

Gov 2002: Quantitative Social Science Methods II

Contents

- Head Instructor
- Teaching Fellows
- Description
- Expectations
- Course objectives
- Prerequisites
- Credit
- Grading
- Lectures
- Sections
- Problem Sets
- Midterm Exam
- Final Exam
- Discussion
- Regrading Policy
- Office Hours and Availability
- Books
- Computing
- Mental Health
- Academic Honesty

Head Instructor

Prof. Matthew Blackwell

mblackwell@gov.harvard.edu

<https://www.mattblackwell.org>

Office Hours: Mondays, 1-3pm ET [Sign up here](#)

Teaching Fellows

Soubhik Barari

sbarari@g.harvard.edu

<http://soubhikbarari.github.io>

Office Hours: TBD

Georgina Evans

georginaevans@g.harvard.edu

<https://sites.google.com/view/georginaevans>

Office Hours: TBD

Description

Quantitative methods in the social sciences are *exploding*. Each year researchers are deploying new and exciting methods to answer important substantive questions. But to use (and not misuse) these novel methods, it is crucial to have a firm understanding of the basic building blocks of quantitative methods in the social sciences: probability, statistical inference, and (more often than not) the linear model. This course, the second in the four-course quantitative methods sequence for PhD students in the Government department, provides this rigorous foundation necessary for the rest of the sequence and the rest of your careers. After reviewing the basic probability and statistical inference, we offer a systematic introduction to the linear model and its variants – the workhorse models for social scientists. We will cover the material with enough mathematical rigor to understand the intuition and concepts, but also cover how to use statistical computing to apply the methods.

Expectations

In this course, you will be expected to

- complete 10 weekly problem sets,
- take one midterm exam,
- take one final exam,
- and participate in the course via Zoom lectures and discussion forums.

Course objectives

After taking this course we hope you will:

- Understand the key concepts of probability for quantitative social science at the level of Stat 110.
- Have a solid understanding of the core foundations of frequentist statistical inference at the level of Stat 111.
- Be able to implement, interpret, and critically evaluate the use of the linear regression model.
- Have a deeper knowledge and familiarity with professional tools for data analysis such as R, git, Rmarkdown, and LaTeX.

Prerequisites

The equivalent of Gov 51 is required and Gov 2001 is highly recommended. We will also assume some familiarity with calculus and linear algebra at the level of the [Gov Math Refresher](#). Because we'll be using more linear algebra than in 2001, we have also set up a mostly self-guided [Gov January Linear Algebra Review](#) for students to complete ahead of the course to get up to speed with those concepts. We assume basic familiarity with R, Rmarkdown, and LaTeX.

No matter your background, you should be prepared to engage the class material on a regular, almost daily basis even beyond the time dedicated to assignments and exam review. This material can be challenging for many students (it was for me!).

Credit

This course satisfies the Methods requirement for the PhD program in the Government department and also can count toward the [methods course out for general exams](#).

Grading

Category	Percent of Final Grade
Participation	10%
Four Problem Sets	55%
Midterm Exam	15%
Final Exam	20%

We will use [Gradescope](#) for submission of the various assignments throughout the semester. Once enrollment is finished, Gradescope will automatically connect through Canvas.

Lectures

Lectures will be Mondays and Wednesdays 10:30am until 11:45am ET, via Zoom. All aspects of lectures except for breakout room sessions will be recorded and made available to students as soon as possible after the lecture.

Sections

We will have weekly section meetings where the Teaching Fellows will guide you through worked problems that are similar to the problem sets. These sections are vital to learning the material and you are strongly encouraged to attend.

Problem Sets

Methods are tools and it isn't very instructive to read a lot about hammers or watch someone else wield a hammer. You need to get your hands on a hammer or two. Thus, in this course, you will have problem sets on a (roughly) weekly basis. They will be a mix of analytic problems, computer simulations, and data analysis.

Given the (waves hands generically at the world) situation, we understand that circumstances might make things difficult this semester. Accordingly, we will be dropping your lowest **two** problem set scores.

The schedule for the problem sets will be:

Problem Set	Release Date	Due Date
Problem Set 1	Thu, Jan 28th 12:00pm ET	Wed, Feb 3rd 11:59pm ET
Problem Set 2	Thu, Feb 4th 12:00pm ET	Wed, Feb 10th 11:59pm ET
Problem Set 3	Thu, Feb 11th 12:00pm ET	Wed, Feb 17th 11:59pm ET
Problem Set 4	Thu, Feb 18th 12:00pm ET	Wed, Feb 24th 11:59pm ET
Problem Set 5	Thu, Mar 11th 12:00pm ET	Wed, Mar 17th 11:59pm ET
Problem Set 6	Thu, Mar 18th 12:00pm ET	Wed, Mar 24th 11:59pm ET
Problem Set 7	Thu, Mar 25th 12:00pm ET	Wed, Mar 31st 11:59pm ET
Problem Set 8	Thu, Apr 1st 12:00pm ET	Wed, Apr 7th 11:59pm ET
Problem Set 9	Thu, Apr 8th 12:00pm ET	Wed, Apr 14th 11:59pm ET
Problem Set 10	Thu, Apr 15th 12:00pm ET	Wed, Apr 21st 11:59pm ET



Midterm Exam

The midterm exam will be an open-book and open-internet checkout exam that is designed to be completed in 75 minutes. Given the stress and potential for technological problems, however, you will have 3 hours to complete it. Since there is a wide distribution of availability and time zones in the class, we will make the exam available to checkout over a 36-hour window on Canvas. Once you check out the exam, you will have 3 hours to complete it. There is no discussion or collaboration with other students or humans permitted on the midterm exam. The exam is **tentatively scheduled** for the sixth week of term (probably March 3-4).

