PS 581: Quantitative Political Methodology I

Spring 2023

Ted Enamorado Jordan D. Wong

Department of Political Science
Washington University in St. Louis

This course is the first course in our political methodology sequence. Students will learn a variety of *cross-section* regression models including linear regression models, instrumental variables models, and models for missing data. In this class, we will emphasize the connections between these methods and causal inference, which is a primary goal of social science research.

1 Contact Information

Ted Enamorado
Email: ted@wustl.edu

Jordan D. Wong
Email: jordan.d.wong@wustl.edu

2 Logistics

- Lectures: Tuesdays and Thursdays from 10:00–11:20 am. Location: Seigle L006

  We will discuss concepts in depth but will also talk about data analysis. Lecture slides will be posted on Canvas prior to each class.

- Lab sessions: Fridays from 9:00–9:50 am. Location: Seigle 111

  We will take a hands-on approach to data analysis, and put into practice what we have learned during lecture.

- Ted's Office Hours: Tuesdays from 3 to 5 pm (or just make an appointment). Location: Seigle 244.

- Jordan's Office Hours: Wednesdays from 4 to 6 pm. Location: Seigle 278.

  Please, be respectful of your TA's time. We are eager to help, to that end, please write down all your questions before coming to office hours so that we can better assist you.

3 Questions and Announcements

In addition to lab sessions and office hours, please use the **Piazza** Discussion Board at https://piazza.com/ when asking questions about lectures, problem sets, and other course materials. This allows all students to benefit from the discussion and to help each other understand the materials. Both students and instructors are encouraged to participate in discussions and answer any questions that are posted.

To join the PS581 Piazza site, click on "Search Your Classes" from the Piazza homepage. After specifying Washington University in St. Louis as your school, search for "`PS581`" You will then be prompted to enter your `wustl.edu` email address to confirm your registration. All class announcements will be posted on Piazza and Canvas will be used for hosting all class materials.

Some useful tips for Piazza include:

- Piazza has apps available for the iOS and Android platforms. The apps are free downloads and provide complete access to all of Piazza's message board features.
- To insert LaTeX-formatted text in a post, place a double dollar sign ($$) on both ends of the relevant text, or click the fx button in the Details toolbar above your post.
- To add formatted **R** code to a post, click the "pre" button in the Details toolbar above your post. A grey text box will open up where you can paste code from **R**.
- You can classify a post using pre-selected tags, or you can generate your own by prepending a hash (#) to your chosen label. Posts can then be sorted by these tags using the search bar in the left-hand column.

4 Course Requirements

Your final grade is based on lab attendance, the problem sets, and two exams:

- **Labs** (10 points): There will be a weekly lab session where you will work through a guided practical exercise designed to help with the understanding of the materials covered that week and the problem sets. You are required to attend ALL lab sessions. Attendance: 5 points; Lab exercises: 5 points. The grading scheme each lab assignments is $\checkmark-$ (50% of the points), $\checkmark$(75% of the points), $\checkmark+$ (100% of the points).
- **Problem sets** (50 points): There will be four problem sets throughout the semester. Each problem set will contribute equally to the final grade and will contain analytical and data analysis questions.

  For this class, problem sets provide opportunities for individual learning and group collaboration. The following instructions will apply to all problem sets:

  - *Collaboration.* Problem sets for this course present opportunities for students to discuss questions and collaborate to find a solution. Discussions with other people are permitted and encouraged. However, when the time comes to write code that solves the problem, such discussions are no longer appropriate: the code must be your own work.
  - *Submission.* Submit your answers and code as a PDF file (**R Markdown** is highly recommended) to Canvas. Please ensure your code adheres to the tidyverse's **R** Style Guide rules (`https://style.tidyverse.org/`). Late submission will not be accepted unless you obtain a prior approval from the instructor.

- **Exams** (40 points): Two exams, covering the first and second half of the course materials, respectively. Each exam is equally weighted (20 points each).

