Rutgers, The State University of New Jersey  
School of Social Work  
Advanced Statistical Methods II:  
Applied Regression and Related Multivariate Methods  
19:910:639, Spring 2016

Class  
Instructor: Lenna Nepomnyaschy, Associate Professor  
Email: lennan@ssw.rutgers.edu  
Time: Mondays 2:00 – 4:00 pm  
Weekly drop-in individual help sessions: 1:20 – 2:00pm  
Location: Social Work Annex – Seminar room

Required Lab  
Instructor: Louis Donnelly, Post-Doctoral Scholar  
Email: louisjd@princeton.edu  
Time: Mondays 4:15 – 5:30  
Location: Social Work Annex – Seminar room

Course Overview  
This course, the second of the statistics sequence for social work doctoral students, will focus on applied regression analysis and related multivariate methods. Linear regression will be covered in depth, including regression assumptions, model specification, and diagnostics. Methods for dichotomous and categorical dependent variables, including logistic, probit, and multinomial regression will also be covered. Advanced methods including multi-level models, survival analysis, fixed and random effects models, and others will be introduced. Students will learn to use the Stata statistical package for all analyses and class assignments.

Required Software:  
This course requires that students learn and use the Stata Statistical Software Package for hands-on data analysis and statistics applications for class assignments.

- Stata is available to all students in any Rutgers computer lab as well as the Doctoral Student Computer lab in the SSW Annex  
- Stata is available for all employees (GRAs, staff, faculty) for free download from the OIRT software portal: https://software.rutgers.edu/  
- For remote access, Stata is available via remote browser which can be accessed from any internet (or wifi)-connected computer with an activated Rutgers NetID: http://apps.rutgers.edu  
- To make sure your NetID is activated for this application (most likely NOT), go to: https://netid.rutgers.edu/cgi/activate_netid.cgi

Students may also purchase their own version of Stata at discounted rates through the Rutgers Office of Instructional Technology: http://www.stata.com/order/new/edu/gradplans/gp-direct.html
Required Texts


Suggested Supplementary Texts
Writing about Quantitative Analysis

OTHER USEFUL STATISTICS AND DATA ANALYSIS BOOKS


MORE ADVANCED METHODS TEXTS FOR SPECIALIZED TECHNIQUES


GENERAL STATA TEXTS


Course Requirements
Students will be evaluated on the following course requirements (detailed instructions to follow).

Homework Assignments (6 total) = 50%
There will be SIX (6) homework assignments, which will be based on the skills and concepts introduced during the previous class and on the required readings. Assignments will include hands-on application of statistical formulas to basic data, writing syntax to create Stata output from a dataset, creating tables and graphs from output, interpreting output, and writing up methods and results of analyses. Assignments will build on one another leading to the mid-term and final assignment that will include most of the previous elements.
Homework #1: Selecting and describing variables – univariate analysis
Homework #2: Calculating and interpreting bivariate regression coefficients by hand
Homework #3: Estimating, interpreting, and predicting parameters for bivariate regressions
Homework #4: Estimating and interpreting results from a series of multiple regression models
Homework #5: Estimating and interpreting results from regressions with binary outcomes
Homework #6: Interaction effects (moderation) with continuous and binary dependent variables

Mid-term Assignment = 20%
The mid-term assignment will be the first part of the final assignment. Students will formulate a question, select and construct appropriate variables for inclusion in a multiple regression model with a continuous dependent variable, present basic descriptive characteristics of the variables, estimate a number of multiple regression models, and interpret their results. This assignment will build on the elements of the weekly homework assignments.

Final Written Assignment = 20%
The final assignment will consist of a complete data analysis project which will build on the mid-term assignment and all prior weekly homework assignments. Analyses will be extended to include estimation of interaction effects, and a variety of models with binary dependent variables. The written assignment will take the form of the methods and results sections of a journal-style quantitative empirical paper. Students will describe their measures and analytic strategy, describe the sample characteristics, interpret results from their bivariate and multivariate models, and provide a brief discussion of the answer to their question.

Presentation of Final Assignment = 5%
Students will present their final assignment to the class in a powerpoint presentation (10 minutes maximum, similar to a conference presentation) on the last day of class.

Peer-reviewed article submission and discussion = 5%
Each student will find a peer-reviewed scholarly article that utilizes multiple regression estimation for the class to discuss. All students are expected to read each article to discuss in a specific class. Articles must be submitted one week prior to discussion date.

