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

PADM-GP 2902:
MULTIPLE REGRESSION AND INTRODUCTION TO ECONOMETRICS
Spring 2016

FACULTY INFORMATION

Professor:
Timothy Roeper
Email: tar310@nyu.edu
(Please include "PADM-GP 2902" in the subject line of all emails)
Office hours: 7-8 PM on Tuesdays and 11 AM-12 PM on Thursdays (Room 3045)

Teaching Colleagues:

Andrew Brown	Jennifer Hur	Carlos Quirola
ab1958@nyu.edu	jh173@nyu.edu	carlos.quirola@nyu.edu

COURSE DESCRIPTION
Multiple regression is the core statistical technique used by policy and finance analysts in their work. In this course, you learn the theory and practice of econometric analysis. Specifically, you learn 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 professors, 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

Timothy Roeper   Spring 2016   PADM-GP 2902

LECTURES

Sec 1: Thur. 9:00am–10:40am   Sec 2: Tue, 4:55pm–6:35pm
7E12 Room LL23   194M Room 203

Preparation for Lecture

The slides with the next week's lecture will be available by Friday afternoon each week. You should print out the slides before class with the goal of being able to answer all of the questions in blue. You should also read the indicated chapters in the Studenmund text, see the class schedule below.

ECONOMETRICS RECITATIONS (PADM-GP 2902)

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 001: Thursdays, 8:35-9:35pm, SILV, Room 208 (Hur)
Section 002: Fridays, 12:35-1:35pm, TISC, Room LC11 (Quirola)

STATA RECITATION (PADM-GP 2902)

All of you are expected to learn and use Stata in this course, but it will not be covered in lecture. The best way to learn Stata is to try doing it yourself and get help when you get stuck. The computer exercises will give you all an opportunity to use Stata, and when you have problems you should come to the Stata recitation on Monday night. You should also use the Monday night recitation as a resource when you are using Stata for your final project.

Stata Recitation: Mondays, 8:30-9:30 PM, Waverly, Room 668 (Brown)

TUTORING

We also offer free tutoring from the teaching colleagues.
   - Mondays, 5pm – 6pm (Quirola) – Jersey Conference Room (**except for 2/29 and 4/11, on those days the Mulberry Conference Room**), Puck Building
   - Wednesdays, 6:30pm – 7:30pm (Brown) – The Study, Puck Building
   - Thursdays, 5:30pm – 6:30pm (Hur) – 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, class notes, 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 midterm. **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 on NYU Classes before the beginning of the class for which they are listed in the course schedule below**. No late assignments will be accepted.

   e) For Stata output, submit only the last "run" of the analysis. For problem sets, submit your answers on the answer sheet provided for each set. Write your mailbox number on your submissions. Graded assignments will be available on NYU Classes.

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

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

An in-class exam will be given on **Tuesday, 11/17 & Wednesday, 11/18** (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 five to six, 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. You are also strongly encouraged to individually complete an optional assignment that will get you started with your analysis in Stata.)

   a) Complete your project form (at this link) by Saturday, January 30th, at 6:00pm with your rankings of preferred data sets, and I will put you in to groups. If you do not fill out the project form, I will simply assign you to a group. See datasets and their

descriptions in *NYU Classes, Resources, Project Descriptions and Datasets.*

b) As a group, contact me to set up a time to meet to discuss the project in week four. The day before our meeting, email me a project proposal that is no longer than 1 page double-spaced (it is perfectly fine if it is shorter). 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.

e) Go to the Stata recitation (Mondays at 8:30 PM), and get started analyzing your dataset by running the regression you proposed in our meeting, as well as two alternative specifications, and exporting your results to Excel using the outreg2 or estout commands. If you cannot go to the Stata recitation and need help, go to Andrew Brown's tutoring hours (Wednesdays 6:30 PM or one of the other Teaching Colleagues' tutoring hours). Everyone in the group should run your regressions separately and then compare results. This will ensure that you all get practice using Stata and that your regressions are run correctly.

f) Present your results during one of the last two class sessions to get feedback before writing. **Note:** All students are expected to attend each day of presentations.

g) Write an 8-10 page paper, double-spaced, organized into five sections as shown in the *Paper Outline* below. The table 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: Non-Stata related questions about the group project should be addressed to the professor, not the TCs. The TCs are an excellent resource for any Stata-related questions that may arise while doing your project.

COURSE MATERIALS

1. **Required: A. H. Studenmund,** *Using Econometrics: A Practical Guide*, 6th ed, cited as S. ISBN: 0131367730.

