NOTE: This syllabus is revised, effective 3/21/2020.

PL SC 503: "Multivariate Analysis for Political Research"

Spring 2020

Professor Christopher Zorn
Department of Political Science
Pennsylvania State University
E-mail: zorn@psu.edu
Tuesday & Thursday, 1:00-2:30 p.m.
Boucke Building, Room 301

Course Description

This is the second (full) course in quantitative methods in Penn State's political science Ph.D. program. The course introduces students to linear regression models for the analysis of quantitative data, and provides a basis of knowledge for more advanced statistical methods. It will also have a substantial programming/computation focus. The course assumes basic math literacy, including familiarity with probability theory, properties of estimators, rudimentary calculus, and linear algebra, as well as mastery of the basic statistics taught in PLSC 502. The bulk of the course will focus on regression models for continuous response variables, and will include discussions of the mathematical bases for such models, their estimation and interpretation, model assumptions and techniques for addressing violations of those assumptions, model diagnostics, and topics related to model specification and functional forms. We'll conclude with an introduction to the idea of maximum likelihood, and a review of generalized linear models (logit, probit, etc.).

Note that all course materials (including this syllabus, slides, notes, data, computer code, homework exercises, etc.) will be available on a dedicated Github repo, which can be found at: https://github.com/PrisonRodeo/PLSC503-2020-git. Throughout this syllabus, hot links are in Penn State Blue.

Texts

Required:

Weisberg, Sanford. 2013. *Applied Linear Regression*, 4th Ed. New York: Wiley. (*ALR*'s Wiley page.)

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

Additional readings as necessary, all of which will be available via JSTOR™or on the course github repo.

The Weisberg text will be the primary text for the course. It's expensive, but a good reference, and is available both used and in an e-book version. Faraway is for the latter part of the course, and (among other things) is also used in PLSC 504. There is a second edition in print; either version will work for this course.

Recommended:

Kennedy, Peter. 2003. *A Guide to Econometrics*, 5th Ed. Cambridge: MIT Press. "Cliff's notes" for linear regression. There will be some readings assigned from Kennedy below, but these will be made available on the github repo.

Fox, John. 2008. *Applied Regression Analysis and Generalized Linear Models*, Second Edition. Thousand Oaks, CA: Sage Publications. Nice to have if you can get it cheaply; previous versions of this course used this as its main text.

Fox, John, and Sanford Weisberg. 2011. *An R and S-Plus Companion to Applied Regression*, Second Edition. Thousand Oaks, CA: Sage Publications. A companion to the Fox text, for S-Plus™ and R users.

Nagler, Jonathan. 1996. "Coding Style and Good Computing Practices." *The Political Methodologist* 6(2):2-8. Contains words to live by.

Other Good Regression Texts:

Chatterjee, Samprit, and Ali S. Hadi. 2006. *Regression Analysis by Example*, 4th Ed. New York: Wiley.

Cohen, Jacob, Patricia Cohen, Stephen G. West, and Leona S. Aiken. 2002. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, 3rd Ed. Lawrence Erlbaum.

Gelman, Andrew, and Jennifer Hill. 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York: Cambridge University Press.

Montgomery, Douglas C., Elizabeth A. Peck, and G. Geoffrey Vining. 2006. *Introduction to Linear Regression Analysis*, 4th Ed. New York: Wiley.

Seber, George A.F., and Alan J. Lee. 2003. *Linear Regression Analysis*, 2nd Ed. New York: Wiley.

A Few Other Useful References:

Chang, Winston. 2013. *The R Graphics Cookbook*. Sebastopol, CA: O'Reilly Media.

Crawley, Michael J. 2012. *The R Book*, 2nd Ed. New York: Wiley.

Gardener, Mark. 2012. *The Essential R Reference*. New York: Wiley.

Teetor, Paul. 2011. *The R Cookbook*. Sebastopol, CA: O'Reilly Media.