*It goes without saying that students are expected to attend every class, arrive on time and participate in class discussions and exercises. While there is not specific course credit associated with attendance and participation, absences, being late to class, and lack of preparation will impact students’ overall grades.*
| Week 1: | Jan 25 | Class: Introduction to Course & Review  
|        |        | *Lab: Intro to Stata and constructing variables*  |
| Week 2: | Feb 1  | Class: Bivariate regression – Introduction  
|        |        | *Lab: Constructing and describing variables continued*  |
| Week 3: | Feb 8  | Class: Bivariate regression – Estimation and Interpretation, **HW #1 Due**  
|        |        | *Lab: Bivariate regression*  |
| Week 4: | Feb 15 | Class: Bivariate regression – Hypothesis testing, **HW #2 Due**  
|        |        | *Lab: Bivariate regression continued*  |
| Week 5: | Feb 22 | Class: Multiple regression – Introduction and model building  
|        |        | *Lab: Multiple regression continued*  |
| Week 6: | Feb 29 | Class: Multiple regression – Dummy variables & Nonlinear functions, **HW #3 Due**  
|        |        | *Lab: Multiple regression continued*  |
| Week 7: | Mar 7  | Class: Multiple regression – Diagnostics, Confounding & Mediation, **HW #4 Due**  
|        |        | *Lab: Multiple regression – diagnostics & review for mid-term assignment*  |
| **MARCH 14** | | **SPRING BREAK – NO CLASS**  |
| Week 8: | Mar 21 | Class: Binary dependent variables  
|        |        | **MID-TERM ASSIGNMENT DUE**  
|        |        | *Lab: Binary Dependent Variables*  |
| Week 9: | Mar 28 | Class: Binary dependent variables continued  
|        |        | *Lab: Binary dependent variables continued*  |
| Week 10: | April 4 | Class: Interaction (moderation) effects w/ continuous DVs, **HW #5 Due**  
|        |        | *Lab: Interaction effects w/continuous DV’s*  |
| Week 11: | April 11 | Class: Interaction effects w/ binary dependent variables, Missing data  
|        |        | *Lab: Interaction effects w/binary DVs*  |
| Week 12: | April 18 | Class: Categorical dependent variables, **HW #6 Due**  
|        |        | *Lab: Categorical DVs and review for presentations*  |
| Week 13: | April 25 | Class: Advanced methods  
|        |        | *Lab: Review for final assignment*  |
| Week 14: | May 2  | Class: Advanced methods  
|        |        | *Lab: Review for final assignment*  |
| Week 15: | May 9  | Class: **STUDENT PRESENTATIONS; Final assignment Due**  
|        |        | NO LAB  |
Detailed Course Outline

Please note: In addition to the required readings for each week, there are sample empirical articles listed. I will be adding (or substituting) relevant peer-reviewed empirical papers that use the various methods that we are covering as we go, including those submitted by students. Thus, each week there will be alternate journal articles that students will be required to read.

Week 1: January 25, 2016
Topics: Introduction to course, Concepts, Examples, Terminology, Variables

Required Readings:


Gordon, Chapter 1: Examples of Quantitative Research in the Social and Health Sciences (p. 5-7). Briefly skim One Literature excerpt to discuss in class.

Gordon, Chapter 5: Basic Descriptive Statistics (p. 97 – 135), SKIM & REVIEW: Types of Variables: Nominal, Ordinal, Interval (and the ways in which each should be described)

Supplementary Readings, Statistics Review:
Gordon, Chapter 6, Sample, Population, and Sampling Distributions (p. 143 – 188)


Other Supplementary Reading – Just for Fun

Week 2: February 1, 2016
Topics: Introduction and overview of regression analysis, Calculating bivariate regression coefficients

Required Readings:
Lewis-Beck, Chapter 1: Bivariate Regression: Fitting a Straight Line (p.1 – 22)
Gordon, Chapter 8: Basic Concepts of Bivariate Regression (p. 241 – 256). **STOP at 8.4.4**

Examine the descriptive tables (Table 1, generally) and how the authors “narratively describe” the variables in these TWO papers:

- Fomby & Cherlin. 2007. Family Instability and Child Well-being. *American Sociological Review*. **Page 191, Results, Descriptive Results and Table 1.**

- Nepomnyaschy, Magnuson & Berger. 2012. Child Support and Child Development. *Social Service Review*. **Page 16, Results, Sample Description and Table 1.**

**Week 3: February 8, 2016**

**Topics**: Bivariate regression analysis (continued): Regression Assumptions, Interpreting Coefficients

**Homework #1 Due, (February 11): Selecting and Describing variables**

**Required Readings:**
Lewis-Beck, Chapter 2: Bivariate Regression: Assumptions and Inferences, **1st PART ONLY** (p. 23 – 29).

- Studenmund, Chapter 4: The Classical Model (p. 93 – 110).


**Week 4: February 15, 2016**

**Topics**: Bivariate regression continued: Hypothesis testing; Rescaling and transforming variables

**Homework #2 Due (February 18): Calculating bivariate regression coefficients by hand**

**Required Readings:**

- Studenmund, Chapter 5: Hypothesis Testing (p. 121 – 150, 159 – 166)

- Gordon, Chapter 8: Basic Concepts of Bivariate Regression (p. 260 – 281, **STOP at 8.5**)

**Sample papers**

**Week 5: February 22, 2016**

**Topics:** Multiple regression, Model specification

**Required Readings:**
Gordon, Chapter 9: Basic Concepts of Multiple Regression (p. 294 – 330)

Lewis-Beck, Chapter 3: Multiple Regression (p. 55 – 74).

