foster a diverse and inclusive environment for students, faculty, and staff. Consistent with University Policy AD29, students who believe they have experienced or observed a hate crime, an act of intolerance, discrimination, or harassment that occurs at Penn State are urged to report these incidents as outlined on the University's Report Bias webpage (http://equity.psu.edu/reportbias/).

Course Schedule

Readings should be completed prior to coming to class on the assigned day. Note that we will not, in general, hew closely (or at all) to the readings themselves, other than topically. Links are generally to DOIs or to stable PDFs at JSTOR.

As a rule, we will cover one broad topic per week. Readings will be assigned weekly. We will typically spend Monday covering statistical / theoretical topics, and Wednesday delving into practicalities (software, applications, diagnostics, etc.). It will generally be a good idea to have completed the assigned readings prior to Monday's class each week.

Part One: Nonlinear Models and GLMs

August 26: Introduction and Overview

- **Readings**
  - *Required*:
    - None.
  - *Recommended*:
    - None. (Read Long, Chapter 1, for background, if you're so inclined.)

August 28: Overview / Review: Maximum Likelihood

- **Readings**
  - *Required*:
    - Long, Chapter 2 (esp. pp. 25-33), pp. 52-61, and Chapter 4.
    - Faraway, pp. 279-285.

        - Buse, A. 1982. "The Likelihood Ratio, Wald, and Lagrange Multiplier Tests: An Expository Note." *The American Statistician* 36(3):153-57.
    - *Recommended*:
        - Eliason (1993), pp. 1-28.
        - Greene (2003), §17.4.
        - King (1989), Chapter 4.
        - Breusch, T. S. 1979. "Conflict Among Criteria for Testing Hypotheses: Extensions and Comments." *Econometrica* 47(1):203-07.
        - Greene (2003), pp. 484-496.
        - Meeker, William Q. and Luis A. Escobar. 1995. "Teaching About Approximate Confidence Regions Based on Maximum Likelihood Estimation." *The American Statistician* 49(1):48-53.

September 2: No Class – Labor Day

September 4: Models for Binary Responses

- **Readings**
    - *Required*:
        - Long, pp. 34-52, 61-112.
        - Faraway, pp. 25-38.
    - *Recommended*:
        - Aldrich and Nelson (1984), pp. 9-30.
        - Eliason, pp. 39-45.
        - Greene (2003), pp. 665-680.
        - Griffiths, William E., R. Carter Hill, and Peter J. Pope. 1987. "Small Sample Properties of Probit Model Estimators." *Journal of the American Statistical Association* 82(399):929-37.
        - King (1989), pp. 97-114.
        - Nagler, Jonathan. 1994. "Scobit: An Alternative Estimator to Logit and Probit." *American Journal of Political Science* 38(1):230-55.
        - Berry, William D., Jacqueline H. R. DeMeritt, and Justin Esarey. 2010. "Testing for Interaction in Binary Logit and Probit Models: Is a Product Term Essential?" *American Journal of Political Science* 54(January): 248-66.
        - Hagle, Timothy M., and Glenn E. Mitchell. 1992. "Goodness of Fit Measures for Probit and Logit." *American Journal of Political Science* 36(3):762-84.
        - Herron, Michael C. 2000. "Postestimation Uncertainty in Limited Dependent Variable Models." *Political Analysis* 8(1):83-98.
        - King, Gary, Michael Tomz, and Jason Wittenberg. 2000. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." *American Journal of Political Science* 44(2):347-61.
- **Exercise One: Estimate and interpret binary logit and probit models.**

