University of Michigan School of Education

EDUC 799: CATEGORICAL AND LIMITED DEPENDENT VARIABLE MODELING

Fall 2022

Lecture: Tuesday
1:00-3:50pm
SOE 2328

Lab: Friday
10:00am to Noon
SOE 2114

Lecture Instructor:

Steve DesJardins
Professor, CSHPE
2108C School of Education
Email: sdesj@umich.edu
Office: (734) 647-1984
Cell: (734) 904-7857
Office hours: by appointment

Lab Instructor:

Josh Skiles
Doctoral Student, CSHPE
2022-M School of Education
Email: jskiles@umich.edu
Office: ----------
Cell: (276) 224-9630
Lab office hours: Fridays 12:15pm to 1:15 pm

COURSE DESCRIPTION:

This course focuses on regression methods in which the dependent variable is non-continuous and the application of those methods to education-related research questions. These types of outcomes are ubiquitous in education and social science research and include Likert scales, binary variables like whether a student graduates from high school or not, categorical variables such as the type of college or major a student enrolls in, or which political candidate a person votes for. However, employing the general linear model to study these types of outcomes frequently violates underlying assumptions and may produce misleading results. Non-linear regression methods are now widely used as alternatives to linear regression when studying non-continuous outcome variables. Detailed knowledge of such non-linear methods is necessary for quantitative researchers, but also for qualitative researchers who want to be critical consumers of studies that employ such approaches.

As an education course, we will illustrate and practice these methods using education data, but students are welcome to use any data they wish for the final project. All assignments and lab activities will use Stata 17 available on lab machines around campus and via the virtual enclave. The course is open to graduate students who are comfortable with multiple linear regression methods, which is the foundation and point of departure for the methods we cover. Like most statistics classes, the learning in this class is cumulative, with early content providing the foundation for later learning. Therefore, if you find yourself struggling with any of the course material don't wait to approach us for help. We strive to create a learning environment of high expectations but also high support.

LEARNING OBJECTIVES:

The course is designed with three learning objectives in mind:

(a) Understanding the need for and statistical properties of non-linear probability models,
(b) Estimating, interpreting, and optimizing multivariate models with categorical and limited dependent variables as their outcomes, and
(c) Applying these methods to answer research questions and questions of practice about education.

The curriculum, instruction, and assessments therefore emphasize technique, interpretation, and implications of results for decision-making.

ORGANIZATION OF THE COURSE:

Our Tuesday meetings will emphasize (1) theory underlying the methods we study, (2) sometimes walking you through an example of the code and output in Stata, and (3) discussion as well as time for questions/answers. Preparation for lecture primarily involves reading. I strongly encourage you to read the Long text **before** coming to class, even if you need to browse it again later.

Much of the class is devoted to development of technical skill in and theoretical understanding of a body of methods. These readings, and the discussions we will have about them in lecture, will foster awareness of and critical thinking about how scholars have conducted the type of work we are learning about and the context in which the practice of statistics occurs. We are more than number crunchers or analysts; we are part of a long history and conversation about knowledge production.

Lab time on Friday will be devoted to (1) reinforcing material discussed in lecture, (2) exercises using Stata, and (3) consultation about assignments and final projects. Both instructors are also available during office hours to provide assistance with assignments and final projects.

The course is organized as follows. Please note that the dates on which these topics will be discussed may be adjusted based on how quickly or slowly we progress through the materials, conflicts within the semester, and particular learning needs of the group.

