

# PSCI 7095-001: Advanced Political Data Analysis (Data II)

University of Colorado Boulder

Spring 2018

<b>Time:</b>	Thursday, 4:35-7:05
<b>Location:</b>	KTCH 1B31
<b>Instructor:</b>	Dr. Andrew Q. Philips
<b>Office:</b>	KTCH 144
<b>Email:</b>	andrew.philips@colorado.edu
<b>Office hours:</b>	Tuesday, 12:00-1:30, or by appointment
<b>Teaching Assistant:</b>	Kimberlee Chang (kimberlee.chang@colorado.edu)
<b>TA Office Hours:</b>	TBA

**COURSE DESCRIPTION:** Quantitative analysis is an important component of nearly all political science research. This course is designed to introduce you to the fundamental tools used for data analysis. We will review and build on what you have learned in Data I. Most of this course focuses on understanding the approach of ordinary least squares (OLS). We will start by reviewing some topics you may already know, then move onto OLS estimation. Then, we will cover violations of OLS assumptions, as well as various approaches to addressing such violations. In addition, we will spend a lot of time discussing how to present and interpret regression results, as well as substantive and statistical significance.

While it is crucial for political scientists to master OLS, in reality most research involves using alternative modeling techniques, many of which were developed for use in situations where OLS is inappropriate. Therefore, in the second half in the course, we cover a variety of alternative extensions.

By the end of this course you should be able to:

- Understand what is going on “under the hood” of OLS, and interpret regression results
- Diagnose and address violations of the regression assumptions
- Have a variety of models to add to your “toolkit”
- Apply what you have learned to your research.

**PREREQUISITES:** This an graduate level course; students should have a background in introductory regression (i.e., Data I). We will be working in matrix algebra notation throughout much of the course, although prior experience with this is not necessary.

**SOFTWARE:** We will use R for most of this course. Although familiarity with R is not necessary, it is a plus. Those unfamiliar with this program may want to purchase or borrow the suggested textbooks that cover working with R, although there are copious amounts of information available for free online. Please download both R (<https://cran.r-project.org/>) and RStudio (<https://www.rstudio.com/>) before the first class session. We will probably also use some Stata, although it is not necessary to purchase it for this course. Although there will not be a substantial amount of writing, students are encouraged to write up any assignments using  $\text{\LaTeX}$ .

**GRADES:** Course grades will be based on the following. Participation and homework assignments make up 40% of the final grade. About halfway through the semester, a mid-term exam will be given

that is worth 30% of the final grade. At the end of the semester, there will be a final exam worth 30% of the final grade. There are no opportunities for extra credit.

Participation and Homework	40%
Midterm Exam	30%
Final Exam	30%

The following scale will be used to turn numerical grades into letter ones. Note that I will round up a letter should your grade fall on the number (but on or above 0.5) between two letters (e.g., 89.5 up to 90 rounds up to an A-).

#### Grade Scale

A	95-100
A-	90-94
B+	87-89
B	84-86
B-	80-83
C+	77-79
C	74-76
C-	70-73
D+	67-69
D	64-66
D-	60-63
F	0-59

**PARTICIPATION:** Participation is an integral component of graduate courses. Students are expected to come to every class *having already read the assigned readings for that day*, and should be prepared to discuss them. Graduate-level courses only are successful when all students participate actively in the discussion.

**HOMEWORK:** Throughout the semester, there will be various homework assignments. Some of these will be group assignments, others on your own. Most will involve some form of data analysis and interpretation/presentation of regression results. We will discuss more specifics on homework in class.

**MIDTERM EXAM:** About halfway through the semester there will be a mid-term exam. This will be closed book.

**FINAL EXAM:** At the end of the semester, there will be a comprehensive final exam. The final exam will be open book. You may consult textbooks, articles, and your notes for this exam.

**TEACHING ASSISTANT:** The teaching assistant this semester will be Kimberlee Chang.

**ATTENDANCE AND LATE POLICY:** Attendance is a key component of succeeding in graduate school. I provide slides for each class, but we will have a much more comprehensive discussion than what appears on the slide. Attendance is mandatory, with the exception of university-excused absences. If you need to miss a class, you should—if possible—let me know in advance so that we can make arrangements.

Assignments are due on the day listed in the syllabus. Late assignments will not be accepted.