September 9-11: Nominal and Ordinal Responses

- **Readings**
  - *Required*:
    - Long, pp. 114-186.
    - Faraway, pp. 97-112.
  - *Recommended*:
    - Alvarez. R. Michael, and John Brehm. 1998. "Speaking in Two Voices: American Equivocation about the Internal Revenue Service." *American Journal of Political Science* 42(2):418-52.
    - Alvarez, R. Michael, and Jonathan Nagler. 1998. "When Politics and Models Collide: Estimating Models of Multiparty Elections." *American Journal of Political Science* 42(1):55-97.
    - Fry, Tim R., and Mark N. Harris. 1998. "Testing for Independence of Irrelevant Alternatives: Some Empirical Results." *Sociological Methods and Research* 26(3):401-23.
    - Greene (2003), pp. 724-28.
    - Quinn, Kevin M., Andrew D. Martin, and Andrew B. Whitford. 1999. "Voter Choice in Multi-Party Democracies: A Test of Competing Theories and Models." *American Journal of Political Science* 43(4):1231-47.
    - Dow, Jay K., and James W. Endersby. 2004. "Multinomial Probit and Multinomial Logit: A Comparison of Choice Models for Voting Research." *Electoral Studies* 23(1):107-22.
    - Glasgow, Garrett. 2001. "Mixed Logit Models for Multiparty Elections." *Political Analysis* 9(2):116-36.
    - Gelpi, Christopher. 1997. "Crime and Punishment: The Role of Norms in Crisis Bargaining." *American Political Science Review* 91(2):339-60.
    - Jones, Bradford S., and Michael E. Sobel. 2000. "Modeling Direction and Intensity in Semantically Balanced Ordinal Scales: An Assessment of Congressional Incumbent Approval." *American Journal of Political Science* 44(1):174-85.
    - Sanders, Mitchell S. 2001. "Uncertainty and Turnout." *Political Analysis* 9(1):45-57.
    - Liao (1994), pp. 25-469.
    - Winship, Christopher, and Robert D. Mare. 1984. "Regression Models with Ordinal Variables." *American Sociological Review* 49(4):512-25.
    - Greene (2003), pp. 723-24.
    - Whitten, Guy B., and Harvey Palmer. 1996. "Heightening Comparativists' Concerns for Model Choice: Voting Behavior in Great Britain and the Netherlands." *American Journal of Political Science* 40(1):231-60.
- **Exercise Two: Estimate and interpret ordered and unordered logit and probit models.**

September 16-18: Event Counts

- **Readings**
  - *Required*:
    - Long, pp. 217-250.
    - Faraway, pp. 55-66.

        - Zorn, Christopher. 1998. "An Analytic and Empirical Examination of Zero-Inflated and Hurdle Poisson Specifications." *Sociological Methods and Research* 26(3):368-400.
    - *Recommended*:
        - Cameron and Trivedi (1998), Chapter 3.
        - Corpas?Burgos, Francisca, Gonzalo Garcia?Donato, and Miguel A. Martinez?Beneito. 2018. "Some Findings on Zero?Inflated and Hurdle Poisson Models for Disease Mapping." *Statistics in Medicine*, forthcoming.
        - Gowa, Joanne. 1998. "Politics at the Water's Edge: Parties, Voters and the Use of Force Abroad." *International Organization* 52(2):307-24.
        - King, Gary. 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):838-63.
        - King, Gary. 1989. "Variance Specification in Event Count Models: From Restrictive Assumptions to a Generalized Estimator." *American Journal of Political Science* 33(3):762-84.
        - King, Gary, and Curtis Signorino. 1996. "The Generalization in the Generalized Event Count Model, With Comments on Achen, Amato, and Londregan." *Political Analysis* 6(1):225-52.
        - Liao (1994), pp. 70-79.
        - King, Gary. 1989. "Event Count Models for International Relations: Generalizations and Applications." *International Studies Quarterly* 33:123-47.
        - Sheingate, Adam D. 2006. "Structure and Opportunity: Committee Jurisdiction and Issue Attention in Congress." *American Journal of Political Science* 50(October):844-59.

- **Exercise Three: Estimate and compare some event count models.**

September 23-25: Models for Sample Selection

- **Readings**
    - *Required*:
        - Berk, R. A. 1983. "An Introduction to Sample Selection Bias in Sociological Data." *American Sociological Review* 48(June):386-398.
        - Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47(January):153-161.
    - *Recommended*:
        - Berinsky, Adam J. 1999. "The Two Faces of Public Opinion." *American Journal of Political Science* Vol. 43(October):1209-1230.
        - Leeman, Lucas. 2014. "Strategy and Sample Selection: A Strategic Selection Estimator." *Political Analysis* 22(3):374-397.
        - Reed, William. 2000. "A Unified Statistical Model of Conflict Onset and Escalation." *American Journal of Political Science* 44(January):84-93.
        - Sartori, Anne E. 2003. "An Estimator for Some Binary-Outcome Selection Models Without Exclusion Restrictions." *Political Analysis* 11(2):111-138.
        - Sigelman, Lee, and Langche Zeng. 1999. "Analyzing Censored and Sample-Selected Data with Tobit and Heckit Models." *Political Analysis* 8(2):167-82.
        - Vella, Francis. 1998. "Estimating Models with Sample Selection Bias: A Survey." *The Journal of Human Resources* 33:127-169.