Week No. & Topic Area	Meeting Dates	Assignments Due this Week
1) Introductions /Linear Regression/ Stata Intro	Aug 30, Sept 2	
2) Linear Probability and Latent Variable Models	Sept. 6, 9	
3) Logistic Regression	Sept. 13, 16	Assign. 1
4) Logistic Regression (cont'd)	Sept. 20, 23	Assign. 2 & Paper proposal
5) Probit Regression	Sept. 27, 30	
6) Hypothesis Testing and Goodness of Fit	Oct. 4, 7	Assign. 3
7) Out-of-Sample Prediction	Oct. 11, 14	Assign. 4, Lit Review
8) Fall Break (Ordinal Outcomes lecture on Friday)	**Oct. 21**	
9) Multinomial Outcomes	Oct. 25, 28	Assign. 5, Add Methods section
10) Multinomial Outcomes (cont'd)	Nov. 1, 4	
11) Interaction Effects	Nov. 8, 11	Assign. 6
12) Conditional Logit	Nov. 15, 18	Assign. 7, Add Results section
13) Thanksgiving Break (no lab)	**Nov. 22**	
14) An Introduction to Causal Inference	Nov. 29, Dec. 2	
15) Causal Inference (cont'd)	Dec. 6	
16) In-Class presentations and turn in papers	Dec. 13	

COURSE REQUIREMENTS:
Grading will be on the A-F scale for five (5) semester credits. Incomplete grades are **strongly discouraged**, but if needed they must be arranged with Steve before the end of the last regular class meeting on December 2. Your grade will be calculated based on the following requirements:

a. ATTENDANCE, ENGAGEMENT WITH READINGS, PARTICIPATION (25% of course grade)
Attendance is required for the lecture and lab, except in cases of contagiousness and conference travel. If you are sick, please stay home and notify the instructor for lecture or lab, as appropriate. If you have conference travel, please work this out well in advance with Steve. More than one absence, except in unusual circumstances that have been discussed in advance, will result in you being asked to drop the class.

Participation means actively supporting our collective learning. Josh and I are committed to creating a focused, non-competitive atmosphere in which we work together to build mastery in methods and think about their use. To this end, I expect engagement through asking and answering questions (of the instructors and one another) and contributing to discussion in ways that demonstrate you have given thought to the weekly readings.

Scholars are frequently called upon to conduct manuscript reviews for their colleagues - either formally for a journal, conference, publishing house, etc. or informally to guide revisions before the manuscript is submitted. However, graduate students rarely get practice in this type of review. Therefore, one element of your participation grade will be writing up a review and leading the class in discussion of an article that uses one of the methods we are studying. You will have the opportunity to sign up for this activity early in the class.

b. DELIVERABLES

i. WEEKLY ASSIGNMENTS (35% of final grade)

a. Sometimes, the weekly assignments will involve conducting statistical analyses of education data, other times they will involve critiquing published research, and still other assignments may be designed for you to demonstrate your ability to write-up the results of statistical analysis. The assignments will give you practice in such important skills as critical reading, coding, interpretation, and communicating your ideas in writing.

b. An electronic copy of the assignment on the specified due dates must be submitted, usually at the start of lecture or lab when they are due. Note that due dates include one lab session in between to allow for discussion of any challenges you encountered in the design, coding, and/or interpretation of results.

c. When naming your assignments, please use the following naming convention:
[Your LastName]_Assignment[#].[docx]
Example: Jones_Assignment1.docx. All assignments must be submitted in MS Word format. For the final project assignments, please use the naming conventions below. If you are working in a group, please list the last names alphabetically.
[Lastname]_Step1.[docx]
[Lastname]_Step2.[docx]
[Lastname]_Step3.[docx]…
[Lastname]_Step4.[docx]
[Lastname]_FinalPresentation.[docx]
[Lastname]_FinalPaper.[docx]

d. Weekly assignments are graded using rubrics. For most assignments, you will receive separate grades on two dimensions: (1) The quality of technical work, and (2) the clarity and accuracy of your communication.

e. Late assignments will negatively affect your grade on the assignments' substantive dimension, and your grade in the course.

ii. COURSE PROJECT (40% of final grade)

The course project begins early in the semester, and will develop into an original, rigorous, conference-ready inquiry by the end of the semester, including a ~20 page paper (double spaced; 1 inch margins; excluding tables, figures, references, appendices) and a 15-20 minute final presentation. The paper and presentation should focus on a novel question about education, and should use one or more of the methods we cover in the course, using any data you wish.