Final Exam

The final exam will be an open-book, open-internet checkout exam similar to the midterm, but designed to take 3 hours to complete and you will be given 6 hours to complete it. As with the midterm, there is no discussion or collaboration with other students or humans permitted on the final exam. The final exam is **tentatively scheduled** for May 12th.

Discussion

We will be using [Ed](#) and [Slack](#) for discussions for this course. You can sign up for the Gov 2002 Ed page using [this link](#) and you can join the Gov 2002 Slack [here](#). To become more familiar with the platforms, please see the [Ed users guide](#) and the [Slack quick start](#).

With two platforms, you might ask: where do I post what? In Gov 2002, Ed will be for help with and discussion of the content and materials of the course, whereas Slack will be for organization and community. Thus, questions and discussions about problem sets, reading, lectures, section, and so on would go on Ed. Slack would be better suited for meeting people, organizing study groups, seeing announcements about class, and various other logistical/social conversations.

When discussing problem sets on Ed, please refrain from posting large portions of your solutions to a question. If, for some reason, you feel you need to post such an excerpt from your solutions to ask your question, you may make the question private (which means only the course staff can view it). Please use this sparingly, since more questions (and answers!) helps the whole group. In addition, please search before posting a question to see if someone else has already posted. Use the categories to help organize the discussions for others to read.

Regrading Policy

If you feel there has been an error in the grading of one your assignment, you may request a regrade of the assignment on Gradescope. A member of the teaching staff will regrade the entire assignment, not just the part you are disputing. Therefore, your regrade might increase or decrease the overall grade on the assignment.

Office Hours and Availability

My office hours are TBD. If you have questions about the course material, computational issues, or other course-related issues please do not hesitate to set up an appointment with either any of us.

If you have a general question, you can also post it on Ed. This is almost always the fastest way to get an answer. However, you can also email me directly at mblackwell@gov.harvard.edu. If the question is of general interest, I will forward the question and my answer to the class. Make sure to tell me explicitly in your email if you would like to stay anonymous.

Books

Unfortunately, there are no perfect references that cover everything we do in Gov 2002. I will mostly try to follow material from a couple of different books, which are currently all freely available on the internet.

- Blitzstein, Joe and Jessica Hwang. *Introduction to Probability, Second Edition*. Freely available to read online or for purchase on [Amazon](#) if desired.
- Hansen, Bruce. *Introduction to Econometrics*. Covers probability and inference.
- Hansen, Bruce. *Econometrics*. Covers regression and beyond.

There are several other books for purchase that may be useful to you.

- Freedman, David. *Statistical Models*. Cambridge University Press.
- Wasserman, Larry. *All of Statistics: A Concise Course in Statistical Inference*. Springer. Available to download via the Springer website (may need to log in through Harvard).
- Hayashi, Fumio. *Econometrics*. Princeton University Press.
- Angrist, Joshua and Jorn-Steffen Pischke. *[Mostly Harmless Econometrics]*. Princeton University Press.
- Aronow, Peter and Benjamin Miller. *Foundations of Agnostic Statistics*. Cambridge University Press. Covers much of the same basic material that we cover in the class.

Sometimes it is helpful to digest certain concepts when they are presented with less mathematical notation. The following books can be extremely useful in this regard:

- Imai, Kosuke *Quantitative Social Science: An Introduction*. Princeton University Press.
- Diez, David M., Christopher D. Barr, and Mine Cetinkaya-Rundel. 2015. *OpenIntro Statistics*. 3rd edition.
<https://www.openintro.org/>
- Freedman, David, Pisani, Robert, and Purves, Roger. 2007. *Statistics*. W.W. Norton & Company. 4th edition.
- Gonick, Larry, and Woollcott Smith. 1993. *The Cartoon Guide to Statistics*. HarperPerennial.

Computing

We'll use R in this class for computing and data analysis. R is free, open source, and available on all major platforms (including Solaris, so no excuses). [RStudio](#) (also free) is a graphical interface to R that is widely used to work with the R language. You can find a virtually endless set of resources for R and RStudio on the internet. For beginners, there are several web-based tutorials. In these, you will be able to learn the basic syntax of R. We'll post more R resources on the course website.

We will also use git and Github to manage our projects, and a combination of LaTeX and Rmarkdown to typeset the problem sets.

Mental Health

Grad school is a stressful time in one's life and mixing it with a global pandemic, remote learning, and dislocation makes this one of the most fraught years any of us have probably faced. We acknowledge that nothing is quite normal and that there may be times when you feel overwhelmed by this course or by life more generally. Please feel free to reach out to any of the course

staff if you want to talk about any issues you are having with the course or anything else. We will always try to help and we are committed to being extra accommodating this semester on course policy issues. Please just get in touch.

Of course, there are other resources at Harvard if you need them. A few are listed below:

- [Counseling and Mental Health Services](#)

Academic Honesty

The work that you do in both the problem sets should be your own work. You may seek help from others so long as this does not result in someone else completing your work for you. When asking for help, you may show others your code to help diagnose a bug or highlight a potential issue, but you should not view their (working) code. You should cite any discussions you have with other students in your problem set and note if they helped you with your code. You should never copy and paste code from another student or elsewhere (e.g., websites, former students).

I also strongly suggest that you make a solo effort at all the problems before consulting others. The exams will be very difficult if you have no experience working on your own. **There is no collaboration allowed on the exams.**

GOV 2002: Quantitative Social Science II (Spring 2021)

[Harvard University](#) . . . [Government Department](#)

 [Matthew Blackwell](#) . . .  mblackwell@gov.harvard.edu

All content licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](#).

Content © 2021

 [View the source at GitHub](#).

 Tuesdays . . . 

 Virtual