Statistical Methods in Political Research II - PS 699
Department of Political Science - University of Michigan
Winter 2015

UPDATED: DECEMBER 27, 2014

Instructor:	Rocío Titiunik, <titiunik@umich.edu>
Instructor's Office:	6733 Haven Hall
Class Location:	3451 Mason Hall
Class Time:	Tue-Thu 2:30–4:00 pm
Instructor's Office Hours:	Thursday 10-12 and by appointment
Prerequisites:	PS599 or equivalent course
GSI:	Diogo Ferrari, <dferrari@umich.edu>
Discussion Section:	Friday 9:00–10:00 am, at 2353 Mason Hall
GSI's Office Hours:	TBA

Course Outline. This course is an introduction to statistical and econometric analysis for social scientists. We will cover linear regression analysis, and nonlinear models as time permits. The course will be both theoretical and practical. There will be problem sets that will be both theoretical and empirical; we will use the software R.

Software. The course will use the statistical software R (see http://cran.r-project.org). All students are expected to learn R and use it to solve the problem sets. We will teach how to use R in the discussion section, but students are expected to be familiar with the language, in particular with the contents of PS514. You can find links to the R website and several online manuals and other resources on my website http://www.umich.edu/~titiunik/R.

Books. The required book for the course is Greene's textbook:

- Greene, William H., . *Econometric Analysis*, 2008. Prentice Hall, 6th Edition.

In addition to the required book, some lectures will also rely on the following books:

- Wooldridge, Jeffrey M., 2010. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, 2nd Edition. (If you want to buy an advanced econometrics book in addition to Greene's, I strongly recommend this one.)

- Cameron, Colin A. and Pravkin K. Trivedi, 2005. *Microeconometrics: Methods and Applications*. Cambridge University Press.
- Angrist, Joshua D. and and Jorn-Steffen Pischke, 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, First Edition.
- Casella, George and Roger L. Berger, 2001. *Statistical Inference*. Duxbury Press, 2nd edition.
- Hogg, Robert V., Allen Craig, and Joseph W. McKean, 2004. *Introduction to Mathematical Statistics*. Prentice Hall, 6th edition.
- Simon, Carl P., and Lawrence Blume, 1994. *Mathematics for economists*. New York: Norton.

If you are finding the level of difficulty in Greene (2011) a little too high, you can refer to the following more basic textbooks:

- Stock, James H. and Mark W. Watson, 2006. *Introduction to Econometrics*. Addison-Wesley Series in Economics, 2nd Edition.
- Fox, John, 2008. *Applied Regression Analysis and Generalized Linear Models*. Sage Publications, Inc, 2nd edition. (If you want to buy a more basic book, I highly recommend this one.)

Requirements. There will be frequent problem sets, a midterm exam and a final paper. All students must complete the problem sets, the midterm exam and the final exam to pass the course. Students are also expected to do the weekly readings, and come to class prepared to participate in lecture and ask questions.

Problem Sets. Problem sets are a very important part of the course. All calculations requiring computing and graphing must be done with the statistical software `R`. All problem sets must be submitted electronically via the `Ctools` website. Late problem sets will be given no points. The main body of your problem set should consist of clearly typed answers to the questions. The `R` code used to solve the questions should be attached as *an appendix at the end of the file*. Graphs should be clearly numbered and labeled, and should be placed on the main body of the file. The raw output from the `R` console will *not* be accepted as a solution. You should think of your problem set as a mini-report that is to be clearly organized and written.

Final Paper. You will be asked to replicate and extend a published paper. You can choose the paper based on your substantive interests, with the restriction that the paper must be published in a prominent political science (or related) journal within the last ten years (the more modern the paper, the easier it will be to obtain the data). You will be asked to do the following:

- Obtain the original data

- Replicate all tables in the original paper
- Provide a critique of the paper's research design, estimation methods and substantive claims
- Propose extensions to the analysis (robustness checks) that directly adress your criticisms

Your replication paper should be organized like a typical paper – i.e, something like: introduction, relevant prior literature, theory and hypotheses, data sources, research design, estimation results, conclusion. Each student is required to meet with the instructor at least once during office hours to discuss the paper chosen.

Before the final version is turned in, students are required to submit progress reports accoding to the following schedule:

Table 1: Replication paper due dates

Assignment	Due date
One-page report: chosen paper and replication plan	Thu Feb 5
4/8-page report: complete replication tables and outline of proposed extensions	Thu March 12
8/14-page report: final paper including replication and extension parts	Wed, April 29

Grading Policy. The final grade will be determined by problem sets (30%), midterm exam (40%) and final paper (30%).