Studenmund, Chapter 6: Choosing the Independent Variables (p. 167 – 190).

**Sample papers**


**Week 6: February 29, 2016**

**Topics:** Multiple Regression (cont’d): Dummy Variables, Nonlinear Relationships

**Homework #3 Due (March 3): Estimating, interpreting, and predicting parameters from bivariate regression models**

**Required Readings**
Gordon, Chapter 10: Dummy Variables (p. 334 – 372, UP TO 10.5)

Gordon, Chapter 12: Nonlinear Relationships (p. 434 – 456)

Miller, Chapter 9: Quantitative Comparisons for Multivariate Models *(p. 193 – 199 ONLY)*

Miller, Chapter 10: The Goldilocks Problem in Multivariate Regression (p. 211 – 229)

**Sample papers**

**Suggested Readings:**
Hardy, M. 1993. *Regression with Dummy Variables*, Chapters 1, 2, and 3 (p. 1 – 28)

**Week 7: March 7, 2016**

**Topics:** Multiple Regression continued: Regression Diagnostics, Confounding, and Mediation

**Homework #4 Due (March 10): Estimating and interpreting parameters from multiple regression models**

**Required Readings:**
Gordon, Chapter 14: Outliers, Heteroskedasticity, and Multicollinearity (p. 481-520).

Lewis-Beck. Chapter 4: Multiple Regression: Special Topics (p. 75-95)

Gordon, Chapter 13: Indirect Effects and Omitted Variable Bias (p. 461 – 480)

Miller, Chapter 3: Causality, Statistical Significance and Substantive Significance (p. 34 – 48)

Studenmund, Chapter 11: Running Your Own Regression Project, Practical Advice for Your Project (p. 383 – 393).

**AND A Regression User’s Checklist and Guide, (p. 395 – 400).**


**Sample papers**

**MARCH 14, 2016 SPRING BREAK NO CLASS**

**Weeks 8 & 9: March 21 & March 28, 2016**

**Topics:** TWO weeks: Dichotomous dependent variables: Odds ratios, Risk Ratios, Logistic Regression, Linear probability models, Probit Regression, Marginal Effects

**Week 8: MID-TERM ASSIGNMENT DUE**
**Required Readings:**
THESE READINGS WILL BE COVERED OVER THE NEXT TWO WEEKS

Gordon, Chapter 16: Dichotomous Outcomes (p. 552 – 602).


**Sample papers**


**Suggested Readings:**
Pampel, *Logistic Regression: A Primer*, Chapters 1, 2, 4 and *SKIM* 3 (p. 1 – 68)

---

**Week 10: April 4, 2016**

**Topic:** Interaction (moderation) effects w/continuous dependent variables

**Homework #5 Due:** *Estimating and Interpreting Results from Regressions with Binary Outcomes*

**Required Readings:**
Gordon, Chapter 11: Interactions (p. 381-425).

Miller, Chapter 16: Writing About Interactions (p. 339 – 365).

**Sample papers**
Greenfield, E. & Marks, N. 2010. Sense of Community as a Protective Factor against Long-Term
Psychological Effects of Childhood Violence. *Social Service Review* 84(1): 129-147. (interaction effects)


**Supplementary Reading**


---

**Week 11: April 11, 2016**

**Topic:** Interactions with Binary Dependent Variables; Missing Data

**Required Readings:**


**Sample papers**


---

**Week 12: April 18, 2016**

**Topics:** Polytomous dependent variables (categorical, ordered, count),

**Homework #6 Due:** Estimating and comparing output from different models for binary dependent variables

**Required Readings:**

Gordon, Chapter 17: Multi-Category Outcomes – Multinomial Logit & Ordered Logit, (p. 609-646 ONLY).

**Sample papers**

Bell, Janice & Frederick Zimmerman. 2010. Shortened Nighttime Sleep Duration in Early Life and Subsequent Childhood Obesity. *Archives of Pediatric and Adolescent Medicine*, 164:9(840-845). (ordered logit)

**Weeks 13 & 14: April 25 & May 2, 2016**

**THESE TOPICS WILL BE COVERED OVER THE NEXT TWO WEEKS**

**Topics:** Introduction to advanced methods: Longitudinal analyses and addressing challenges to causal inference

A. Cross-sectional data: Propensity Score Matching, Difference in Difference Models, Instrumental Variables, Multi-level Models for Grouped data

B. Longitudinal data: Survival Analysis, Fixed Effects, Lagged Dependent Variables, Cross-Lagged Models, Multi-level Models for Longitudinal data

**Required Readings:**


Chapter 2: Exploring Longitudinal Data on Change (p. 16 – 44)
Chapter 3. Introducing the Multilevel Model for Change (p. 51 - 74).


**Sample Empirical Papers**


**Supplementary Readings**


---

**Week 15: May 9, 2016**

STUDENT PRESENTATIONS OF FINAL ASSIGNMENT
WEEK 15, MAY 9, 2016: FINAL ASSIGNMENT DUE