Oregon State University
School of Public Policy
Spring 2014

PPOL 622: Advanced Policy Analysis (4 Credits)
Lecture: Monday TBD
Lab: Wednesday TBD

Dr. Alison Johnston
Gilkey Hall 300b
E-mail: Alison.Johnston@oregonstate.edu
Office Hours: TBD

This course is introduces students to advanced quantitative modelling techniques used in policy analysis. ***Students are not expected to have a thorough understanding of matrix algebra or calculus, although the required readings may present topics in these formats.*** The course focuses predominantly on temporal dynamics within quantitative research, and graduate students are expected to apply these topics to a research topic. In lecture, we will discuss the assumptions and conditions of temporal multivariate statistical modelling, how they have been applied to academic papers and research agendas, and how policy analysis has adapted to innovations in modelling techniques over the past 3 decades. In lab we will work with various datasets, exploring the methods that we learn in lecture as well as discuss required article readings in greater depth. This course will focus predominantly on: Poisson regression (counts modelling), tobit and negative binomial regression (zero-clustered dependent variables), time series, and panel regressions for continuous and limited-response dependent variable.

1. Student Learning Outcomes:

By the end of the course students will:

1. Evaluate and master methods of temporal multivariate quantitative analysis.
2. Evaluate and critique the progression of different modelling techniques over time within the field of policy analysis and political science
3. Develop an innovative research paper which applies one of the methods learned in the course to a data-set of the student's choosing.
4. Judge empirical papers that utilize different regression techniques
5. Conduct statistical modelling within the regression software program STATA, identify proper model specification for a dependent variable, and interpret how alternative model design influences empirical results.

2. Enforced Prerequisites

Completion of one of the following sequences: SOC 516 and Econ 524 with a B+ or higher; PPOL 621 (with a C+ or higher), or; AEC 523 and 525 (with a C+ or higher)

3. Required Texts

Wooldridge, J. (2006). *Introductory Econometrics: A Modern Approach*. Thomson, Couth-Western. Mason, Ohio

Wooldridge, J. (2002) *Econometric Analysis of Cross Section and Panel Data.* MIT Press. Cambridge, Massachusetts

Long, J.S. and Freese, J. (2006) *Regression Models for Categorical Dependent Variables Using STATA.* STATA Press. 2nd Edition.

For those of you with a more advanced background in matrix algebra and/or calculus, I highly recommend the following text as a supplement for this course: Greene, W. (2003) *Econometric Analysis.* Prentice Hall, New Jersey.

In addition to the core text, there will be supplementary articles assigned for each week where relevant methods are used. *You are expected to read at least one supplementary article per week.*

4. Assessment of Outcomes

- Paper replication homeworks (15% each) (60%)
- Article presentation (10%)
- Final Paper (30%)
    - Selection of dataset (Due ***by week 6 in lab***)
    - Presentation of testable hypothesis (5% - Due by ***week 6 in lab***)
    - Final report (25% - Due ***on Wednesday of finals week***)

Lab homeworks: Homeworks revolve around the empirical replication of papers published in top peer reviewed political science and policy journals. Questions relate to the material discussed in the relevant weeks, and require the empirical replication, analysis and critique of methods used in the paper.

Article Presentation: You are required to present one of the supplementary articles in a 15 minute presentation. Your presentation must outline the research question, empirical method and model specifications, and the results. You must also outline how the methods used in the article relate to the current week's (as well as previous weeks) – i.e. whether the authors decided for or against particular model specifications in line with what has been discussed in the course and course readings.

Final paper (due on Wednesday in finals week): Your final paper (maximum of 5,000 words INCLUDING output tables and references) will apply one of the methods to an original research question, and present it in the form of a journal article. In your article, you must include: a research question and testable hypothesis (or hypotheses), a brief review (6-10 articles) of the literature relevant to your hypothesis, a description of the data and the regression model you developed to test your hypothesis including the results, and concluding remarks that discuss the implications of your research. This paper must rely upon the use of one econometric method, along with relevant tests associated with the method's assumptions, discussed in class. The dataset you will use for your paper as well as your hypotheses you aim to test are ***due in Week 6's lab***. You will be able to test the methods used in labs on your dataset after completion of lab assignments.

