PUAD 602: Quantitative Methods for Policy Analysis 2
Spring 2014

Course Information:
Thursday 5:30-8:00 PM (Ward 105)
Instructor: Daniel Puskin
Prerequisites: PUAD 601 or equivalent

Instructor Information:
Office Hours: Monday 8:00-8:45, Tuesday 4:30-6:15, Wednesday 4:30-6:15, and by appointment.
Office: Ward 329
Email: puskin@american.edu
Phone: 202-885-6210

Learning Objectives:
* Continue to change the way you think about data
* Strengthen your understanding of bivariate Ordinary Least Squares regression
* Gain recognition of when classical assumptions may be violated
* Learn tools to conduct hypothesis tests for relationships between multiple variables
* Learn tools to conduct hypothesis tests for nonlinear relationships between variables
* Learn tools to address selection biases
* Be able to communicate statistical findings to a wide range of audiences
* Address limits of statistical tests and the implications for policy analysis

Course Description and Roadmap:
You finished 601 by describing the linear relationship between two variables. The types of questions you explored included the strength of the linear relationship, how the y-variable changes when the x variable changes by one unit, and whether the relationship between the x-variable and y-variable is significantly different from what random chance would predict.

In this course, we will revisit those questions and then delve much deeper. In 601, we made very strong assumptions about the relationships between our variables of interest. This course is largely about relaxing these Classical Assumptions.

When looking at just two variables, you easily can be missing so much of the story because other factors are not being incorporated into the analysis. For example, during the first lecture, we will try to explain why some homes cost more than others. Size of the house matters and the regression we run will highlight that relationship. However, the location of the house, newness of the kitchen, number of bathrooms, size of the yard, whether there is a pool, and whether the house is on a busy street will also matter. Multivariate regression allows for other factors like these to be included. Thus, if we want

a more accurate prediction of how much a house should cost based on market forces, we can come up with a model that does considerably better than just size alone would generate. These additional factors allow more variation in our outcome of interest to be explained.

Not only does controlling for other variables help us make better predictions of what our outcome variable will be, but it helps us better understand what the impact of a change in an explanatory variable will have on our outcome variable. In the bivariate case, the impact of square footage on its own may be overstated. Why? Because larger homes tend to be better along other dimensions, such as having larger yards and better kitchens and being in more expensive neighborhoods with better schools. If we want to know how much a change in square footage causes prices to go up, we better control for these other factors.

The relationships between our variables of interest are not always linear—i.e. a one unit change in an explanatory variable does not necessarily have a constant corresponding change in y. Sometimes we might be interested in relationships where the slope differs depending on the value of x—e.g. perhaps our relationship is exponential or logarithmic and not linear. Sometimes, different groups might have different slopes. For example, perhaps men and women, on average, have different returns to schooling. Sometimes our outcome of interest can only be a zero or a one. During this term, you will learn how our empirical tests can account for these possibilities.

Under the classical assumptions, when using OLS, we make very strong assumptions about the population residuals (the part of the variation in our outcome variable that cannot be explained by the explanatory variables). We will review these assumptions and examine what can be done when they do not hold. Serial correlation (when one observation's error term depends on another observation's error term) and heteroskedasticity (the variance of the error terms depends on the x-values) are two concerns that we will address.

A major assumption made in the Classical model is that there is nothing in the residual component that influences the x-values or vice versa. Multivariate regression addresses part of this potential problem. However, we cannot always simply throw more variables at the empirical question. Sometimes we need other techniques to address these selection concerns. We will use panel data, where we have multiple data points, for each individual. We will look at methods that take advantage of natural variation in our data. Difference-in-differences and instrumental variables are techniques that allow us to address concerns about selection bias.

Throughout the term, we will discuss the strength of the relationship between variables and how to interpret our coefficient of interest. For example, how does controlling for other variables change that interpretation and impact the size, magnitude and

significance of our estimate? Being able to provide a description of what our Stata results tell us is a critical skill and is dependent upon what test we run.

This is the next step in your progression to becoming great story tellers (non-fiction, fiction, and something in between). How do we take a bunch of numbers and tell a story about them? Throughout the course, we will be challenging what we observe. Statistical relationships by no means imply causality, but rather document a mathematical relationship between variables. You will now have more tools at your disposal to suggest causality. Nonetheless, even with these additional tools, you can tell a lot of different stories with numbers. Think critically about what you see, using your institutional knowledge and analytical skills. Ideally, we can get closer to establishing causal channels. When you have the true story, you will be equipped as policy experts to take proper action. Importantly, you will be able to challenge those with misguided or disingenuous interpretations of data.

Blackboard:
Please check Blackboard regularly. I will post assignments, course documents, and supplemental readings. I will also use Blackboard to send out e-mails. Please have an ability to access your AU account.

Readings:
Two texts are required for the course.

- Studenmund, A.H. *Using Econometrics*. Addison Wesley. Sixth edition, 2009 (ISBN: 9780131367739)

- Wheelan, Charles. *Naked Statistics: Stripping the Dread from the Data*. WW Norton, 2013. (ISBN 9780393071955)

Additional required readings and optional readings will be provided on Blackboard. If you would like other readings to help understand the material or to dig deeper into a subject matter, please let me know.