Part Two: Longitudinal Data

September 30 - October 2: Introduction to Time Series Analysis

- **Readings**
  - *Required*:
    - TBA
  - *Recommended*:
    - TBA
- **Exercise Four: Fit and interpret some time series models.**

October 7: Survival: Introduction and Parametric Models

- **Readings**
  - *Required*:
    - Box-Steffensmeier, Janet M., and Bradford S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists*, Chapter 3.
    - Alt, James, and Gary King. 1994. "Transfers of Governmental Power: The Meaning of Time Dependence." *Comparative Political Studies* 27(2):190-210.
  - *Recommended*:
    - Bennett, D. Scott, and Allan C. Stam III. 1996. "The Duration of Interstate Wars." *American Political Science Review* 90(June):239-57.
    - Bueno de Mesquita, Bruce, and Randolph M. Siverson. 1995. "War and the Survival of Political Leaders: A Comparative Study of Regime Types and Political Accountability." *American Political Science Review* 89(2):841-55.
    - McCarty, Nolan and Rose Razaghian. 1999. "Advice and Consent: Senate Responses to Executive Branch Nominations." *American Journal of Political Science* 43(October):1122-43.
    - Teachman, Jay D., and Mark D. Hayward. 1993. "Interpreting Hazard Rate Models." *Sociological Methods and Research* 21(February):340-71.

October 9: Survival: The Cox Model (and Extensions)

- **Readings**
  - *Required*:
    - Box-Steffensmeier, Janet M., and Bradford S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists*, Chapters 4-5 and 8-11.
  - *Recommended: Discrete-Time Approaches*:
    - Alt, James E., Gary King and Curtis S. Signorino. 2001. "Aggregation Among Binary, Count and Duration Models: Estimating the Same Quantities from Different Levels of Data." *Political Analysis* 9(Winter):21-44.
    - Beck, Nathaniel, Jonathan N. Katz, and Richard Tucker. 1998. "Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable." *American Journal of Political Science* 42(October):1260-88 (and erratum).

- Lindsey, J. K. 1998. "Counts and Times to Events." *Statistics in Medicine* 17:1745-51.
- Signorino, Curt, and David Carter. 2010. "Back to the Future: Modeling Time Dependence in Binary Data." *Political Analysis* 18(3):271-292. Also read response by Beck and rejoinder by Signorino & Carter.
- Singer, Judith D., and John B. Willett. 1993. "It's About Time: Using Discrete-Time Survival Analysis to Study Duration and the Timing of Events." *Journal of Educational Statistics* 18(Summer):155-95.

◦ *Recommended: Proportional Hazards Models*:

- Box-Steffensmeier, Janet M., and Christopher Zorn. 2001. "Duration Models and Proportional Hazards in Political Science." *American Journal of Political Science* 45(October):951-67.
- Cox, David Roxbee. 1972. "Regression Models and Life Tables." *Journal of the Royal Statistical Society, Series B* 34(2):187-220.
- Desmarais, Bruce A., and Jeffrey J. Harden. 2012. "Comparing Partial Likelihood and Robust Estimation Methods for the Cox Regression Model." *Political Analysis* 20(1):113-135. DOI:10.1093/pan/mpr042
- Grambsch, Patricia M., and Terry M. Therneau. 1994. "Proportional Hazards Tests and Diagnostics Based on Weighted Residuals." *Biometrika* 81(3):515-26.
- Grambsch, Patricia M., Terry M. Therneau, and Thomas R. Fleming. 1995. "Diagnostic Plots to Reveal Functional Form of Covariates in Multiplicative Intensity Models." *Biometrics* 51(December):1469-82.
- Idris, Muhammad, and Christopher Zorn. 2018. "Proportional Hazards Analysis of Survival Data with Tied Survival Times: Theory and Best Practices." Working paper: Pennsylvania State University.
- Keele, Luke J. 2010. "Nonproportionally Difficult: Testing for Nonproportional Hazards In Cox Models." *Political Analysis* 18:189-205.
- Licht, Amanda A. 2011. "Change Comes with Time: Substantive Interpretation of Nonproportional Hazards in Event History Analysis." *Political Analysis* 19(2):227-243.