**REQUIRED TEXTS:** The following text is required for the course. Any additional readings will be

made available to you on the first day of class or as needed. This text is advanced, but will be a helpful reference after the semester is over.

- Greene, William H. *Econometric analysis*. 2011. 7th edition. Pearson.

Note that it is expected to read the week's required readings before coming to class.

**RECOMMENDED TEXTS:** The following texts are not required, but may be helpful to some. In the schedule below there are additional texts in the "suggested readings".

- Gujarati, Damodar N. and Dawn C. Porter. *Essentials of econometrics*. 2009. 4th edition. McGraw-Hill Education. A good introductory econometrics text.
- Angrist, Joshua D. and Jorn-Steffen Pischke. *Mostly harmless econometrics: An empiricist's companion*. 2009. Princeton University Press. A quite accessible introduction to causal inference and various other topics.
- Wooldridge, Jeffrey M. *Econometric analysis of cross section and panel data*. 2010. 2nd edition. MIT Press. A Greene-level text that focuses on cross-sectional and cross-sectional time series data.
- Kennedy, Peter. *A guide to econometrics*. 2008. 6th edition. Wiley-Blackwell. A good introductory econometrics text.
- Monogan, James E. 2015. *Political analysis using R*. Springer.
- Philips, Andrew Q. 2016. "AT<sub>E</sub>X: A brief introduction". Available [here](#).

**TENTATIVE SCHEDULE:** Note that this schedule is subject to change. We will spend as long as we need to on a topic, and many of the topics in the second half of the course may not take a full course day to cover.

## **Week 1: Course Introduction, Regression Assumptions, Introduction to Matrix Algebra and R**

Required Readings:

- Greene, Appendix A

Suggested Readings:

- Burns, Patrick. 2011. *The R Inferno*. Available at: <http://www.burns-stat.com/documents/books/the-r-inferno/>.
- Philips, Andrew Q. 2017. "R: A brief introduction."

## **Week 2: Under the Hood: OLS**

Required Readings:

- Greene, Ch. 1 and 2

## **Week 3: Under the Hood: OLS (continued)**

Required Readings:

- Greene, Ch. 3

## **Week 4: OLS in Practice**

Required Readings:

- Greene, Ch. 4

## **Week 5: OLS in Practice (Continued)**

Required Readings:

- Greene Ch. 5 and 6
- Brambor, Thomas, William Roberts Clark, and Matt Golder. 2006. "Understanding interaction models: Improving empirical analyses." *Political Analysis* 14(1): 63-82.
- Berry, William D., Matt Golder, and Daniel Milton. 2012. "Improving tests of theories positing interaction." *Journal of Politics* 74(3): 653-671.

Suggested Readings:

- Franzese, Robert, and Cindy Kam. 2009. *Modeling and interpreting interactive hypotheses in regression analysis*. University of Michigan Press.

## **Week 6: Generalized Least Squares**

Required Readings:

- Greene Ch. 9

Suggested Readings:

- White, Halbert. 1980. "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity." *Econometrica: Journal of the Econometric Society*: 817-838.

## **Week 7: Binary Dependent Variables**

Required Readings:

- Berry, William D., Jacqueline HR 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(1):248-266.

## **Week 8: MID-TERM EXAM**

### **POST MID-TERM: Weeks 9-15**

The second half of the course will consist of a variety of topics that build on the first half of the course. Depending on time, we will cover everything below (and maybe some extra topics). There will not be extensive coverage of each topic; in fact most topics could comprise an entire semester-long course. Instead, it is to help familiarize you with the various methodological tools that are out there for you to use, depending on your research question.

Class will be held March 15, April 12, 19, 26, and May 3. Class will not be held March 22, March 29 (Spring BREAK), or April 5 (MPSA), though we will probably reschedule one of these class times.

## Time Series:

Required Readings:

- Greene Ch. 20 and 21 (skim both)
- 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.
- DeBoef, Suzanna, and Luke Keele. 2008. "Taking time seriously." *American Journal of Political Science* 52(1): 184-200.
- Philips, Andrew Q. 2017. "Have Your Cake and Eat It Too? Cointegration and Dynamic Inference from Autoregressive Distributed Lag Models." *American Journal of Political Science*. doi:10.1111/ajps.12318

## Spatial Statistics:

Required Readings:

- Beck, Nathaniel, Kristian Skrede Gleditsch, and Kyle Beardsley. "Space is more than geography: Using spatial econometrics in the study of political economy." *International Studies Quarterly* 50.1 (2006): 27-44.
- Franzese, Robert J., and Jude C. Hays. 2007. "Spatial-econometric models of cross-sectional interdependence in political science panel and time-series-cross-section data." *Political Analysis* 15(2): 140-164.

