NEW YORK UNIVERSITY
ROBERT F. WAGNER GRADUATE SCHOOL OF PUBLIC SERVICE

PADM-GP 2902: MULTIPLE REGRESSION AND INTRODUCTION TO ECONOMETRICS
Spring 2017

FACULTY INFORMATION

Professor:
Jesse Margolis
Email: jmargolis@nyu.edu
Office hours: Puck 3017, Tue., 3:00 – 4:30pm and by appointment

Teaching Colleagues:
Amy Ganz (alg638@nyu.edu)
Madeline Stoddart (mjs1026@nyu.edu)

COURSE DESCRIPTION

Multiple regression is the core econometric technique used by policy and financial analysts. In this course, you will learn how to use and interpret this critical statistical method. Specifically, you learn how to build and estimate multiple regression models, how to evaluate whether regression coefficients are biased, whether standard errors (and thus t statistics) are valid, and whether regressions used in policy and finance studies support causal arguments.

In addition, employing one consistent dataset for all your computer exercises, you perform statistical analyses discussed in class using Stata, an econometric statistical package, and you see how the results reflect econometric concepts. Finally, with a group of your classmates and project datasets provided by your professor, you do a project that involves estimating your own regression model and applying the techniques we learn in class.

By the end of the course, you should be able to:

- Understand what an Ordinary Least Squares (OLS) regression does and why it is useful
- Write and interpret mathematical equations representing various regression models
- Interpret regression results as they are typically represented in statistical software packages and academic articles
- Use Stata to run your own regressions to suggest answers to interesting policy questions
- Think critically about the assumptions underlying your (or another researcher's) interpretation of regression output and test whether these assumptions are likely to hold
- Conduct a research project in which you formulate, estimate, write about, and present an econometric model
- Have a better understanding of the statistics that underlie research in your field of interest

LECTURES

Sec 1: Tue. 11:00am – 12:40pm
12 Waverly Place, Room L120

Sec 2: Tue, 4:55pm–6:35pm
Silver Room 208

STATA TUTORIALS

In addition to a video tutorial posted to *NYU Classes*, I will lead two *optional* in-person tutorials:

- Monday, 1/30, 6:30pm-8pm (12 Waverly Place, Room L111)
- Monday, 2/6, 1:45pm-3:15pm (12 Waverly Place, Room L111)

To take full advantage of the tutorial, I recommend that you bring a laptop with Stata installed.

RECITATIONS (PADM-GP 290)

You are required to register for recitation. Attendance is optional but highly recommended. In these sessions you will discuss the learning objectives of the problem set (and some specific questions) that is due the following week. As time allows, TCs will also answer follow-up questions from the week's lecture and Stata.

Section 003: Thursdays, 8:35-9:35pm, Silver, Room 208 (Stoddart)

Section 004: Fridays, 12:35-1:35pm, Meyer, Room 102 (Ganz)

The Friday recitation will meet on Tuesday evenings instead of Friday afternoon twice during the semester:

- Tuesday, 2/14, 8:35pm – 9:35pm (Meyer, Room 102)
- Tuesday, 2/28, 8:35pm – 9:35pm (Meyer, Room 102)

TUTORING

We also offer free optional tutoring from the teaching colleagues.

- Mondays, 5:30pm – 6:30pm (Stoddart) – The Study, Puck Building
- Wednesdays, 5:30pm – 6:30pm (Ganz) – The Study, Puck Building

PREREQUISITE: CORE-GP 1011 or equivalent

NYU CLASSES

You will need to have access to the NYU Classes found under "Academics" on your NYU Home site (https://home.nyu.edu/). The most up-to-date syllabus and all class related documents (problem sets, computer exercises, databases, solutions, PowerPoints, etc.) will be posted there.

SUMMARY OF COURSE GRADING
1. 20% Problem Sets and Computer Exercises
2. 35% Midterm Exam
3. 45% Regression Project

COURSE REQUIREMENTS AND GRADING

1. ***(20%) Fourteen problem sets and computer exercises. There are 100 possible points for each problem set (PS) and for each computer exercise (CE).***

    a) Thus the grades on all your problem sets and computer exercises will be added, divided by 14 (or 12 after two are dropped – see c. below) and multiplied by 0.2.

    b) Problem Sets and Computer exercises will be graded for completion of the entire exercise. You should try to obtain correct answers, however, as this will help you on the exam. **Copying others or previous semesters will result in a zero**.

    c) I will drop the two lowest grades.

    d) Answers to problem sets and computer exercises **must be submitted online through NYU Classes by midnight on Monday prior to class**. No late assignments or emailed assignments will be accepted.

    e) For Computer Exercises, submit only a write up (Word document or PDF) and your log file for the final "run" of the analysis. You do not need to submit your "do" file. For problem sets, submit your answers on the answer sheet provided for each set.

    f) Solutions will be posted to *NYU Classes* on Tuesday afternoons.