◦ *Recommended: Repeated Events*:

- Box-Steffensmeier, Janet M., and Christopher Zorn. 2002. "Duration Models for Repeated Events." *Journal of Politics* 46(November):1069-94.
- Box-Steffensmeier, Janet M., Suzanna Linn, and Corwin D. Smidt. 2014. "Analyzing the Robustness of Semi-Parametric Duration Models for the Study of Repeated Events." *Political Analysis* 22:183-204.
- Cleves, Mario. 1999. "Analysis of Multiple Failure-Time Data with Stata." *Stata Technical Bulletin* 49:30-39.
- Kelly, Patrick J. and Lynette L-Y. Lim. 2000. "Survival Analysis for Recurrent Event Data." *Statistics in Medicine* 19:12-33.
- Metzger, Shawna K., and Benjamin T. Jones. 2016. "Surviving Phases: Introducing Multistate Survival Models." *Political Analysis* 24(4):457-477.
- Wei, L. J. and David V. Glidden. 1997. "An Overview of Statistical Methods for Multiple Failure Time Data in Clinical Trials." *Statistics in Medicine* 16:833-39.

◦ *Recommended: Competing Risks*:

- Crowder, Martin. 2012. *Multivariate Survival Analysis and Competing Risks*. New York: Chapman & Hall/CRC.
- David, H. A., and M. L. Moeschberger. 1978. *The Theory of Competing Risks*. New York: MacMillan.

- Diermeier, Daniel, and Randy T. Stevenson. 1999. "Cabinet Survival and Competing Risks." *American Journal of Political Science* 43(4) October: 1051-68.
- Gordon, Sanford C. 2002. "Stochastic Dependence in Competing Risks." *American Journal of Political Science* 46(January):200-17.
- Pintilie, Melania. 2007. "Analysing and Interpreting Competing Risk Data." *Statistics in Medicine* 26:1360-67.
- Wolbers, Marcel, et al. 2014. Competing Risks Analyses: Objectives and Approaches. *European Heart Journal*.
- Zorn, Christopher and Steven R. Van Winkle. 2000. "A Competing Risks Model of U.S. Supreme Court Vacancies, 1789-1992." *Political Behavior* 22(June):145-66.

◦ *Recommended: Duration Dependence*:

- Heckman, James J. 1991. "Identifying the Hand of the Past: Distinguishing State Dependence from Heterogeneity." *American Economic Review* 81(May):75-79.
- Warwick, Paul. 1992. "Rising Hazards: An Underlying Dynamic of Parliamentary Government." *American Journal of Political Science* 36(November):857-76.
- Zorn, Christopher. 2000. "Modeling Duration Dependence." *Political Analysis* 8(Autumn): 367-380.

◦ *Recommended: Cure Models*:

- Box-Steffensmeier, Janet M., Roman Ivanchenko, and Christopher Zorn. 2006. "Cure Models for Political Science Research." Working paper: Ohio State University.
- Box-Steffensmeier, Janet M., Peter Radcliffe, and Brandon Bartels. 2005. "The Incidence and Timing of PAC Contributions to Incumbent U.S. House Members, 1993-94." *Legislative Studies Quarterly* 30(November):549-79.
- Hettinger, Virginia, and Christopher Zorn. 2005. "Explaining the Incidence and Timing of Congressional Responses to the U.S. Supreme Court." *Legislative Studies Quarterly* 30(February):5-28.
- Maller, R. A. and S. Zhou. 1996. *Survival Analysis with Long-Term Survivors*. New York: Wiley.
- Schmidt, Peter and Anne D. Witte. 1989. "Predicting Recidivism Using 'Split-Population' Survival Time Models." *Journal of Econometrics* 40(1):141-59.
- Tsodikov, A. 1998. "A Proportional Hazards Model Taking Account of Long Term Survivors." *Biometrics* 54:1508-15.

◦ *Recommended: Frailty Models*:

- Bennett, D. Scott. 1997. "Testing Alternative Models of Alliance Duration, 1816-1984." *American Journal of Political Science* 41(July):846-78.
- Box-Steffensmeier, Janet M., and Suzanna De Boef. 2005. "Repeated Events Survival Models: The Conditional Frailty Model." *Statistics in Medicine* 25(December):3518-33. DOI: 10.1002/sim.2434.
- Box-Steffensmeier, Janet M., Suzanna L. De Boef and Kyle A. Joyce. 2007. "Event Dependence and Heterogeneity in Duration Models: The Conditional Frailty Model." *Political Analysis* 15(3):237-256.
- Carpenter, Daniel. 2002. "Groups, the Media, Agency Waiting Costs and FDA Drug Approval." *American Journal of Political Science* 46(July):490-505.
- Chiozza, Giacomo, and Hein E. Goemans. 2004. "International Conflict and the Tenure of Leaders: Is War Still Ex Post Inefficient?" *American Journal of Political Science* 48(July):604-18.