Grading Scale

A	B	C	D	F
	B+: ≥ 87	C+: ≥ 77	D+: ≥ 67	F : < 60
A : ≥ 94	B : ≥ 83	C : ≥ 73	D : ≥ 63	
A−: ≥ 90	B−: ≥ 80	C−: ≥ 70	D−: ≥ 60	

5 Statistical Software

In this course, we support a statistical computing environment, called **R**. **R** is available for any platform and without charge at http://www.r-project.org. We choose **R** for its flexibility and power. However, students may use other statistical software such as STATA, Python, Julia, etc., for the problem sets and the final project, but at their own risk; that is, we will not be able to answer your software-related questions. Of course, there will be no penalty for using different statistical software. What matters is the analysis you present rather than the software you use.

6 Textbooks

There is no single textbook for this course. However, you may find the relevant parts of the following textbooks useful. You do NOT need to buy any of them.

1. Probability and Statistics

   David A. Freedman. *Statistical Models: Theory and Practice*. Cambridge University Press, Cambridge, 2005.

   Larry Wasserman. *All of Statistics: A Concise Course in Statistical Inference*. Springer, New York, 2005.

2. Econometrics

   Jeffrey M. Wooldridge. *Introductory Econometrics: A Modern Approach*. South-Western Cengage Learning, Mason, OH, 5th edition, 2012.

   Jeffrey M. Wooldridge. *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, Cambridge, MA, 2nd edition, 2010.

   Fumio Hayashi. *Econometrics*. Princeton University Press, Princeton, 2000.

   Bruce Hansen. *Econometrics*. Manuscript, 2019. URL `https://www.ssc.wisc.edu/~bhansen/econometrics/Econometrics.pdf`.

3. Causal Inference

Guido W. Imbens and Donald B. Rubin. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press, 2015.

Stephen L. Morgan and Christopher Winship. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge University Press, New York, 2007.

Joshua D. Angrist and Jörn-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, Princeton, 2009.

7 Other Class Policies

- **Attendance:** you will not be graded directly on your lecture attendance. However, I strongly suggest students expecting to receive an A in this course attend all lectures and arrive early. Lab attendance *is* mandatory. Your grade will reflect missed lab sessions.
- **Technology in the classroom:** During some lecture periods and software training, you will frequently use computers in this course. Please be respectful to your instructor and your peers by using your computers only for class-related purposes. Put your phone away before class starts.
- **Students with disabilities:** Students with disabilities needing academic accommodation should 1) contact the office of Disabilities Resources (`disabilityresources@wustl.edu`); 2) bring a letter to the instructor indicating the need for accommodation and what type. This should be done during the first two weeks of class.
- **Religious observances:** Some students may wish to participate in religious observances during this semester. If you have a religious observance that conflicts with your participation in the course, please meet with the instructor before the end of the second week of the semester to discuss appropriate accommodations.

8 Course Outline

If time permits, we will cover the following topics:

1. Probability and Mathematical Statistics
   (a) Random variables
   (b) Probability distributions
   (c) Expectation & Variance operators
   (d) Law of Large Numbers & Central Limit Theorem
2. Statistical Modeling with Cross-section Data
   (a) Simple Regression
   (b) Review of Linear Algebra
   (c) Multiple Regression and the Geometry of the Linear Model
   (d) The Gauss-Markov Theorem
   (e) Regression Diagnostics
   (f) Model Fit and Model Selection

(g) Cross Validation
(h) Shrinkage estimators for the Linear Model: Ridge and LASSO
(i) Instrumental Variables
(j) Inference for the Linear Model
(k) Models with interactions
(l) Resampling methods: the bootstrap
(m) Missing Data and Multiple Imputation

3. Intro to Causal Inference

(a) Causality and the Linear Regression
(b) Permutation Tests
(c) Mediation Analysis
(d) Instrumental Variables (revisited) and the Wald Estimator
(e) Sharp Regression Discontinuity Designs