Plagiarism: Academic work must be your own. It is plagiarism to claim work (such as writing, exams or projects) done by anyone other than the author(s) named. Plagiarism also includes cutting and pasting information from websites without attribution of ***AND*** paraphrasing someone else's ideas or writing. It is not sufficient to re-arrange or re-state someone else's writing or ideas. A zero tolerance policy will be applied towards plagiarism and any work which is plagiarized will automatically result in a ***COURSE GRADE OF F***. For more information on how the university handles academic misconduct, go to http://oregonstate.edu/studentconduct/offenses.

Disrespectful behaviour: Disrespectful behaviour towards students on grounds of race, gender, economic background, age, sexual orientation, religion, or any other factors which individuals have no choice or are irrelevant to the class will not be tolerated. ***Disrespectful behaviour can result in course expulsion***. For more information on the university's policy regarding academic conduct go to http://oregonstate.edu/studentconduct/offenses

Students with Disabilities: Accommodations are collaborative efforts between students, faculty and Disability Access Services (DAS). Students with accommodations approved through DAS are responsible for contacting the faculty member in charge of the course prior to or during the first week of the term to discuss accommodations. Students who believe they are eligible for accommodations but who have not yet obtained approval through DAS should contact DAS immediately at 737-4098."

Grading Scale

- A: 94-100%
- A-: 90-93%
- B+: 87-89%
- B: 83-86%
- B-: 80-82%
- C+: 77-79%
- C: 73-76%
- C-: 70-72%
- D+: 67-69%
- D: 63-66%
- D-: 60-62%
- F: <60%

Course Schedule:

Week 1: Review of Multivariate Linear Regression

Review of (cross-sectional) ordinary least squares and logistic regression as well as a brief introduction to matrix algebra and derivative notation.

Core Reading:

- Wooldridge (2002) Chapter 4 (the Single-Equation Linear Model and OLS Estimation) and 15 (Discrete Response Models – read sections 15.1-15.7)

Supplementary Reading:

- Gerner Hariri, J. (2012) "The Autocratic Legacy of Early Statehood", *American Political Science Review*. 106(3): 471-494.
- Hamann, K., Johnston AL, and Kelly, J. (2013) "Striking Concessions from Governments: Explaining the Success of General Strikes in Western Europe, 1980-2009." *Comparative Politics,* 46(1): 23-41

Lab Exercises (using data from Hamann, Johnston and Kelly, 2013):

- First week's lab will consist of a review of logistic regression analysis, as well as a practice replication paper

Week 2: Poisson Regression Models (Counts Data), Negative Binomial Regression, Zero Inflated Counts, and Tobit Regression models

Review of counts modelling (Poisson and negative binomial regression). Discussion of over-dispersed Poisson and negative binomial models, zero-clustering, and the zero-inflated counts and tobit models.

Core Reading:

- Coxe, S., West, S., and Aiken, L. (2009) "The Analysis of Count Data: A Gentle Introduction to Poisson Regression and its Alternatives". *Journal of Personality Assessment*. 91(2): 121-136.
- Wooldridge, J. (2006). Chapter 17.2 (The Tobit Model)
- Long, J.S. and Freese, J. (2006): Chapter 8-8.3 and 8.6 (Models for Counts Outcomes)

Supplementary Reading:

- Drury, A., and DeLisi, M. (2008). "Gangkill: An Exploratory Empirical Assessment of Gang Membership, Homicide Offending, and Prison Misconduct", *Crime and Delinquency*. 57(1): 130-146.
- Kenkel, D. (1991) "Health Behavior, Health Knowledge, and Schooling", *Journal of Political Economy*. 99(2): 287-305.

- Neumayer, E. and Plümper, T. (2011) "Foreign terror on Americans", *Journal of Peace Research*. 48(1): 3-17.

