

## CCJS 621: General Linear Models in Criminal Justice Research

Instructor	Greg Midgette ( <a href="mailto:gem@umd.edu">gem@umd.edu</a> ) 2220L LeFrak Hall
Lecture	Monday 4 pm – 6:45 pm 2165E LeFrak Hall (Large Conference Room)
TA	Madisen Placzkowski ( <a href="mailto:mplacz@umd.edu">mplacz@umd.edu</a> ) 2220C LeFrak Hall
Lab	TBA
Office Hours	Greg: Wed 2 pm – 3:30 pm and by appointment Madisen: TBA
Course Description	<p>This course is an in-depth exploration of applied linear regression analysis and the broader class of generalized linear models. Linear regression analysis, and inferential statistics more generally, are based on the desire to estimate the relationship between two or more variables. We would like these estimates to be unbiased, efficient, and consistent. Short of those goals, we should learn what can be done when they're not. That's the aim of this course.</p> <p>You will learn econometric techniques and how to interpret their results. These skills are necessary, but rarely sufficient to be good researchers. Rather, you need to <i>at least</i> understand the theoretical issues involved in the basic model using the simplest possible case, the bivariate model, and you should be familiar with further complexities of statistical inference from regression models. You will gain fluency with fundamental theoretical issues, familiarity with the general model, and comfort with the computer application of multivariate regressions and binary response models are basic goals.</p>
Instruction Format	Weekly in-person instruction led by Prof. Midgette and a lab with recitation with examples led by the TA. For the sake of continuity and uniformity, we'll use Stata in this course.
ELMS	The course syllabus, links to lecture material, and your grades will be available on <a href="http://elms.umd.edu">elms.umd.edu</a> for all students that are registered for the class.

- Course Prerequisites** Prior exposure to statistics at the level of CCJS 620. Familiarity with personal computers and the BSOS computing lab will be assumed. Assistance with Stata will be provided within the context of the class.
- Texts** Wooldridge, Jeffrey M., *Introductory Econometrics: A Modern Approach*, 6<sup>th</sup> or 7<sup>th</sup> ed., Cengage Learning (pdf of 6<sup>th</sup> ed. will available on ELMS per tradition, but I'd buy a physical copy).
- Note: there are many potentially useful complements to Wooldridge (e.g., Greene's *Econometric Analysis*, Kennedy's *A Guide to Econometrics*, and Agresti & Finlay's *Statistical Methods for the Social Sciences*).
- Assessment** Your grades will be based on your performance on periodic homework assignments, one project, and two examinations. The weights for each assignment in your final grade are listed below. Take-home assignments are to be handed in via ELMS by 4PM on Monday (project due dates are listed below and homework assignments are due one week after they are assigned). Late assignments will lose 20 points each day they are late.
- Homework (15%).** You can collaborate on homework assignment, but each student must turn in a separate assignment and is responsible for the content of that assignment. *To receive full credit, you must show your work and include your Stata do and log files.*
- Project (25%).** The project will be completed in three stages and will be peer reviewed. Each student will complete a project that demonstrates knowledge of the methods learned in class. This project may take the form of an empirical section "attached" to a thesis proposal or the actual thesis using pre-existing data. Your TA will provide an overview of how to find research data. The due dates and specific instructions for each segment are at the end of the syllabus.
- The assignment is designed to facilitate your work with the data that you intend to use in your thesis. Of course, this might prove impossible. Those of you who aspire to collect your own data for your thesis will probably need to find something to use for this class. It is also possible that the dataset you're interested in using is be very large and complicated. In that case, it is appropriate to use a manageable subset of the data, or a related, more accessible dataset. The point of the project is to give you some practice working with quantitative data in the context of a criminological problem that interests you.*

**Midterm and Final (30% each).** The final exam will focus on the material since the midterm exam. However, understanding of this material will require you to integrate knowledge from the earlier part of the course. The exams will be open notes and open book, and they will be strictly limited to 2 hours and 45 minutes.