        - Manton, Kenneth G., Eric Stallard and James W. Vaupel. 1981. "Methods for Comparing the Mortality Experience of Heterogeneous Populations." *Demography* 18(August):389-410.
        - Omori, Yasuhiro and Richard A. Johnson. 1993. "The Influence of Random Effects on the Unconditional Hazard Rate and Survival Functions." *Biometrika* 80(4):910-14.
        - Sastry, Naryan. 1997. "A Nested Frailty Model for Survival Data, With an Application to the Study of Child Survival in Northeast Brazil." *Journal of the American Statistical Association* 92(438):426-35.
        - Vaupel, James W., Kenneth G. Manton, and Eric Stallard. 1979. "The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality." *Demography* 16:439-54.

- **Exercise Five: Fit and interpret parametric and Cox survival models.**

October 14: Panel/TSCS: Overview and Unit Effects

- **Readings**
    - *Required*:
        - Hsaio, Cheng. 2003. *Analysis of Panel Data*. Chapters 1 and 3.
        - Stimson, James. 1985. "Regression in Space and Time: A Statistical Essay." *American Journal of Political Science* 29:914-47.
        - Zorn, Christopher. 2001. "Estimating Between- and Within-Cluster Covariate Effects, with an Application to Models of International Disputes." *International Interactions* 27(4):433-45.
    - *Recommended*:
        - Bartels, Larry M. 1996. "Pooling Disparate Observations." *American Journal of Political Science* 40(August):905-42.
        - Blackwell, Matthew, and Adam N. Glynn. 2018. "How to Make Causal Inferences with Time-Series Cross-Sectional Data under Selection on Observables." *American Political Science Review* 112: forthcoming.
        - Neuhaus, J. M., and J. D. Kalbfleisch. 1998. "Between- and Within-Cluster Covariate Effects in the Analysis of Clustered Data." *Biometrics* 54:638-45.
        - Nuamah, Nicholas N. N. N. 1986. "Pooling Cross Section and Time Series Data." *The Statistician* 35:345-51.
        - Plumper, Thomas, and Vera E. Troeger. 2007. "Efficient Estimation of Time-Invariant and Rarely Changing Variables in Finite Sample Panel Analyses with Unit Fixed Effects." *Political Analysis* 15(2):124-139.
        - "Symposium on Fixed-Effects Vector Decomposition." 2011. *Political Analysis* 19(2).

October 16: Panel/TSCS: Dynamics

- **Readings**
    - *Required*:
        - 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(September): 634-647.
        - Beck, Nathaniel, and Jonathan N. Katz. 1996. "Nuisance vs. Substance: Specifying and Estimating Time-Series Cross-Section Models." *Political Analysis* 6:1-36.

- Keele, Luke, and Nathan J. Kelly. 2006. "Dynamic Models for Dynamic Theories: The Ins and Outs of Lagged Dependent Variables." *Political Analysis* 14(2):186-205.
- Wawro, Gregory. 2002. "Estimating Dynamic Panel Data Models in Political Science." *Political Analysis* 10(Winter):25-48.

◦ *Recommended*:

- Achen, Christopher. 2000. "Why Lagged Dependent Variables Can Suppress the Explanatory Power of Other Independent Variables." Presented at the Annual Meeting of the Society for Political Methodology, UCLA. Available here.
- Anderson, T.W., and C. Hsiao. 1982. "Formulation and Estimation of Dynamic Models Using Panel Data." *Journal of Econometrics* 18:47-82.
- Beck, Nathaniel. 1991. "Comparing Dynamic Specifications: The Case of Presidential Approval." *Political Analysis* 3:51-87.
- Beck, Nathaniel. 2001. "Time-Series Cross-Section Data: What Have We Learned in the Past Few Years?" *Annual Review of Political Science* 4:271-293.
- Beck, Nathaniel, and Jonathan Katz. 2011. "Modeling Dynamics in Time-Series-Cross-Section Political Economy Data." *Annual Review of Political Science* 14:331-52.
- Blais, Andre, Donald Blake, and Stephane Dion. 1996. "Do Parties Make a Difference: A Reappraisal." *American Journal of Political Science* 40:514-520.
- Burkhart, Ross E., and Michael S. Lewis-Beck. 1994. "Comparative Democracy: The Economic Development Thesis." *American Political Science Review* 88:903-910.
- Smith, Mark A. 2001. "The Contingent Effects of Ballot Initiatives and Candidate Races on Turnout." *American Journal of Political Science* 45(3): 700-706.
- Wawro, Gregory, and Ida Pagter Kristensen. 2006. "Lagging the Dog?: The Robustness of Panel Corrected Standard Errors in the Presence of Serial Correlation and Observation Specific Effects." Working paper: Columbia University. Contact Dr. Wawro (`gjw10@columbia.edu`) if you're interested in this paper.
- Wilson, Sven E., and Daniel M. Butler. 2007. "A Lot More to Do: The Sensitivity of Time-Series Cross-Section Analyses to Simple Alternative Specifications." *Political Analysis* 15(2):101-123.