Some "Econometrics" Texts (can generally be ignored):

Dougherty, Christopher. 2007. *Introduction to Econometrics*, 3rd Ed. New York: Oxford University Press.

Greene, William. 2008. *Econometric Analysis*, 6th Ed. New York: Prentice-Hall.

Gujarati, Damodar. 2003. *Basic Econometrics*, 4th. Ed. New York: McGraw-Hill.

Hill, R. Carter, William E. Griffiths, and Guay C. Lim. 2007. *Principles of Econometrics*, 3rd Ed. New York: Wiley.

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

Maddala, G. S. 2001. *Introduction to Econometrics*, 3rd Ed. New York: Wiley.

Stock, James S. and Mark W. Watson. 2011. *Introduction to Econometrics*, 3rd International Edition. New York: Pearson.

Wooldridge, Jeffrey. 2005. *Introductory Econometrics: A Modern Approach*, 3rd Ed. Mason, OH: South-Western College Publishing.

Most of these are generally similar to Fox (2008), though with more of an "econometric" flavor (more emphasis on proofs, less emphasis on visualization, etc.).

The Methods Preceptor

Tae Gyoon Kim is the methods preceptor for PLSC 503. He is a PLSC / SODA Ph.D. candidate who studies computational social science and social media. He will serve as a "first line of defense" in the course: He can assist you with course material, software and programming issues, and other

matters related to the course work. He can be reached via e-mail at txk290 [at] psu [dot] edu.

Grading

Grading will be based on a total of 1000 points, divided as follows:

- Homework exercises: Ten worth 50 points each.
- A final paper/project, worth 500 points.

Details for the homework assignments and the final project will be announced in class.

Some Other Useful Resources

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 (http://www.icpsr.umich.edu).

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 (at http://polmeth.wustl.edu).

The **Comprehensive R Archive Network** (CRAN) (http://cran.r-project.org/) is the place to go for downloads, packages, and documentation. Similarly, the **Stata**^{TM} homepage (http://www.stata.com) is a valuable resource for questions about Stata statistical 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 (here). For further information, please visit the Student Disability Resources website (here).

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 here. 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

Linear Regression: Basics

- **January 14**: *Introduction: Regression, A Conceptual Overview*
  Readings (for background):
  - Preface to the 4th Ed. of Weisberg.
  - Berk, Richard. 2010. "What You Can and Can't Properly Do with Regression." *Journal of Quantitative Criminology* 26(4):481-487.
  - Weisberg, Chapter 1 and Appendix A.1 and A.2.
- **January 16**: *Bivariate Regression: A (Re)Introduction*
  Readings:
  - Weisberg, Chapter 2, pp. 21-30 and Appendix A.3.
- **January 21**: *Bivariate Regression: Inference*
  Readings:
  - Weisberg, Chapter 2, pp. 30-38 and Appendix A.4.
- **January 23**: *Model Fit (plus "Stupid Regression Tricks")*
  Readings:
  - Weisberg, Chapter 2, pp. 30-38 and Appendix A.4.
  - Lewis-Beck, Michael S. and Andrew Skalaban. 1990. "When to Use R-Squared." *The Political Methodologist* 3(2):11-12.
  - King, Gary. 1990. "When Not to Use R-Squared." *The Political Methodologist* 3(2):9-11.
  - Luskin, Robert C. 1991. "R-Squared Encore." *The Political Methodologist* 4(1):21-23.

- **January 28**: *No Class (PSU Faculty Senate).*

- **January 30**: *Special Topic: Bootstrapping And Other Delights*
  Readings:
  - No readings assigned.

  Homework One due.

Multivariate Linear Regression

- **February 4**: *Multivariate Regression: Estimation*
  Readings:
  - Weisberg, Chapter 3, pp. 51-68 and Appendix A.8.
- **February 6**: *Multivariate Regression: Inference*
  Readings:
  - Weisberg, Chapter 6, pp. 133-150.

  Homework Two due.