Grade Definitions	A+ 98% - 100%	B- 80% - 82%	D 63% - 67%
	A 93% - 97%	C+ 78% - 79%	D- 60% - 62%
	A- 90% - 92%	C 73% - 77%	F Less than 60
	B+ 88% - 89%	C- 70% - 72%	
	B 83% - 87%	D+ 68% - 69%	
Academic Integrity	The University of Maryland, College Park Code of Academic Integrity sets standards for academic integrity at Maryland for all graduate students. Academic dishonesty will not be tolerated.		
Class Conduct	We are expected to treat each other with respect. You are expected to adhere to the Code of Student Conduct. For more information, see: <a href="http://president.umd.edu/administration/policies/section-v-student-affairs/v-100b">president.umd.edu/administration/policies/section-v-student-affairs/v-100b</a>		
Religious Observances	The University of Maryland policy on religious observance states that students should not be penalized in any way for participation in religious observances and that, whenever feasible, they should be allowed to makeup academic assignments that are missed due to such absences. However, the student must personally notify the instructor within two weeks of the start of the semester. The request should not include travel time.		
E-mail	<b><i>Don't email me through ELMS.</i></b> Please send all messages to my university email address and include "CCJS621" in the subject line of all emails regarding this course. I will generally respond quickly to your emails, but there may be times when I am unable to do so.		
Self-Identification	The University of Maryland recognizes the importance of a diverse student body, and we are committed to fostering equitable classroom environments. I take this very seriously. I invite you, if you wish, to tell me how you want to be referred to both in terms of your name and your pronouns (e.g., she/her, they/them, he/him). Visit <a href="http://lgbtq.umd.edu">lgbtq.umd.edu</a> to learn more.		
Disability Support	Please contact me by February 14 to discuss a plan for accommodations for any documented disability consistent with UMD Accessibility and Disability Service policy.		

## Course Outline

Week	Topic	Chapter	Project Milestones
1 1/29	Review	1; Math Refreshers B & C	
2 2/5	Fundamentals of regression	2	
3 2/12	Multiple regression	3	
4 2/19	Inference from multiple regression	4	Part 1 due
5 2/26	Asymptotics (consistency and efficiency)	5	
6 3/4	Functional form and residuals	6	
3/11	<i>Midterm Exam</i>		
3/18	Spring Break		
7 3/25	Dummies as Predictors / Project Planning	7	
8 4/1	Dummies as Outcomes (LPM)	7	Part 2 due
9 4/8	Heteroskedasticity	8	Part 2 peer reviews due
10 4/15	Some measurement challenges in regression	9	
11 4/22	Limited dependent variables	17	
12 4/29	Panel methods	13; 14.1-3	Part 3 due
5/6	<i>Project presentations</i>		
5/13	<i>Final Exam</i>		

## Project Timeline

All project components should be submitted via ELMS in a Word file (.docx). If you use some other software, give yourself time to convert your file to Word. I'll give you feedback as annotations, and for me that's easiest in Word. *Your text should be double-spaced with 12-point Times/Times New Roman and you should use 1-inch margins.*

***Part 1 (due February 19): 2-page description of the data including a short description of the problem.*** This should include a theoretical justification of the hypothesized relationship, identification of the data source, the current status of data (e.g., downloaded, in the mail, etc.), an enumeration of the sample (count of observations and description of key variables), a description of the strengths and weaknesses of the sample relative to the problem at hand, and a list of the relevant measures including their measurement units. *Be sure to define your unit of analysis and your conceptual independent and dependent variables.* The dependent variable can be either continuous or binary and you should have a minimum of 30 observations (preferably  $\gg 30$ ). The problem should be stated in terms of causality where you are interested in the causal impact of *one* independent factor on *one* dependent variable, while keeping in mind that your analysis will include other “control” variables and will probably fall short of establishing causality (No offense. Most studies fall short of establishing causality.)

***Part 2 (due April 1): 3-to-4-page double-spaced written description of the data.*** Tables and figures should be attached at the end of the document and all tables should be formatted in Word—not just pasted output or an image from Stata, R, or Excel. As a guide, look at tables in *Criminology*, *Criminology & Public Policy*, and *Journal of Quantitative Criminology*. The emphasis should be on describing the relationship between 2 key variables using prose to discuss tables and graphs, but you should also explore the other model covariates to identify possible problems in the regression. For example, high correlation between covariates may lead to problems with multicollinearity, and skew in the dependent variable could lead to heteroskedastic error terms. These are solvable problems.

Regression output in this part of the project is not required, but preliminary analyses are OK. This document will be shared with a peer reviewer who will provide comments in addition to my comments. Start with a brief description of the problem you are addressing, the theoretical motivation for hypothesis, and a brief introduction to the data. This is especially important because the outside reviewer will probably not be familiar with your topic. *You may change your*

*topic between Part 1 and Part 2, but you should discuss this decision with both your TA and me.*  
**Peer reviews are due on April 8.**

***Part 3 (due April 29): 6-to-8-page written description (inclusive of tables) of a multivariate regression analysis with hypothesis test(s) on the variables of interest.*** In addition, *tests on the assumptions should be performed in relation to the estimate for the primary independent variable and addressed if there is a problem.* Furthermore, model diagnostics should be performed to determine if your findings are driven by any outliers. Finally, a written description of ways to improve the analysis with alternative methods should be included. The econometric reasoning for each approach should be made explicit. As always, be sure to introduce your problem and explain the nature of the data. You'll present this work in class at the end of the semester (the parameters of presentation depend on enrollment).