- **Exercise Six: Fit and discuss some panel / TSCS data models.**

October 21 - 23: Panel Data Models for Binary, Count, and Other Odd Responses, Including GEEs

- **Readings**

◦ *Required*:

- Beck, Nathaniel, Jonathan N. Katz, and Richard Tucker. 1998. "Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable." *American Journal of Political Science* 42(October):1260-88.
- Cameron, A. Colin, and Pravin K. Trivedi. 1998. *Regression Analysis of Count Data*. New York: Cambridge University Press. Chapter 9.
- Hsaio, Cheng. 2003. *Analysis of Panel Data*. Chapter 7, §7.1-7.3 and Chapter 8.
- Neuhaus, J. M., J. D. Kalbfleisch, and W. W. Hauck. 1991. "A Comparison of Cluster-Specific and Population-Averaged Approaches for Analyzing Correlated Binary Data." *International Statistical Review* 59(1):25-35.

- Zorn, Christopher. 2001. "Generalized Estimating Equation Models for Correlated Data: A Review with Applications." *American Journal of Political Science* 45(April):470-90.

- *Recommended*:
  - Baker, Andy, and Kenneth F. Greene. 2011. "The Latin American Left's Mandate: Free-Market Policies and Issue Voting in New Democracies." *World Politics* 63(1):43-77.
  - Ballinger, Gary A. 2004. "Using Generalized Estimating Equations for Longitudinal Data Analysis." *Organizational Research Methods* 7:12750.
  - Caldeira, Gregory A., John R. Wright, and Christopher Zorn. 1999. "Strategic Voting and Gatekeeping in the Supreme Court." *Journal of Law, Economics and Organization* 15(3):549-72.
  - Green, Donald P., Soo Yeon Kim, and David Yoon. 2001. "Dirty Pool." *International Organization* 55:441-68 (and commentary by Beck & Katz, Oneal & Russett, and King).
  - Katz, Ethan. 2001. "Bias in Conditional and Unconditional Fixed Effects Logit Estimation." *Political Analysis* 9(Autumn):379-84 (and also see Coup'e, Tom (2005) "Bias in Conditional and Unconditional Fixed Effects Logit Estimation: A Correction." *Political Analysis* 13(Summer):292-95).
  - Korre, A.K., and V.G.S. Vasdekis. 2018. "Goodness of Fit Tests for Random Effect Models with Binary Responses." *Statistics in Medicine*. Forthcoming.
  - Li, Quan, and Drew Schaub. 2004. "Economic Globalization and Transnational Terrorism: A Pooled Time-Series Analysis." *Journal of Conflict Resolution* 48:230-258.
  - Martin, Andrew D. 2003. "Bayesian Inference for Heterogeneous Event Counts." *Sociological Methods and Research* 32:30-63.
  - Wawro, Gregory. 2001. "A Panel Probit Analysis of Campaign Contributions and Roll Call Votes." *American Journal of Political Science* 45(July):563-579.
  - Whitford, Andrew B., Jeff Yates, and Holona L. Ochs. 2006. " Ideological Extremism and Public Participation." *Social Science Quarterly* 87(1):36-54.
  - Wooldridge, Jeffrey. 1999. "Distribution-Free Estimation of Some Nonlinear Panel Data Models." *Journal of Econometrics* 90(May):77-97.

Part Three: Measurement

October 28-30: Principal Components and Factor Analysis

- **Readings**
  - *Required*:
    - Everitt and Hothorn 2011. Chapters 3 and 5. (Also scan chapter 7.)
    - Greenacre, Michael. 2012. "Biplots: The Joy of Singular Value Decomposition." *Wiley Interdisciplinary Reviews: Computational Statistics* 4:399-406.
    - Henson, Robin K., and J. Kyle Roberts. 2006. "Use of Exploratory Factor Analysis in Published Research: Common Errors and Some Comment on Improved Practice." *Educational and Psychological Measurement* 66:393-416.
  - *Recommended*:
    - Flora, David B., and Jessica K. Flake. 2017. "The Purpose and Practice of Exploratory and Confirmatory Factor Analysis in Psychological Research: Decisions for Scale Development and Validation." *Canadian Journal of Behavioural Science* 49:78-88.