- **February 11**: *Multivariate Regression: Dichotomous Covariates*
  Readings:
  - Weisberg, Chapter 5, pp. 98-123.
- **February 13**: *(Non-)Linearity and Data Transformations*
  Readings:
  - Weisberg, Chapter 4, pp. 67-93; Chapter 8, pp. 185-199.

  Homework Three due.

- **February 18**: *Variance Issues*
  Readings:
  - Weisberg, Chapter 7, pp. 156-179.

- **February 20**: *Collinearity, etc.*
  Readings:
  - Kennedy, Chapter 11, pp. 205-217.

  Homework Four due.

- **February 25**: *Multivariate Regression: Residuals, Outliers, and Diagnostics*
  Readings:
  - Weisberg, Chapter 9, pp. 204-226.
- **February 27**: *Specification Error: Endogeneity / Simultaneity*
  Readings:
  - Kennedy, pp. 107-109; 180-191.

- **March 3**: *Variable Selection*
  Readings:
  - Weisberg, Chapter 10, pp. 234-248.
- **March 5**: *Multiplicative Interactions, I*
  Readings:
  - Friedrich, Robert J. 1982. "In Defense of Multiplicative Terms in Multiple Regression Equations." *American Journal of Political Science* 26(November):797-833.
  - Brambor, Thomas, William R. Clark, and Matt Golder. 2006. "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis* 14:63-82.
  - Hainmueller, Jens, Jonathan Mummolo, and iqing Xu. 2019. "How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice." *Political Analysis* 27:163-192.

  Homework Five due.

- **March 10**: **No Class – Spring Break**
- **March 12**: **No Class – Spring Break**

- **March 17**: **No Class – PSU Faculty Senate**

- **March 19**: *NO CLASS: Global Pandemic.*

Likelihood, etc.

- **March 24**: *Multiplicative Interactions, II*
  Readings:
  - No readings assigned.
- **March 26**: *Maximum Likelihood - Introduction*
  Readings:
  - Fox, Appendix D6, pp. 92-95.
  - Weisberg, Appendix A.11.

- **March 31**: *MLE: Estimation / Optimization*
  Readings:
  - No readings (but see above for some suggestions).
- **April 2**: *MLE: Inference (+ "Robust" Revisited)*
  Readings:
  - Buse, A. 1982. "The Likelihood Ratio, Wald, and Lagrange Multiplier Tests: An Expository Note." *The American Statistician* 36(3):153-57.
  - Freedman, D. A. 2006. "On the So-Called 'Huber Sandwich Estimator' and 'Robust' Standard Errors." *The American Statistician* 60:299-302.
  - King, Gary, and Margaret E. Roberts. 2014. "How Robust Standard Errors Expose Methodological Problems They Do Not Fix, and What To Do About It." *Political Analysis* 22:1-21.

- **April 7**: *Binary Responses, I.*
  Readings:
  - Weisberg, Chapter 12, pp. 270-279.

- Faraway, pp. 25-38.

Homework Six due.

- **April 9**: *Binary Responses, II*.
  Readings:
  - No readings assigned

- **April 14**: *Nominal Responses*.
  Readings:
  - Faraway, pp. 97-103.
- **April 16**: *Ordinal Responses*.
  Readings:
  - Faraway, pp. 106-112.

  Homework Seven due.

- **April 21**: *Event Counts*.
  Readings:
  - Faraway, pp. 55-66.
- **April 23**: *Generalized Linear Models*
  Readings:
  - Weisberg, Chapter 12, pp. 279-285.
  - Gill, Jeff. 2000. *Generalized Linear Models: A Unified Approach*. Thousand Oaks, CA: Sage. pp. 51-61.

  Homework Eight due.

- **April 28**: **No Class – PSU Faculty Senate**

- **April 30**: *Wrap-Up, Catch-Up, and Review*.
  Readings:
  - None / TBD.

Homework Nine due.

- **May 6**: **Final Papers / Projects Due**.