2. ***(35%) Exam (100 possible points)***

An in-class exam will be given on **Tuesday, April 11** (see Course Schedule below). You may bring a non-graphing calculator and two pages (single-sided) of notes.

3. ***(45%) Group Project (100 possible points)***

In groups of four to five, you will conduct a regression analysis, present your results, and write a paper. (**Note:** All group members will complete peer evaluations that will factor into grades.)

    a) Please fill out this Project Data Set Preference Form (https://goo.gl/forms/820NALQrr2KsqCri1) by **Monday, January 30, at 9am** and I will put you in groups. See datasets and their descriptions in *NYU Classes, Resources, Project Descriptions and Datasets.*

b) Once groups have been assigned, I will send out a sign-up sheet to meet with me during the week of **Class 4 or 5** to discuss the project. Please present the question you propose to answer and at least one specification that will be estimated.

c) Read chapter 11, "Running Your Own Regression Project," in the course text.

d) See me during office hours or after class for quick check-ins on your work. You may also schedule a second meeting during the week of the exam or the week following to share a draft of your presentation and receive feedback.

e) Present your results during one of the last two class sessions to get feedback before you finish writing. **All students are expected to attend both days of presentations.**

f) Write an 8-10 page paper, double spaced, organized into five sections as shown in the *Paper Outline* below. The paper should include at least two tables and may include one or more figures (i.e. graphs). Tables and figures will appear at the end of the paper and do not count toward the page limit.

Paper Outline

I. Introduction: What is the goal of your regression study? Why is it interesting? Why do we care? (This does not have to be momentous – but you should explain why the results could be interesting or valuable to someone.)

II. Data: Describe your sources and discuss the descriptive statistics in Table 1.

III. Model and Empirical Strategy: What is your model (equation), and how does it achieve the goal of your analysis? Why are the specific variables used and measured as they are? Do you have any prior expectations about the signs of coefficients? How will you estimate this model? (Usually OLS with fixed effects.)

IV. Results: Discuss the results that are presented in Table 2.

V. Conclusions: What does your model say about your goal or issue? What is the next step in this research?

Appended at end of paper:

i) Figures (with title): Histograms, bar charts, scatterplots, or other figures you choose to include.

ii) Table 1 (with title): Descriptive statistics of all the variables in your model(s).

iii) Table 2 (with title): Results of your models, presented in four or five columns.

iv) Final cleaned up, annotated Stata log file of your results.

Note: questions about the group project should be addressed to the professor, not the TCs.

COURSE MATERIALS

1. **Required Text: A. H. Studenmund, *Using Econometrics: A Practical Guide***, 6th ed, cited as S. ISBN: 0131367730 (note: there is a newly-released 7th edition, which should be fine to use, but the syllabus will refer to chapters and page numbers as they appear in the 6th edition)

2. **Required Software: Stata/IC 14**, purchased and loaded onto your computer by week one.

You should purchase this software here in order to obtain a student discount: http://www.stata.com/order/new/edu/gradplans/student-pricing/

Purchase Stata/IC 14 (**not** Small Stata). The least-cost option is a 6-month license, at $75. If you are planning to take Estimating Impacts, Advanced Empirical Methods, or the Research Capstone, you may want to consider a one-year or perpetual license. Stata 14 is not platform-dependent and will run on either Windows or Mac operating systems.

No previous knowledge of Stata is necessary. In addition to learning Stata through the problem sets and computer exercises and in class, the Data Services Studio in Bobst (http://library.nyu.edu/dataservice/) offers short courses (tutorials) and on-site help.

Stata is also available free-of-charge at NYU's Virtual Computer Lab (VCL). For more information, see: http://www.nyu.edu/life/resources-and-services/information-technology/instructional-technology-support/instructional-technology-tools-and-services/virtual-computer-lab.html

If you choose to use Stata on the VCL, please note that:

- Specifying a file path is slightly different (there is an additional prefix: \\Client\$...).
- You must not save your work on the VCL but on your actual computer. Otherwise, you will lose your work once you log out or get disconnected for any reason.

