GVPT 729: Quantitative Analysis of International Political Economy and International Security
Spring 2022

Tuesdays 11:00 am to 1:45 pm (EST) in Susquehanna Hall 1101
(There will be a short break at 12:00pm.)

Instructor: Dr. Jacob Aronson
Email: jacobaro@umd.edu
Office hours: **Tuesday** 2:00 pm to 3:00 pm and by appointment
Office: 2117H Chincoteague Hall (in the CIDCM section)

Course Description and Goals
This course builds on knowledge gained during 622. In 622, students learned the basics of statistical analysis, the relationship between data distributions and regression, and how to interpret linear models. In 729, students will learn how to analyze real-world datasets, how to run and interpret generalized linear models ("glm"), how to produce useful predictions from models, and the basics of causal analysis.

During the semester, this course will combine teaching the principles of statistics as used in political science with hands-on data analysis. Some class periods (noted in the syllabus below) will involve in-class exercises. The class will also continue student training in how to use R for data processing and for statistical analysis.

The class is divided into four sections.

- The first section introduces the class and reviews what was learned in 622. We will also discuss how ChatGPT may be used.
- The second section will teach students how to use and understand generalized linear models. For the rest of the course, students will need to know how to install and use R and the important R package 'Tidyverse.'
- The third section will teach students about the incorporation of data in a research design and will cover important topics such as the use of post-stratification in prediction.
- The fourth section will cover the basics of causal analysis, identify how bias affects interpretation of results, and present the logic of matching.

Once 729 is completed, students will know not only how to use models for continuous outcomes but also how to use models for outcomes that are binomial or are counts. This covers most uses in political science and related fields.

Course Reading and Software Requirements
One book listed below is required for this course. This is the same book that was utilized in GVPT 622. You must also have access to (and use) the statistical package R. Students are

required to conduct work on their own laptop, which they will need bring to class when we work with RStudio. R can be downloaded for free at: https://cran.r-project.org/. RStudio, an integrated development environment (IDE) that makes it easier to program, can be downloaded for free at: https://www.rstudio.com/products/rstudio/download/. You know the drill!

One book is required for this course:

- Andrew Gelman, Jennifer Hill, and Aki Vehtari. 2020. *Regression and Other Stories*. Cambridge University Press.

While the book introduces students to both classical and Bayesian inference, we will focus on sections of assigned chapters that discuss classical inference as it is more common in applied political science.

Any reading materials aside from the textbook will be posted on the course ELMS page.

Simple R tutorials are available at https://www.w3schools.com/r/default.asp and at https://education.rstudio.com/learn/. The back of the textbook (Appendix A) also provides information on how to program in R. Resources on Tidyverse (the R package we will use extensively) is available at https://www.tidyverse.org/learn/. More complicated tutorials on using Tidyverse for data science are available at https://www.tidymodels.org/books/moderndive/. Programming in R is a challenging but beneficial skill to know—it is worth the investment.

Integration of ChatGPT (GPT-3.5)

In week two, we will briefly cover how to use the most recent version of the ChatGPT language model. Use of ChatGPT is neither required nor needed to get a maximum grade in this class. Its use, however, is *allowed*—and may be helpful!

The main link that allows access is https://chat.openai.com/chat. To use the service, you will need to sign up for a (free) account. Try to sign up as soon as possible. Due to its popularity, it may not be possible to sign up on that page. An alternative is to use the beta link at https://beta.openai.com/playground. The beta has additional features such as the ability to read and interpret files.

This tool is new. Nobody currently knows exactly how to integrate it into their course, but many believe it will become ubiquitous. For this tool to be beneficial to you, it will be important to figure out the right questions to ask and how to properly integrate a response into your own work (including homework and exams). Pasting a response verbatim will result in a low grade. We will learn together and modify course standards (e.g., grading criteria) as needed. Do not hesitate to raise suggestions or concerns.

My belief is that use of this tool will raise the floor of human achievement. Basic tasks will be easier to conduct, which will allow students to focus on answering more challenging questions. This is analogous to the introduction of the calculator. Students were able to focus less on memorizing ‘times tables’ and focus more on other beneficial tasks/problems.