An optional text of particular interest, particularly for those of you interested in taking more statistics courses is

- Wooldridge, Jeffrey. Introductory Econometrics: A Modern Approach. Fifth Edition, Cengage Learning, 2012 (**ISBN-13:** 978-1111531041)

Calculators and Stata:
A basic calculator may be useful for certain exercises and exams. You may prefer a graphing calculator that allows you to see more of a trail of your calculations. We will be getting our hands dirty with data. This will involve labs and homework assignments where you will be asked to use Stata to work with data and interpret output. Stata is available in the School of Public Affairs Lab in the Ward Circle Building and the Social Science Research Lab in the Hurst Building. You may also purchase your own copy of Stata at a discounted price.

Grades:
Problem Sets (15 percent).
Most weeks will have problem sets for this course. These will be distributed a week before they are due. You may discuss assignments with classmates, but must hand in your own original assignment. Problem Sets will be due at the beginning of class or at some designated time. Late submissions will not be accepted. However, your lowest grade will be dropped. I accept problem sets submitted by email, so long as they are in a form that is easy to open.

Each problem set will be evaluated on a check plus (2 out of 2 points), check (1.8/2), check minus (1.6/2), minus (1/2) scale. A check plus demonstrates strong command over almost all the material. A check grade demonstrates command over some of the material, but contains conceptual errors. A check minus contains major deficiencies. A minus demonstrates inadequate effort. Show your work, write legibly (or type it out), and organize your thoughts—provide a professional work product.

Class Participation (5 percent) Show up and be engaged

Video Presentation (10 percent). Each student will submit a presentation on an article. Details will be given early in the semester.

Exams (45 percent). The class will have two exams – a midterm (20 percent), and a Final (25 percent).

Policy Memo (25 percent). There are two separate components of the policy memo: a proposal for a research agenda, relevant descriptive statistics and bivariate regressions of interest and plans for additional analysis 10 percent of the class grad; an extension of the first paper running multivariate regressions and robustness checks 15 percent of the class grade. The policy memo will ask you to pose a research question and answer it using appropriate methods. I will provide a few datasets for you to select from. More information forthcoming including details about the prizes for the best papers. Each stage will be reduced by 1/3 of a letter grade for each day it is late.

Tutors:
These are PhD students and are an excellent resource to help you in this class. They also have larger insights into the program and the policy world. They will be available in Ward 307 at the hours listed below.

Chris Birdsall: Mondays 2:00-7:00pm   cb0491a@student.american.edu

Ray Zuniga: Tuesdays 3:00-8:00pm   rz0014a@student.american.edu

Liz Crowe: Wednesdays 1:30-6:30pm   ec2856a@student.american.edu

Katie Vinopal: Thursdays: 1:00-6:00pm   ky1065a@student.american.edu

Support Services:
If you experience difficulty in this course for any reason, please do not hesitate to consult with me or the tutor. In addition, American offers a wide range of support services.

Academic Support Center: (x3360, MGC 243) offers study skills workshops, individual instruction, tutor referrals, and services for all students with learning disabilities. Writing assistance is also available.

Counseling Center: (x3500, MGC 214) offers counseling and consultations regarding personal concerns, self-help information, and connections to off-campus mental health resources.

Disability Support Services: (x3315, MGC 206) offers technical and practical support and assistance with accommodations for students with physical, medical, or psychological disabilities.

Writing Center: (x2991, Battelle 228) offers friendly, free, and confidential individual coaching for all types of writing.

Campus Life: check their website (http: //www.american.edu/life) for a listing of information that may be helpful to you throughout the semester.

If you have a disability and might require accommodations in this course, please notify me with a letter from Disability Support Services or the Academic Support Center early in the semester so that we can make arrangements to address your needs.

Academic Honesty:
As the Code states, "Academic Integrity is the heart of intellectual life." Respect it! You must adhere to the rules of the University's Academic Integrity Code (http://www.american.edu/academics/integrity/). You may work on homework assignments with others, but must submit an assignment in your own words. You may not copy information from a

book, article, newspaper, website, other student papers, or other sources without using quotations and citations. If you paraphrase someone else's work, cite it. If you have any questions about the Code, please ask me. Potential violations will be treated very seriously. Again from the Code: "American University views academic integrity as integral to its mission; treating it as far more serious than a disciplinary matter."

Course Schedule

1/16	Week 1	Course introduction	Studenmund Chapter 1
1/23	Week 2	Review of OLS and Notation	Studenmund Chapters 2,3
1/30	Week 3	Tests of Significance	Studenmund Chapter 5 Supplement from Wooldridge
2/6	Week 4	Classical Assumptions	Studenmund Chapter 4 Wooldridge for Proofs
2/13	Week 5	Omitted Variable Bias	Studenmund Chapter 6
2/20	Week 6	Non-Linearities	Studenmund Chapter 7
2/27	Week 7	Dummies	Studenmund Chapter 7
3/6	Week 8	Multicollinearity	Studenmund Chapter 8
3/13		No Classes Spring Break	
3/20	Week 9	Probability Modeling	Studenmund Chapter 13
3/27	Week 10	Review for Midterm	Memo Proposal Due
4/3	Week 11	**Midterm**	
4/10	Week 12	Serial Correlation and Heteroskedasticity	Chapters 9, 10
4/17	Week 13	Addressing Selection Bias (Panel Data)	Chapter 16
4/24	Week 14	Addressing Selection Bias	To Be Determined
5/1	Week 15	Final Exam	