You are invited to collaborate with another member of the class on the course project, although this is not required. The project involves deadlines in several stages, and at each stage of the process you will receive feedback from one or both of the instructors. Only the final paper and presentation will be graded. You will have the rubric for the final paper and presentation well in advance of their due date. Please use APA format for all assignments turned in related to your course project (unless you are from a discipline in which another publication style is normative; if this is the case, please notify the instructors).

i. Abstract/ Proposal: Sept 20
ii. Literature review: Oct 11
iii. Revised literature review and draft of proposed methodology. You will turn this in and briefly present it to the instructors and class for feedback: October 25
iv. Revised literature review and methodology, and draft of results: Nov 18
v. FINAL WEEK: Complete paper, including introduction and conclusion and a formal 15-20 minute presentation, including feedback from instructors and class: Dec 6 or maybe

Dec 13. **(This is a hard deadline, to which no exceptions will be made except in the case of prior arrangement with the instructors or a documented emergency.)**

TEXTBOOKS:

Required:

J. Scott Long (1997). *Regression Models for Categorical and Limited Dependent Variables*. Advanced Quantitative Techniques in the Social Sciences, Volume 7. Sage Publications. ISBN 0-8039-7374-8.

J. Scott Long and Jeremy Freese (2014). *Regression Models for Categorical Dependent Variables Using Stata, **3rd Edition***, Stata Press, ISBN-978-1-59718-111-2.

Make your life easier this semester and in the years to come and buy this companion to the textbook. It is widely regarded as one of the most useful books that a quantitatively oriented Ph.D. student in the School of Education can buy. It contains a wealth of practical information on coding in Stata, model estimation, postestimation testing, and so forth.

Both of these required books are available at multiple sources online. Additionally, many students have taken the course so it may be possible to get a used copy from them.

Additional Required Readings:

You will find additional weekly readings on Canvas. Look under Files/Additional Readings tab.

Optional:

M. Mitchell (2012). *Interpreting and Visualizing Regression Models Using Stata*. Stata Press, ISBN1- 59718-107-2

Another Stata companion. Most of this book is not about logistic or other categorical DV modeling, and I am making copies of the required chapters available on Canvas. However, it is such a great resource for using *margins* and *marginsplot* to model and display findings and interaction terms across many types of regression that I am making it available in the bookstore.

The three books mentioned above are also available directly from Stata at: http://www.stata.com/bookstore/cldv.html. In the past, the Stata website has had the lowest prices for new copies.

For additional background on the estimation technique that underlies the methods we will study, pick up Scott Eliason's little green Sage book, Maximum Likelihood Estimation: Logic and Practice, Sage Publications ISBN 0803-941072. It is available at Amazon:

https://www.amazon.com/Maximum-Likelihood-Estimation-Quantitative-Applications/dp/0803941072/ref=sr_1_1?s=books&ie=UTF8&qid=1533585821&sr=1-1&keywords=Maximum+Likelihood+Estimation%3A+Logic+and+Practice

For an additional primer on categorical and limited dependent variable modeling, please see the cited handbook chapter below. It is on Canvas in the Files/Additional Reading tab.

Rodriguez, A., Furquim, F., & DesJardins, S. L. (2018). Categorical and Limited Dependent Variable Modeling in Higher Education. In *Higher Education: Handbook of Theory and Research* (pp. 295-370). Springer.

If you are new to Stata, you may want to purchase an introduction. A good one is: A *Gentle Introduction to Stata* by Alan C. Acock (see http://www.stata.com/bookstore/gentle-introduction-to- stata/). However, *many* websites (e.g., https://stats.oarc.ucla.edu/stata/) contain information about using Stata, including coding hints, sample programs, etc., so use your judgment about whether you need this additional book or not.

Also, Stata Press has a YouTube channel that offers helpful videos and short tutorials on how to use the software. The channel is available at: https://www.youtube.com/user/statacorp/featured

SOFTWARE:
For this course, you will need access to a recent copy of Stata, the statistical software with which we will conduct analyses. The Sites campus computing labs, including the one on the second floor of the School of Education, all have Stata's SE software installed. You can also access the software from any computer with a reliable internet connection using the library's remote desktop system, virtual sites: https://documentation.its.umich.edu/node/312. However, you may find this a somewhat unreliable way to work, as the system will drop you if the internet connection is broken for any reason.