Students can also ask the AI anything. If a concept mentioned in class or an R task is not clear, the AI can be asked to help explain it discretely and in real time. The AI can even explain code that has been cut-and-paste into its prompt.

Campus Policies
It is our shared responsibility to know and abide by the University of Maryland's policies that relate to all courses, which include topics like:

- Academic integrity
- Student and instructor conduct
- Accessibility and accommodations
- Attendance and excused absences
- Grades and appeals
- Copyright and intellectual property

Please see the University's website for graduate course-related policies at: https://gradschool.umd.edu/course-related-policies.

Course Requirements and Evaluation
Your final grade will reflect the sum of points earned from each of the following categories:

Assignments:

- In-class Assignment (February 28th): 5%
- Homework Assignment #1 (March 7th): 10%
- Homework Assignment #2 (March 28th): 10%
- Homework Assignment #3 (April 18th): 10%
- Homework Assignment #4 (May 2nd): 10%

Exam and participation:

- Class Participation: 15%
- Midterm Exam (April 4th): 15%
- Final Exam (May 16th): 25%

Homework assignments
The four homework assignments will be composed of several questions pulled from or inspired by the list of exercises at the end of each chapter in Gelman, Hill, & Vehtari (2020). The homework will include questions only from the weeks specified in the assignment (listed in the week it is due). Students are free to work together to solves these problems. Because exams are individual, it will be important for students to understand how homework answers were derived and why.

All homework assignments must be turned in BEFORE CLASS BEGINS ON THE DAY THAT THE ASSIGNMENT IS DUE. Late assignments will not be accepted. It is much better to get half credit than no credit! I will go over each homework assignment in class on the day that it is

due. If a student has an excused absence based on university policies, a makeup assignment can be submitted the following week.

Midterm exam
The in-class midterm exam will involve questions from the exercises at the end of each chapter of Gelman, Hill, & Vehtari (2020) and questions about data used in international relations for a total of ~12-15 questions. The midterm will cover the first three sections, including chapters 13-17 of the textbook. This exam will need to be completed individually in-class during the class period. There will be *no* surprise questions so students who read the assigned articles, feel comfortable with the homework assignments, and are familiar with the exercises at the end of each chapter will be well-prepared. Questions will require you to both write a response and produce and run R code.

Final exam
The in-class final exam will involve questions from the exercises at the end of each chapter of Gelman, Hill, & Vehtari (2020) for a total of ~15-18 questions. These questions will be comprehensive, which means that several of the questions may be like those on the midterm. This exam will need to be completed individually in-class during the final exam period at the end of the semester. There will be *no* surprise questions so students who read the assigned articles, feel comfortable with the homework assignments, and are familiar with the exercises at the end of each chapter will be well-prepared. The exam format will be very similar to the final exam in 622. Questions will require you to both write a response and produce and run R code.

Class participation
Students must be prepared to develop and run R code during select class periods. This means that students will need to bring a laptop into class on those weeks. Students who successfully conduct in-class activities will receive full credit for class participation. Do not hesitate to ask questions. Questions help everyone to better understand the material and contribute to your grade!

Course Evaluation
Please submit a course evaluation through CourseEvalUM to help faculty and administrators improve teaching and learning at Maryland. All information submitted to CourseEvalUM is confidential. Campus will notify you when CourseEvalUM is open for you to complete your evaluations for fall semester courses. Please go directly to the Course Eval UM website to complete your evaluations.

Grading Scale
Grades will be based on a straight scale.

100 to 98 = A+	74 to 76 = C
94 to 97 = A	70 to 73 = C-
90 to 93 = A-	67 to 69 = D+
87 to 89 = B+	64 to 66 = D
84 to 86 = B	60 to 63 = D-
80 to 83 = B-	0 to 59 = F

77 to 79 = C+

Course Schedule

Section 1: Introduction

Week 1 Introduction and Review of Linear Regression January 31st, 2022

This session will describe the format of the class, the topics covered, and identify class goals.