        - Gabriel, K.R. 1971. "The Biplot Graphic Display of Matrices with Application to Principal Components Analysis." *Biometrics* 58:453-467.
        - Greenacre, Michael J., and Patrick J. F. Groenen. 2016. "Weighted Euclidean Biplots." *Journal of Classification* 33:442-459.
        - MacCallum, R.C. 1974. "Relations Between Factor Analysis and Multidimensional Scaling." *Psychological Bulletin* 81: 505-516.

November 4: Scaling

- **Readings**
    - Required:
        - Everitt and Hothorn. 2011. Chapter 4.
        - Mair, Patrick, and Jan De Leeuw. 2015. "Unidimensional Scaling." In *Wiley StatsRef: Statistics Reference Online*. New York: Wiley. pp. 1-3.
        - Sijtsma, Klaas. 2009. "On the Use, Misuse, and the Very Limited Usefulness of Cronbach's Alpha." *Psychometrika* 74:107-120.
    - Recommended:
        - Borg, Ingwer, and Patrick Groenen. 2005. *Modern Multidimensional Scaling: Theory and Applications*, Second Edition. New York: Springer.
        - Coombs, Clyde H. 1950. "Psychological Scaling Without a Unit of Measurement." *Psychological Review* 57: 145-158.
        - De Leeuw, J., and P. Mair. 2009. "Multidimensional Scaling Using Majorization: SMACOF in R." *Journal of Statistical Software* 31:1-30.
        - McIver, John, and Edward G. Carmines. 1981. *Unidimensional Scaling*. New York: Sage.
        - Poole, Keith T. 1984. "Least Squares Metric, Unidimensional Unfolding." *Psychometrika* 49: 311-323.
        - Spector, Paul E. 1992. *Summated Rating Scale Construction*. New York: Sage.
        - Young, Forrest W. 1984. "Scaling." *Annual Review of Psychology* 35: 55-81.

November 6: Cluster Analysis

- **Readings**
    - *Required*:
        - Everitt and Hothorn. 2011. Chapter 6.
        - Ahlquist, John, and Christian Breunig. 2012. "Model-based Clustering and Typologies in the Social Sciences." *Political Analysis* 20:92-112.
    - *Recommended*:
        - Jakulin, Alecs, W. Buntine, T. Pira, and H. Brasher. 2009. "Analyzing the U.S. Senate in 2003: Similarities, Clusters, and Blocs." *Political Analysis* 17:291-310.
        - Ristei Gugiu, M., and M. Centellas. 2013. "The Democracy Cluster Classification Index." *Political Analysis* 21:334-349.
        - Wierzchon, Slawomir, and Mieczyslaw Klopotek. 2018. *Modern Algorithms of Cluster Analysis*. New York: Springer.
- **Exercise Seven: Do a little bit of multivariate statistics.**

November 12-14: Item Response Models

- **Readings**
  - *Required*:
    - Hambleton et al. (1991), pp. 7-46, 53-88, 109-122.
  - *Recommended*:
    - Lord, Frederic M. 1983. "Unbiased Estimates of Ability Parameters, of Their Variance, and of Their Parallel Forms Reliability." *Psychometrika* 48:477-82.
    - Martin, Andrew D., Kevin M. Quinn, and Jong Hee Park. "`MCMCpack`: Markov Chain Monte Carlo Package."
    - Poole, Keith, and Howard Rosenthal. 1985. "A Spatial Model of Legislative Roll Call Analysis." *American Journal of Political Science* 29(2):357-384
    - Poole, Keith. 2005. *Spatial Models of Parliamentary Voting*. New York: Cambridge University Press.
    - Rasch, Georg. 1961. "On General Laws and the Meaning of Measurement in Psychology." *Proceedings of the IV Berkeley Symposium on Mathematical Statistics and Probability* 4:321-333.
    - Rizopoulos, Dimitris. 2006. "`ltm`: An R Package for Latent Variable Modeling and Item Response Theory Analyses." *Journal of Statistical Software* 17(5).
- **Exercise Eight: Fit and discuss some item response models.**

November 18-20: Network Analysis

- **Readings**
  - *Required*:
    - TBA
  - *Recommended*:
    - TBA
- **Exercise Nine: Fit and discuss some network models.**

November 25-27: No Class: Thanksgiving Break

December 2-4: Text Analysis: Introduction and Overview

- **Readings**
  - *Required*:
    - Denny, Matthew J., and Arthur Spirling. 2018. "Text Preprocessing For Unsupervised Learning: Why It Matters, When It Misleads, And What To Do About It." *Political Analysis* 26:168-189.
    - Grimmer, Justin, and Brandon M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis* 21:267-297.