All students enrolled in this course in Fall 2022 have the option to receive a complimentary version of Stata 17 SE through the CSHPE. Students will receive detailed instructions on how to purchase and install the software on their personal computer prior to the first lab meeting. Note that Stata 17 is the current version, but different versions have different capabilities (see https://www.stata.com/products/which-stata-is-right-for-me/ for details). If you would prefer to make other arrangements or purchase a different version, the University of Michigan has a campus-wide agreement with Stata allowing students to obtain a copy of Stata, either with an annual or perpetual license. These prices are discounted from the already-discounted academic price of the software. See current pricing and availability at https://www.stata.com/order/new/edu/profplus/student-pricing/. Stata's phone number is: 1-800-782-8272 (Monday through Friday 8:00 to 5:00 Central Time), however, ordering online is easiest. If you order online, be sure to include your **UMICH.EDU** email address to ensure you receive the best price. Once your order is processed, you will be contacted by a StataCorp salesperson about how to obtain your copy. You may pay using a credit card (Visa, MasterCard, American Express or Discover) or a faxed university purchase order.

APA STYLE:
For all assignments (including lab assignments and the final project), use APA 7 style unless you come from a discipline that uses a different citation style. If that is the case, please notify Steve and the GSI early in the semester. If you need help with APA style, please see the links below:

http://www.apastyle.org/manual/

https://owl.purdue.edu/owl/research_and_citation/apa_style/apa_formatting_and_style_guide/general_format.html

STUDENT RIGHTS AND RESPONSIBILITIES:
You should understand what your rights and responsibilities are as a student. Information on these topics can be found at: http://www.umich.edu/~oscr/. You may also want to read the University's General Catalogue, especially the section that details your rights as a student **and** the section that discusses the University's expectations of you (See http://www.rackham.umich.edu/policies).

EMAIL POLICY:
In general, the instructors will try to answer emails as quickly as possible during both the week and weekend. During the week, we will return emails within 24 hours. During the weekend, we will continue to strive for a 24-hour response time; however, it is possible that our responses may be delayed until the following Monday. If multiple students have similar questions, an announcement (via CANVAS) will be sent to the entire class.

CANVAS DISCUSSION BOARD
For assignments, final project questions, and general questions, students should use the Discussion Board feature on our Canvas page. As the semester progresses, Josh will create boards for each individual assignment, as well as an ongoing board for questions related to lecture, course content, or the final assignment. When posting, students should only ask one question per post to allow other students to find/review the information more easily.

ACCOMMODATIONS FOR SPECIAL NEEDS:
We are happy to discuss with you any accommodations for special needs that will help you succeed in this course. Please approach one of us during the first week of the semester and see https://ssd.umich.edu/ for more information about the University's services for students with disabilities.

Course Schedule

WEEK 1 (August 30 & Sept 2)

INTRODUCTIONS, EXPECTATIONS FOR THE COURSE & REVIEW OF LINEAR REGRESSION

We will discuss the syllabus, introduce data sets we'll be using during the class, and review linear regression.

REQUIRED READINGS:
Long, Chapters 1 & 2.
Long & Freese, Chapters 1 & 2.

ASSIGNMENTS: Some of the Stata commands we will use are Long and Freese's user-written commands not built into the current version of Stata. In their supplement to the text, Long and Freese tell you how to load these programs on to any computer you will use. Please familiarize yourself with this by next week. You may also want to check out Long's web site that contains information about these "SPOST" commands (see https://jslsoc.sitehost.iu.edu/spost13.htm).

Also, please familiarize yourself with the Canvas site for this class (named "EDUC 799 001 FA22"; Link: https://umich.instructure.com/courses/556809) and peruse the materials that are there as you will need to interface with this site on a regular basis. Please fill out the pre-class survey sent out by Josh if you haven't already done so.

If you are new to Stata or need a refresher, focus on the files in the Stata Materials folder—especially "Introduction to Stata with 50 Basic Commands," "Getting Started with Stata 15," "Stata Command Reference," or this URL https://www.stata.com/features/overview/quick-starts/

DUE: Nothing.