Incomplete policy. No incompletes will be given.

Overview of covered topics

1. Review of basic matrix algebra, statistics and asymptotic theory
   (a) Random variables
   (b) Distribution, density and mass functions
   (c) Expectation, Median and Variance
   (d) Conditional distributions
   (e) Convergence in probability and distribution
   (f) Weak law of large numbers and central limit theorem
   (g) Asymptotic distribution
2. Maximum Likelihood Estimation
   (a) General approach
3. Classical Linear Regression
   (a) Ordinary Least Squares (OLS) estimation and inference
   (b) Finite-sample properties of OLS
   (c) Large-sample properties of OLS
   (d) Basic statistical inference
   (e) Some model specification issues
4. Departures from classical assumptions in linear regression model
   (a) General model with non-spherical disturbances
   (b) Heteroskedasticity
   (c) Serial correlation
   (d) Endogeneity (instrumental variables)
5. Non-linear regression functions
   (a) Non-linear least squares
   (b) Discrete choice models
6. Panel Data
7. Causal Inference

Course outline

Note 1: This outline may change, please check periodically for updates.
Note 2: Starred readings are suggested and encouraged, but not required.

First half: January 8th-February 26th

1. Review of matrix algebra, statistics, and asymptotics

- *Greene*, appendices A, B, C, and D.
- *Casella and Berger*, Chapters 5, 7-10.
- *Hogg et al*, Chapters 1-5.

2. Maximum Likelihood Estimation

- Maximum Likelihood Estimation in general, *Greene*, Chapter 16, pp. 482-507.
- Applications of MLE: classical linear regression with normal errors, *Greene*, Chapter 16, pp. 518-522, *Fox**, Chapter 9, 197-198.

3. Classical linear regression model

- The model and its assumptions, *Greene*, Chapter 2, *Fox**, Chapter 5.
- Ordinary least squares, definition and algebraic properties *Greene*, Chapter 3, *Fox**, Chapter 10.
- Ordinary least squares, statistical properties *Greene*, Chapter 4, *Fox**, Chapter 9.
- Ordinary least squares, inference *Greene*, Chapter 5, *Fox**, Chapter 6.
- Omitted variable bias and other model specification issues, *Greene*, Chapter 7, *Fox**, Chapter 6, pp. 110-112.

4. Generalizing the classical linear regression model

- Generalized least squares, *Greene*, Chapter 8, pp. 148-158, *Fox**, Chapter 12.
- Heteroskedasticity, *Greene*, Chapter 8, pp. 158-175.
- Autocorrelation, *Greene*, Chapter 19, pp. 626-634, 640-643, *Fox**, Chapter 16.
- Endogeneity (Instrumental variables estimation) *Greene*, Chapter 12, pp. 314-325, 350-352.

Winter break: Monday March 2nd through Friday March 6th

Second half: March 10th-April 21st

5. Non-linear regression

- Non-linear least squares, *Greene*, Chapter 11, pp. 285-300, *Fox*, Chapter 17.
- Binary choice models, *Greene*, Chapter 23, pp. 770-793, *Fox**, Chapter 14, pp. 335-355.
- Multinomial choice models, *Greene*, Chapter 23, pp. 831-862, *Fox**, Chapter 14, pp. 355-378.

Take-home midterm exam: Friday March 27th to Monday March 30th

6. Panel Data

- Linear panel data models, *Greene*, Chapter 9, *Wooldridge*, Chapter 10.

7. Causal Inference

- Imbens, Guido W. and Jeffrey M. Wooldridge. 2009. "Recent Developments in the Econometrics of Program Evaluation", *Journal of Economic Literature*, 47(1).
- Holland, Paul W., 1986. "Statistics and Causal Inference" (with discussion), *Journal of the American Statistical Association*, 81(396):945-970.
- Angrist and Pischke (2008), Chapters 3 and 4.

Final paper due on Wednesday April 29th

A summary of the relevant due dates is given below:

Table 2: Dates of relevant events and assignments

Event/Assignment	Date/Due date
Replication paper: one-page report	Thu Feb 5
Winter Break:	Mon March 2 through Fri March 6
Replication paper: 4/8-page report	Thu March 12
Midterm exam:	Fri March 27 through Mon March 30
Replication paper: final version	Wed, April 29