Lab Exercises (using data from Neumayer and Plümper, 2011):

- Application of Poisson regression analysis and the comparison of its output with standard OLS
- The calculation of counts probabilities
- Determination of over-dispersion within Poisson models, and the application of over-dispersed Poisson regression and negative binomial regression
- The use of zero-inflated Poisson (ZIP), zero-inflated negative binomial (ZINB) and tobit modelling for zero-skewed counts data

Week 3: Introduction to Time Series (the static model)

Introduction to static time series, event history analysis, and its basic assumptions. Discussion of time stationarity, auto-correlation for static models, and how their presence influence results. Discussions on how to detect (non-)stationarity and (first order) auto-correlation and how to correct for it in linear models via generalized least squares.

Core Reading:

- Wooldridge (2006): Chapters 10 (Basic regression Analysis with Time Series), Chapter 11.1 (Stationarity and Weakly Dependent Time Series), and Chapter 12.1-12.4 (read up to Serial Correlation-Robust Inference after OLS)

Supplementary Reading:

- Birz, G., and Lott, J.R. (2011) 'The effect of macroeconomic news on stock returns: New evidence from newspaper coverage". *Journal of Banking and Finance.* 35: 2791-2800.

Lab Exercises ***(HW 1 due at the beginning of lab)***:

- Setting up and estimating time-series data for analysis
- Determination of time stationarity via the Dickey-Fuller unit root tests, and how to correct for it (via differencing and the inclusion of trends)
- Applications of seasonality to time series data.
- Tests for autocorrelation (Durbin Watson test)
- Applications of the Cochrane-Orcutt and Prais-Winsten estimation technique for generalized least squares

Week 4: Time Series (the dynamic model)

Introduction to the dynamic time series model (i.e. where independent variables include lags of the regressors, or; auto-regressive models, where a lag of the dependent variable is an independent variable). Discussion of how serial correlation operates in dynamic models and how to correct for it. Discussion of higher order serial correlation and dynamic heteroskedasticity in time series models and how to correct for it.

Core Reading:

- Wooldridge (2006): Chapters 11 (Stationary and Weakly Dependent Time Series) and 12.5 onwards (Serial Correlation-Robust Inference after OLS)

Supplementary Reading:

- Schneider, G. and Troeger, V. (2006) "War and the World Economy: Stock Market Reactions to International Conflicts". *Journal of Conflict Resolution*. 50: 623-645.
- Goldstein, J., and Pevehouse, J. (1997) "Reciprocity, Bullying, and International Cooperation: Time-series Analysis of the Bosnia Conflict", *American Political Science Review*. 91(3): 515-529
- Bechtel, M. (2009) "The Political Sources of Systemic Investment Risk: Lessons from a Consensus Democracy", *Journal of Politics*. 71(2): 661-677.

Lab Exercises (using data from Bechtel, 2009):

- Testing for higher order serial correlation in static and dynamic models using the Breusch–Godfrey test
- Employing Newey West standard errors to correct for higher order serial correlation and heteroskedasticity in static models
- Estimating dynamic/autoregressive models in STATA
- The use of ARCH models to account for dynamic heteroskedasticity in time series models

Week 5: Introduction to Panel Regression

Introduction to panel regression analysis and its assumptions. How to create and analyze panels. Discussion of difference-in-difference panel estimators via interactive models and first-difference estimators. Introduction to large t panel analysis and brief overview how the conventional tools of panel analysis have changed in political science over the past 30 years.

Core Reading:

- Wooldridge (2006): Chapter 13 (Pooling Cross Sections across time: Simple Panel Data Methods).
- Stimson, J. (1985) "Regression in Space and Time: A Statistical Essay", *American Journal of Political Science*. 29(4): 914-947.

Supplementary Reading:

- Bechtel, M. and Hainmueller, J. (2011) "How lasting is voter gratitude? An Analysis of the Short- and Long-term Electoral Returns to Beneficial Policy". *American Journal of Political Science*. 55(4): Pg. 852-868.