WEEK 2 (September 6 & 9)

THE LINEAR PROBABILITY MODEL & LATENT VARIABLE MODEL FOR BINARY DEPENDENT VARIABLES

REQUIRED READINGS:
Long, Chapter 3, pp. 34-50.

Dey, E. L. and Astin, A. W. (1993). Statistical alternatives for studying college student retention: A comparative analysis of logit, probit, and linear regression. *Research in Higher Education*, 34(5): 569-581. *(Available on Canvas)*

RECOMMENDED READINGS:
McGrayne, S. (2011). The theory that would not die: How Bayes' rule cracked the enigma code, hunted down russian submarines, and emerged triumphant from two centuries of controversy.
Pt. 1: Causes in the air and The man who did everything.

Silver, N. (2012). The Signal and the Noise. NY, NY: Penguin Books.
Ch. 7: Role models.
Ch. 8: Less and less wrong.
(both available on Canvas in Additional Readings/ Contexts of Statistical Practice)

DUE: Nothing.

WEEK 3 (September 13 & 16)
MAXIMUM LIKELIHOOD ESTIMATION & LOGISTIC REGRESSION

Nonlinear binary models are an appropriate alternative to the linear probability model (discussed in Week 2) when categorical outcomes are being analyzed. We discuss two of the most commonly used nonlinear binary regression techniques, logistic and probit regression.

REQUIRED READINGS:
Long, Chapter 3, pp. 50-83.
Long & Freese, Chapter 5

RECOMMENDED READING:
Eliason, S. (1993). *Maximum Likelihood Estimation: Logic and Practice.* Sage University Papers: Quantitative Applications in the Social Sciences, 07-086. Newbury Park, CA: Sage Publications.

DUE: Tuesday, September 13 (by start of lecture) - Assignment 1.

WEEK 4 (September 20 & 23)
LOGISTIC REGRESSION (cont'd)

We continue our discussion of logistic regression modeling. This week, we'll introduce predicted probabilities and marginal effects as ways to interpret your findings.

REQUIRED READINGS:
Long & Freese, Chapter 6

Pattison, E., Grodsky, E., & Muller, C. (2013). Is the sky falling? Grade inflation and the signaling power of grades. *Educational Researcher 42* (5), 259-265.

RECOMMENDED READING:
Rodriguez, A., Furquim, F. and DesJardins, S. L. (2018). Categorical and limited dependent variable modeling in higher education. In Michael Paulsen (Ed.), *Higher Education: Handbook of Theory and Research* XXXIV: 295-370.

SUPPLEMENTAL READING:
Arcidiacono, P. and Vigdor, J. L. (2004). *Does the river spill over? Estimating the economic returns to attending a racially diverse college*. National Bureau of Economic Research Working Paper located on Canvas and at: http://www.nber.org/~confer/2004/hieds04/riverspilla.pdf

DUE:
- Tuesday, September 20 (by start of lecture) - Assignment 2
- Tuesday, September 20 (by start of lecture) - 2 paragraph proposal/abstract for final paper

WEEK 5 (September 27 & 30)
PROBIT REGRESSION

REQUIRED READINGS:
Reread Long Chapter 3.

Reread Long & Freese Chapter 5

SUPPLEMENTAL READING:
Niu, S., & Tienda, M. (2013). Delayed enrollment and college plans: Is there a postponement penalty? *Journal of Higher Education 84* (1), 1-26. (available on Canvas)

DUE: Nothing. Work on literature review.

WEEK 6 (October 4 & 7)
HYPOTHESIS TESTING AND GOODNESS OF FIT STATISTICS

We review tests of hypotheses that can be used with maximum likelihood estimation techniques. Particularly important is a sound understanding of goodness of fit statistics such as the likelihood ratio test, and how to use this test to make inferences about the relative efficacy of one model vs. another.

REQUIRED READINGS:
Long, Chapter 4.
Long & Freese, Chapter 3 & 5.