## Endogeneity:

Required Readings:

- Greene Ch. 8

## Causality and Causal Inference:

Required Readings:

- Angrist, Joshua D. and Jorn-Steffen Pischke. 2009. *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press. Chapters 1 and 2.
- Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. "The colonial origins of comparative development: An empirical investigation." *American Economic Review* 91(5):1369-1401.
- Meyersson, Erik. 2014. "Islamic rule and the empowerment of the poor and pious." *Econometrica* 82(1): 229-269.

## Pooled Time Series, Clustering, and HLM:

Required Readings:

- Greene Ch. 11
- Arceneaux, K. and Nickerson, D.W., 2009. "Modeling certainty with clustered data: A comparison of methods." *Political Analysis* 17(2):177-190.
- Beck, Nathaniel, and Jonathan N. Katz. 1995. "What to do (and not to do) with time-series cross-section data." *American Political Science Review*: 634-647.
- King, Gary and Margaret E. Roberts. 2015. "How robust standard errors expose methodological problems they do not fix, and what to do about it." *Political Analysis* 23: 159-179.

Suggested Readings:

- Beck and Katz. 2011. "Modeling Dynamics in Time-Series-Cross-Section Political Economy Data." *Annual Review of Political Science*: 331-352.

## Maximum Likelihood Estimation:

Required Readings:

- King, Gary. 1998. *Unifying political methodology: The likelihood theory of statistical inference*. University of Michigan Press: Chapters 1-3.

## Resampling, Quantities of Interest, and Presenting Results:

Required Readings:

- 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):347-361.
- Hanmer, Michael J., and Kerem Ozan Kalkan. 2013. "Behind the curve: Clarifying the best approach to calculating predicted probabilities and marginal effects from limited dependent variable models." *American Journal of Political Science* 57(1):263-277.

Suggested Readings:

- Tomz, Michael, Jason Wittenberg, and Gary King. 2001. "Clarify: Software for interpreting and presenting statistical results."
- Imai, Kosuke, Gary King, and Olivia Lau. 2008. "Zelig: Everyone's statistical software."

## Final Exam: TBA

### STATEMENT ABOUT STUDENTS WITH DISABILITIES

The Americans with Disabilities Act (ADA) is a federal anti-discrimination statute that provides comprehensive civil rights protection for persons with disabilities. Among other things, this legislation requires that all students with disabilities be guaranteed a learning environment that provides for reasonable accommodation of their disabilities. If you believe you have a disability requiring an accommodation, please contact Disability Services—either online at <http://www.colorado.edu/disabilityservices/>—or at the Center for Community, N200, 107 UCB.

To best accommodate students who may require alternative services, it is crucial that you contact me *early in the semester* if you need such accommodations.

### HONOR CODE, COPYRIGHT, AND PLAGIARISM STATEMENTS

*"On my honor, as a University of Colorado Boulder student, I have neither given nor received unauthorized assistance"*

The CU Honor Code is intended to uphold the intellectual reputation of the university by establishing trust among individuals regarding intellectual honesty. As the website states, "The Honor Code secures an environment where academic integrity can flourish and aims to install the principles of honesty, trust, fairness, respect, and responsibility as essential features of the University of Colorado Boulder campus". Violations of intellectual honesty include plagiarism, cheating, and the unauthorized use of materials, all of which erode trust among individuals. If you have any questions about this, please see me, the Honor Code website (<http://www.colorado.edu/honorcode/>), or the Honor Code Office (1B70 Regent Admin Building).

The handouts and lectures used in this course are copyrighted. By "handouts," I mean all materials generated for this class, which include but are not limited to syllabi, exams, in-class materials,

and review sheets. Because these are copyrighted, you do not have the right to copy them or distribute them to others outside class, unless I expressly grant permission. In addition, I do not grant permission to tape class lectures.

## **SYLLABUS CHANGES**

I reserve the right to make changes to the syllabus during the course of the semester as needed and will make the most updated copy available to you and announce said changes during class.

**Last updated:** January 9, 2018