- Huber, E., Ragin, C., and Stephens, J. (1993) “Social Democracy, Christian Democracy, Constitutional Structure, and the Welfare State”. *American Journal of Sociology*. 99(3): 711-749.
- Rodrik, D. (1998) “Why do more open economies have bigger governments?”, *Journal of Political Economy*. 106(5): 997-1032.

Lab Exercises (using data from Bechtel and Hainmueller, 2011 – ***HW 2 due at beginning of lab***):

- Setting up panel analysis within STATA
- Conducting difference-in-difference analyses within STATA using both an interactive model and a first-difference approach.

Week 6: Panel Regression for OLS (the Fixed and Random Effects Model)

Discussion of the analysis of dependent and independent variables which are largely time invariant (i.e. institutional variables), and how to incorporate different model specifications (i.e. period averaging) in order to account for time invariance. How to determine time stationarity within panels. Discussion of fixed versus random effects modelling within panel regression, as well as the advantages and disadvantages of using fixed effects when including national-specific or individual specific characteristics as independent variables within a regression.

Core Reading:

- Wooldridge (2006): Chapter 14 (Advanced Panel Data Methods)
- Garrett, G. and Mitchell, D. (2001) “Globalization, government spending and taxation in the OECD”. *European Journal of Political Research*. 39: 145-177
- Kittel, B. and Winner, H. (2005) “How reliable is pooled analysis in political economy? The globalization-welfare state nexus revisited”. *European Journal of Political Research*. 44: 269-293

Supplementary Reading:

- Neumayer, E. and Plumper, T. (2007) “The gendered nature of natural disasters: The impact of catastrophic events on the gender gaps in life expectancy, 1981-2002”, *Annuls of the Association of American Geographers*. 97(3): 551-566.
- Iversen, T. and Rosenbluth, F. (2008) “Work and Power: The Connection Between Female Labor Force Participation and Female Political Representation”, *Annual Review of Political Science*. 11: 479-495.
- Xu, Z., Hannaway, J., and Taylor, C. (2011) “Making a difference? The effects of Teach for America in high school”. *Journal of Policy Analysis and Management.* 30(3): 447-469

Lab Exercises (using data from Neumayer and Plumper, 2007 and Johnston, Hancké, and Pant, 2015):

- Examining how the incorporation of fixed effects influence time-invariant panel-specific variables
- Means of testing for the use of fixed or random effects manually as well as via a Hausman test
- Creating time-series graphics variables for individual panels (to determine variance and a bird's-eye view of time stationarity)
- Creating manual panel and time fixed effects with the "xi" operator.
- Testing for non-stationarity within panels
- Jack-knife analysis with panels

Week 7: Panel Regression Modelling for OLS (Other Violated Assumptions)

Discussion of the violation of serial correlation (both spatial and time-related), heteroskedasticity and multicollinearity across and within panels and how they influence OLS results. How to test whether OLS/time-series assumptions within panels are violated and how to specify models when these assumptions are violated

Core Reading:

- Beck, N. and Katz, J. (1995). "What to do (and not to do) with Time-Series Cross-Section Data", *American Political Science Review*. 89(3): 634-647.
- Plümper, T., Troeger, V., and Manow, P. (2005). "Panel data analysis in comparative politics: Linking method to theory". *European Journal of Political Research.* 44: 327-354.
- Wilson, S. and Butler, D. (2007) "A lot more to do: The sensitivity of Time-Series Cross-Section Analysis to Simple Alternative Specifications". *Political Analysis.* 15: 101-123.
- Rogers, W (2003). "Regression standard errors in clustered samples", *Stata Technical Bulletin*. **3**: 88-94.