SUPPLEMENTAL READINGS:
Chang, M., Cerna, O., Han, J., & Saenz, V. (2008). The contradictory roles of institutional status in retaining underrepresented minorities in biomedical and behavioral science majors. *Review of Higher Education 31* (4), 433-464.

Peng, C. J., So, T. H., Stage, F. K., and St. John, E. P. (2002). The use and interpretation of logistic regression in higher education journals: 1988-1999. *Research in Higher Education*, 43(3), pp. 259-294.

DUE: Tuesday, October 4 (by start of lecture) - Assignment 3

WEEK 7 (October 11 & 14)
OUT-OF-SAMPLE PREDICTION

We are now going to use logistic regression to predict out-of-sample. This strategy is used extensively by researchers who work in enrollment management within colleges and universities. Particularly important is a sound understanding of goodness of fit statistics such as the likelihood ratio test, and how to use this test to make inferences about the relative efficacy of one model vs. another.

REQUIRED READINGS:
DesJardins, S.L. (2002). An analytic strategy to assist institutional recruitment and marketing efforts. *Research in Higher Education*, *43*(5), pp. 531-554.

DUE:
- Tuesday, October 11 (by start of lecture) - Assignment 4
- Tuesday, October 11 (by start of lecture) - Draft of literature review for final project

WEEK 8 (October 21)
ORDERED LOGIT & ORDINAL OUTCOMES

This week, we will focus on how to statistically model outcomes that can be ordered. It is not appropriate to model ordered dependent variables as though they are interval. Ordered logit and probit regression models allow us to estimate these models more appropriately than when using standard linear regression techniques.

REQUIRED READINGS:
Long, Chapter 5.
Long & Freese Chapter 7.

SUPPLEMENTAL READINGS:
Cheng, S. and Starks, B. (2002). Racial differences in the effects of significant others on students' educational expectations. *Sociology of Education*, *75*(4), pp. 306-327.

A good treatment of logit and probit models. STATA code provided.
Borooah V. K. (2002). *Logit and Probit: Ordered and Multinomial Models*. Sage Publications, Inc.

An example of the use of ordered probit
Brasfield, D. W., Harrison, D. E., and McCoy, J. P. (1993). The impact of high school economics on the college principles of economics course. *The Journal of Economic Education*, *24*(2), pp. 99-111.

DUE: Nothing. Work on literature review and methodology.

WEEK 9 (October 25 & 28)
MULTINOMIAL LOGISTIC REGRESSION

In this section of the course, we move to discuss a statistical model specifically designed to estimate nominal outcomes, that is, dependent variables that are not ordered.

REQUIRED READINGS:
Long, Chapter 6, thru p.178.
Long & Freese Chapter 8 through p. 444.

DUE:

- Tuesday, October 25 (by start of lecture) - Revised literature review & draft of methodology
- Tuesday, October 25 (by start of lecture) - Assignment 5

WEEK 10 (November 1 & 4)
MULTINOMIAL LOGISTIC REGRESSION, cont'd

REQUIRED READINGS:
Perna, L. W. (2004). Understanding the decision to enroll in graduate school: Sex and racial/ethnic group differences. *The Journal of Higher Education*, *75*(5), pp. 487-527.

Posselt, J., et al. (2012). Access without equity: Longitudinal analyses of institutional stratification by race and ethnicity, 1982-2004. *American Educational Research Journal 49*(6), 1074-1111.

SUPPLEMENTAL READING:
Ono, H. (2001). Who goes to college: Features of institutional tracking in Japanese higher education. *American Journal of Education*, *109*(2), pp. 161-195.

DUE: Nothing – work on final papers

WEEK 11 (November 8 & 11)
INTERACTION EFFECTS

REQUIRED READINGS:
Mitchell, Chapter 18 (Nonlinear Models) from *Interpreting and Visualizing Regression Models Using Stata (on Canvas if you opted not to buy the entire book)*

Norton, E. C., Wang, H., & Ai, C. (2004). Computing interaction effects and standard errors in logit and probit models. *Stata Journal*, *4*, 154-167.

Ai, C., & Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics letters*, *80*(1), 123-129.