3. Required: Materials from NYU Classes. Throughout the semester, lecture notes, PowerPoint slides, problem sets, computer exercises, and solutions will be posted on NYU Classes. There will also be one dataset – newschools9810.dta – posted on NYU Classes that we will be using throughout the semester. Please make sure you have access to NYU Classes. See the course schedule below for when assignments are due in class.

There is also a brief video on using Stata created by Professor Dan Smith, who taught this class in the past. Once you have downloaded Stata, you should review this video during the first week of class.

COURSE SCHEDULE

Date	Lecture (details next page)	Reading (Studenmund text)	Due at midnight before class (PS = Problem Set; CE = Computer Exercise)	Covered in Recitation
Tue, 1/24	Class 1: Introduction & Bivariate Regression	Ch 1 & Ch. 2 (34-39 & 46-57)	See "Class Topics," next page	PS 1
Tue, 1/31	Class 2: Multivariate Regression Mechanics	Ch. 2 (39-45) & Ch. 3	PS 1	PS 2
Tue, 2/7	Class 3: Multivariate Regression Assumptions	Ch. 4	PS 2	Stata
Tue, 2/14	Class 4: Hypothesis Testing	Ch. 5 (not including appendix)	CE 1	PS 3
Tue, 2/21	Class 5: Function Form I	Ch. 6 & Ch. 7 (207-213, 218-220, 223-226); Appendix Ch. 5	PS 3	PS 4
Tue, 2/28	Class 6: Functional Form II	Ch. 7 (213-218, 220-223, 226-232)	PS 4 CE 2	PS 5
Tue, 3/7	Class 7: Multicollinearity & Autocorrelation	Ch 8 & Ch. 9	PS 5 CE 3	PS 6
Tue, 3/14	SPRING BREAK			
Tue, 3/21	Class 8: Heteroskedasticity	Ch. 10	PS 6 CE 4	PS 7
Tue, 3/28	Class 9: Panel Data	Ch. 16	PS 7 CE 5	PS 8
Tue, 4/4	Class 10: Binary Dependent Variables	Ch. 13	PS 8 CE 6	Exam Review
Tue, 4/11	Exam			None
Tue, 4/18	Class 11: Quasi-experimental methods	See NYU Classes		None
Tue, 4/25	Student Presentations			None
Tue, 5/2	Student Presentations			None
Fri, 5/5	Final papers due to Professor's mailbox (Puck, 3rd Floor) by 5 p.m.			

CLASS TOPICS

Class 1: **Introduction & Bivariate Regression**
Population vs sample models; OLS estimator; interpreting coefficients; correlation coefficient; R-squared

Before Class 1:
- Fill out the Student Information Sheet (https://goo.gl/forms/jzRJzFXAvNPOIo5D2)
- Purchase and install Stata/IC 14;
- Download and save, from NYU Classes: *newschools9810.dta*,
- Watch the video on NYU Classes about using Stata.

Class 2: **Multivariate Regression Mechanics**
Omitted variable bias; OLS estimator for multiple regression; multiple regression as a prediction model

Class 3: **Multivariate Regression Assumptions**
Classical Liner Model (CLM) Assumptions; Sampling distribution of $\hat{\beta}$

Class 4: **Hypothesis testing**
Standard error of $\hat{\beta}$; significance tests; confidence intervals; F test; R^2; Adjusted R^2

Class 5: **Functional Form Part I: Polynomials and Dummy Variables**
Polynomials; dummy variables; F-Tests of joint significance

Class 6: **Functional Form Part II: Interactions and Logarithmic Transformations**
Interactions of dummies; interactions of continuous and dummy variables; continuous by continuous interactions; logarithmic transformations (percent change or elasticity transformations)

Class 7: **Multicollinearity and Autocorrelation**

Class 8: **Heteroskedasticity**

Class 9: **Panel Data Estimation Part I**
Pooled cross-sectional and time series data; panel data; fixed effects estimation.

Class 10: **Panel Data Estimation Part II; Binary Dependent Variables**
Introduction to Linear Probability, Probit, and Logit Models

Class 11: **A Very Brief Introduction to Experimental & Quasi-Experimental Designs**
Randomized Controlled Trial (RCT); Instrumental Variables (IV), Regression Discontinuity (RD); Matching