Supplementary Readings:

- Hicks, A. and Swank, D. (1992) "Politics, institutions, and welfare spending in industrialized democracies, 1960-82", *American Political Science Review*. 86(3): 658-674
- Pickering, J. and Peceny, M. (2006) "Forging Democracy at Gunpoint", *International Studies Quarterly*. 50: 539-559
- Saideman, S., Lanoue, D., Campenni, M., and Stanton, S. (2002) "Democratization, Political Institutions, and Ethnic Conflict: A Pooled time-series analysis, 1985-1998". *Comparative Political Studies*. 35(1): 103-129
- Johnston AL, Hancké, R, and Pant, S. (forthcoming, 2015) "Comparative Institutional Advantage in Europe's Sovereign Debt Crisis". *Comparative Political Studies*. 48(6)

Lab Exercises (using data from Neumayer and Plumper, 2007 and Johnston, Hancké, and Pant, 2015):

- Testing for serial correlation via the Wooldridge test and correcting for it using Prais-Winsten AR transformations.
- Testing for heteroskedasticity and the discussion of different methods to control for heteroskedasticity (robust, panel corrected, and panel-clustered standard errors)

Week 8: Panel Regression Modelling for Binary Response Variables

Panel regression modelling for dependent variables which are binary. Discussion of random and conditional fixed effects logistic panel modelling.

Core Reading:

- Beck, N., Katz, J., and Tucker, R. (1998). "Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable", *American Journal of Political Science*. 42(4): 1260-1288.
- Oneal, J. and Russet, B. (1997) "The Classical Liberals were right: Democracy, Interdependence and Conflict, 1950-1985", *International Studies Quarterly*. 41: 267-294.

Supplementary Readings:

- Hamann, K., Johnston AL, and Kelly, J. (2013) "Unions Against Governments: Explaining General Strikes in Western Europe, 1980-2006". *Comparative Political Studies.* 46(9): 1030-1057
- Fearon, J. and Laitin, D. (2003) "Ethnicity, Insurgency, and Civil War", *American Political Science Review*. 97(1): 75-90.
- Burke, M., Miguel, E., Satyanath, S., Dykema, J., and Lobell, D. (2009) "Warming increases risk of civil war in Africa". *Proceedings of the National Academy of Sciences*. 106(49): 20670-20674.

Lab Exercises (using data from Burke et al, 2009 – ***HW 3 due at the beginning of lab***):

- Use of random and (conditional) fixed effects estimators on panels within binary dependent variables.
- Discussion of panel exclusion in conditional fixed effects logistic panel modelling.
- Discussion of time effects in panel logistic regression

Week 9: Open lab (possible discussion of hazard/survival models)

Class cancelled Monday due to Memorial Day holiday. Lab on Wednesday will be used for addressing final questions regarding final paper projects as well as review of the material (if desired). Possible introduction and discussion of hazard/survival models in Wednesday's lab.

Week 10: Panel Regression Modelling for Binary Response Variables (temporal dependency, separation problems and other violations)

Discussion of separation problems and temporal dependency within logistic panels. Generalized estimating equation models for correlated data in binary dependent response panels.

Core Reading:

- Carter, D. and Signorino, C. (2010) "Back to the Future: Modelling time Dependence in Binary Data". *Political Analysis*. 18: 271-292
- Zorn, C. (2001) "Generalized Estimating Equation Models for Correlated Data: A Review with Applications", *American Journal of Political Science*. 45(2): 470-490

Supplementary Readings:

- Simmons, B (2000) "International Law and State Behaviour: Commitment and Compliance in International Monetary Affairs", *American Political Science Review*. 94(4): 819-835.
- Oneal, J. and Russett, B. (1999) "Assessing the Liberal Peace with Alternative Specifications: Trade still reduces conflict", *Journal of Peace Research*. 36(4): 423-442.

Lab Exercises (using data from Hamann, Johnston, and Kelly, 2013):

- Determining the effects of separation (when one dummy perfectly predicts the y outcome for a specific year or panel-year) in panel logit
- Discussions of how to control for path dependency via linear trends and time dummies.
- Applications of generalized estimating equation models for logistic analysis with correlated data

Week 11: Finals Week

HW 4 due on Monday, noon of finals week