DUE: Tuesday, November 8 (by start of lecture) - Assignment 6

WEEK 12 (November 15 & 18)
CONDITIONAL LOGIT MODELS

In the multinomial logistic model each explanatory variable has a different effect on each outcome. The conditional logit model is a closely related technique in which the coefficients for the variable are the same for each outcome, but the values of the variables differ for each outcome.

REQUIRED READINGS:
Long, Chapter 6, pp. 178-186.
Long & Freese Chapter 8, pp. 454-464

SUPPLEMENTAL READINGS:
Jacob, B., McCall, B., & Stange, K. (2018). College as country club: Do colleges cater to students' preferences for consumption? *Journal of Labor Economics*, *36*(2), 309-348.

Boskin, M. J. (1974). A conditional logit model of occupational choice. *The Journal of Political Economy*, *82*(2), Part 1, pp. 389-398.

DUE: Tuesday, November 15 (by start of lecture) - Assignment 7
Friday, November 18 (by start of lab) - Literature review, revised methodology, draft findings

WEEK 13 (November 22)
COUNT OUTCOMES: THE POISSON REGRESSION MODEL

Sometimes we study the frequency of an event occurring, such as the number of times a student is late or absent from school. Dependent variables that contain count information are often treated as continuous and linear regression techniques are applied. However, this particular application of the linear regression technique may result in inefficient, inconsistent, and biased estimates. The Poisson regression model redresses these problems.

REQUIRED READINGS:
Long, Chapter 8.

Stage, F. (2007). Moving from probabilities to possibilities: Tasks for quantitative criticalists. *New Directions for Institutional Research 133*, 95-100.

RECOMMENDED READING:
Borsch-Supan, A. (1990) Education and its double-edged impact on mobility. *Economics of Education Review*, *9*(1): 39-53.

SUPPLEMENTAL READINGS:
Toutkoushian, R. T. (1999). Is there a day of the month effect in "Beat It On Down the Line"? in R. Weiner (ed.), *Perspectives on the Grateful Dead: Critical Writings*. Westport, CT: Greenwood Press.

DUE: Nothing – work on final papers

WEEK 14 (Nov 29 & Dec 2)
INTRODUCTION TO CAUSAL INFERENCE

When analyzing observational data, many social scientists have used categorical dependent variable regression models to make inferences about cause and effect. For instance, scholars have estimated a logistic model to analyze whether receiving various types of financial aid causes students to persist to the second year of college. Doing so, however, can be problematic when the independent variable of interest is not randomly allocated (as is the case with financial aid). This week's readings provide an overview of causal inference and an introduction to quasi-experimental methods that employ non-linear models as a first step in a two-stage process for producing estimates that are not tainted by self-selection bias.

REQUIRED READINGS:

DesJardins, S. L., & Flaster, A. (2013). Non-experimental designs and causal analyses of college access, persistence, and completion. In L. W. Perna & A. P. Jones (Eds.), *The State of College Access and Completion* (pp. 190–207). New York, NY: Routledge.

Pay close attention to the sections on the counterfactual framework (p. 195-196) and propensity score matching (p. 200-202).

RECOMMENDED READING:
Zuberi, T. (2000). Deracializing social statistics: Problems in the quantification of race. *Annals of the AAPSS, 568*, 172-185.

Reynolds, C. L., & DesJardins, S. L. (2009). The use of matching methods in higher education research: Answering whether attendance at a two-year institution results in differences in educational attainment. In J.C. Smart (Ed.), *Higher Education: Handbook of Theory and Research* (pp. 47-97). Springer Netherlands.

Bielby, R. M., House, E., Flaster, A., & DesJardins, S. L. (2013). Instrumental variables: Conceptual issues and an application considering high school course taking. In M. B. Paulsen (Ed.), *Higher Education: Handbook of Theory and Research* (pp. 263–321). Springer Netherlands.

DUE: Continue to work on your final papers.

WEEK 15 (December 6th or 13th)
IN CLASS PRESENTATIONS

DUE: Tuesday, December 6 or 13 (by start of presentations) – Final Paper