2. **Optional: Baum, C. H.** *An Introduction to Modern Econometrics Using Stata*, ISBN: 9781597180139

3. **Required: 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 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.

My personal recommendation is not to use the VCL, but to buy Stata for your personal computer.

4. Required: Materials from NYU Classes. Throughout the semester, 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 Smith, who has taught this class in the past. Once you have downloaded Stata, you should review this video prior to the first class.

COURSE SCHEDULE (TUESDAY AND THURSDAY SECTIONS)

Date	Lecture (details next page)	Reading (S= Studenmund text)	Due in Class (PS = Problem Set; CE = Computer Exercise)	Covered in Recitation
Tue, 1/26* Thu, 1/28*	Class 1: Intro and Bivariate Regression	S Chs. 1 & 2 (34-38 & 46-57)	See "Class Topics," next page	Stata
Tue, 2/2 Thu, 2/4	Class 2: Multivariate Regression	S Chs. 2 (39-45) 6 (167-175)		PS 1
Tue, 2/9^ Thu, 2/11^	Class 3: Assumptions of Multivariate Regression	S Ch. 4	PS 1	PS 2
Tue, 2/16 Thu, 2/18	Class 4: Hypothesis Testing	S Chs. 3 & 5 (not including appendix)	PS 2 CE 2	PS 3
Tue, 2/23 Thu, 2/25	Class 5: Functional Form I	S Chs. 6 & 7 (207-213, 218-220, 223-226); Appendix Ch. 5	PS 3	PS 4
Tue, 3/2 Thu, 3/4	Class 6: Function Form II	S Ch. 7 (213-218, 220-223, 226-232)	PS 4 CE 4	PS 5
Tue, 3/8 Thu, 3/10	Class 7: Multicollinearity, Autocorrelation, Heteroskedasticity	S Chs. 8-10	PS 5 CE 5	PS 6
Tue, 3/15 Thu, 3/17	Spring Break			
Tue, 3/22 Thu, 3/24	Class 8: Panel Data	S Ch. 16	PS 6 CE 6	PS 7
Tue, 3/29 Thu, 3/31	Class 9: Binary Dependent Variables I	S Ch. 13	PS 7 CE 7	PS 8
Tue, 4/5 Thu, 4/7	Class 10: Binary Dependent Variables II; Exam Review		PS 8 CE 8	Exam Review
Tue, 4/12 Thu, 4/14	Exam			None
Tue, 4/19 Thu, 4/21	Class 11: Quasi-experimental methods	See NYU Classes		None
Tue, 4/26 Thu, 4/28	Student Presentations			None
Tue, 5/3 Thu, 5/5	Student Presentations			None
Thu, 5/12	Final papers due to Professor's mailbox (Puck, 3rd Floor) by 5 p.m.			

* Fill out the project form as PDF (at the end of syllabus) by Saturday, January 30, at 6:00pm.
^ schedule meeting with me week of class 4 to discuss regression analysis and one equation.

CLASS TOPICS

Class 1:   OLS Bivariate Regression Model with Error Term
Theoretical regression line; deterministic versus stochastic relationships;
population versus sample regression line; error and residual; OLS estimators

Before the first day of class:
- Complete the Student Information Sheet at this link.
- Purchase and install Stata/IC 14 (not small Stata);
- Download and save, from NYU Classes:
  newschools9810.dta,
- Watch the video on NYU Classes about using Stata.

Class 2:   OLS Multiple Regression and Omitted Variable Bias
Interpretation of coefficients; omitted variables; reducing bias;

Class 3:   OLS Multiple Regression and Assumptions about Error Term
BLUE assumptions

Class 4:   Hypothesis Testing in Multiple Regression Context
Significance tests; confidence intervals; F test; R^2; Adjusted R^2; interpretation of
computer output

Class 5:   Functional Form Part I: Polynomials and Dummy Variables
Functional form; using qualitative data (dummies); joint tests of significance;
curvilinear relationships

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, Autocorrelation, and Heteroskedasticity

Class 8:   Panel Data Estimation Part I
Pooled cross-sectional and time series data; panel data; fixed effects estimation.

Class 9:   Binary Dependent Variables
Introduction to Linear Probability, Probit, and Logit Models, Odds Ratios

Class 10:   A Very Brief Introduction to Experimental & Quasi-Experimental Designs
Randomized Controlled Trial (RCT); Instrumental Variables (IV), Regression
Discontinuity (RD); Matching