- Miner, Gary, John Elder IV, Thomas Hill, Robert Nisbet, Dursun Delen, and Andrew Fast. 2012. *Practical Text Mining and Statistical Analysis for Non-structured Text Data Applications*, 1st Ed. Cambridge, MA: Academic Press. Chapters 2 and 3. Available at https://nlp.stanford.edu/IR-book/.
- O'Connor, Brendan, David Bamman, and Noah A. Smith. 2011. "Computational Text Analysis for Social Science: Model Assumptions and Complexity." NIPS Conference.

- *Recommended* :
  - Monroe, Burt and Phillip Schrodt. 2008. "Introduction to the Special Issue: The Statistical Analysis of Political Text." *Political Analysis* 16:351-355.
  - Pathak, Manas A. 2014. *Beginning Data Science with R*. New York: Springer. Chapter 8.
  - Roberts, Margaret E. 2017. "Introduction to the Virtual Issue: Recent Innovations in Text Analysis for Social Science." *Political Analysis* Virtual Issue.

December 9: Text Analysis: Sentiment, Topics, and Scaling (A Sampling)

- **Readings**
  - *Required (choose 3-4)*:
    - Benoit, Kenneth, Kevin Munger, and Arthur Spirling. 2019. "Measuring and Explaining Political Sophistication Through Textual Complexity." *American Journal of Political Science* 63:491-508.
    - Blei, David. 2012. "Probabilistic Topic Models." *Communications of the ACM* 55:77-84.
    - Laver, Michael, Kenneth Benoit, and John Garry. 2003. "Extracting Policy Positions from Political Texts Using Words as Data." *American Political Science Review* 97:311-331.
    - Lowe, Will. 2008. "Understanding Wordscores." *Political Analysis* 16:356-371.
    - Pang, Bo, and Lillian Lee. 2008. "Opinion Mining and Sentiment Analysis." *Foundations and Trends in Information Retrieval* 2:1-135. (read quickly)
    - Slapin, Jonathan and Sven-Oliver Prokschk. 2008. "A Scaling Model for Estimating Time-Series Party Positions from Texts." *American Journal of Political Science* 52:705-722.
  - *Recommended*:
    - Blei, David, Andrew Ng, and Michael Jordan. 2003. "Latent Dirichlet Allocation." *Journal of Machine Learning* 3:993-1022.
    - Dodds, Peter and Christopher Danforth. 2009. "Measuring the Happiness of Large- Scale Written Expression: Songs, Blogs, and Presidents." *Journal of Happiness Studies* 11:441-456.
    - Grimmer, Justin. 2010. "A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases." *Political Analysis* 18:1-35.
    - Lauderdale, Benjamin, and Alexander Herzog, "Measuring Political Positions from Legislative Speech." *Political Analysis* 24:374-394.
    - Lowe, Will. 2016 (etc.) "Scaling Things We Can Count." Manuscript: Princeton University. http://dl.conjugateprior.org/preprints/all-on-the-line.pdf
    - Quinn, Kevin M., Burt L. Monroe, Michael Colaresi, Michael H. Crespin, and Dragomir R. Radev. 2010. "How to Analyze Political Attention with Minimal Assumptions and Costs." *American Journal of Political Science* 54:209-228.
    - Rice, Douglas R., and Christopher Zorn. 2018. "Corpus-Based Dictionaries for Sentiment Analysis of Specialized Vocabularies." *Political Science Research and Methods* 6:forthcoming.

- Roberts, Margaret E., Brandon M. Stewart and Dustin Tingley. 2018 "stm: R Package for Structural Topic Models." *Journal of Statistical Software*, forthcoming.
- Soroka, Stuart, Lori Young, and Meital Balmas. 2015. "Bad News or Mad News? Sentiment Scoring of Negativity, Fear, and Anger in News Content." *The Annals of the American Academy of Political and Social Science* 659:108-121.

- **Exercise Ten: Analyze some text.**

December 11: Text Analysis: Wrap-Up, Catch-Up, and Review

December 18: Final papers/projects are due.