During class we will also briefly review topics covered in GVPT 622. Topics include creating fake data, running and interpreting the results of linear regression, and the connection between data and result.

Week 2 ChatGPT, Datasets February 7th, 2022

This session will cover two important topics: (1) use of ChatGPT, and (2) important datasets used in the study of international relations.

(1) ChatGPT is a natural language model released recently by OpenAI. It is trained using the bulk of human knowledge available through 2021. It cannot produce new knowledge but does a pretty good job of relaying existing knowledge (e.g., it can program for you, conduct statistical operations, and answer most statistics questions).

(2) We will briefly discuss the unit of analysis and the information available in a commonly used set of datasets (UCDP; available at: https://ucdp.uu.se/).

In addition, we will cover interaction terms (chapter 10.3):

- Gelman, Hill, and Vehtari, *Regression and Other Stories*, 2020. Chapter 10.3.

Section 2: Generalized Linear Models

Week 3 Binary Outcomes I: Logistic Regression February 14th, 2022

- Gelman, Hill, and Vehtari, *Regression and Other Stories*, 2020. Chapter 13.

Week 4 Binary Outcomes II: Graphing and Diagnostics February 21st, 2022

- Gelman, Hill, and Vehtari, *Regression and Other Stories*, 2020. Chapter 14.

During this class period, we will also discuss the use of interaction terms in logistics regression models.

Week 5 Binary Outcomes III: Applications February 28th, 2022

- Jason Lyall and Isaiah Wilson III. 2009. "Rage Against the Machines: Explaining Outcomes in Counterinsurgency Wars." *International Organization* 63(1): 67-106.

This class period will apply knowledge we learned about generalized linear models (logit) in the prior two weeks. An in-class exercise will involve downloading replication data and recreating analysis from the identified article.

In-class Assignment

Week 6 Count Outcomes I: Poison Regression March 7th, 2022

- Gelman, Hill, and Vehtari, *Regression and Other Stories*, 2020. Chapter 15.

Homework Assignment #1 (covering weeks 2-5) due.

Section 3: Designing a Research Agenda

Week 7 Research Design I: What the Data Allows March 14th, 2022

- Gelman, Hill, and Vehtari, *Regression and Other Stories*, 2020. Chapter 16.

Week 8 **No Class (Spring Break)** March 21st, 2022

Week 9 Research Design II: Pre- and Post-processing March 28th, 2022

- Gelman, Hill, and Vehtari, *Regression and Other Stories*, 2020. Chapter 17.

Homework Assignment #2 (covering weeks 6-7) due.

Week 10 **Midterm Examination** April 4th, 2022

The midterm exam will be held in class. The midterm will cover material covered in sections 1-3 (including book chapters 13-17).

Midterm Exam

Section 4: Causal Inference

Week 11 Causal Inference I: Theory and Basics April 11th, 2022

- Gelman, Hill, and Vehtari, *Regression and Other Stories*, 2020. Chapter 18.

Week 12 Causal Inference II: Analyzing Experiments April 18th, 2022

- Gelman, Hill, and Vehtari, *Regression and Other Stories*, 2020. Chapter 19.

Homework Assignment #3 (covering weeks 9-11) is due.

Week 13 Causal Inference III: Using Observational Data April 25th, 2022

- Gelman, Hill, and Vehtari, *Regression and Other Stories*, 2020. Chapter 20.1-20.6. (First half of chapter 20.)

Week 14 Causal Inference IV: Matching May 2nd, 2022

- Gelman, Hill, and Vehtari, *Regression and Other Stories*, 2020. Chapter 20.7-20.9. (Second half of chapter 20.)
- Daniel E. Ho, Kosuke Imai, Gary King, and Elizabeth A. Stuart. 2008. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis* 15: 199-236.

Homework Assignment #4 (covering weeks 12-13) is due.

Week 15 Review May 9th, 2022

This class will focus on review of material covered, and address any questions students have from the reading and concerns about the final. If there is enough time, we may also cover a practice problem like one that may appear on the final.

*****Final Exam will be on May 16th from 11:00am to 1:00pm*****