

**Boston College**  
Department of Economics

Introductory Econometrics  
EC228  
Autumn 2009

Mark Kazarosian  
Office Campanella 480 B  
e-mail: [kazarosi@bc.edu](mailto:kazarosi@bc.edu)

Office Hours: Tues, Thurs. 1:30-2:30  
Also by appointment

TA: Shannon Phillips Office Hours:

Monday 3:00-4:00 PM (Carney 302), Wednesday 9:00-10:00 AM (Carney 306), Friday 12:00-1:00 PM (Carney 103)

### **Econometric Methods**

This is a course in the theory and application of regression analysis--a sequel to your statistics course. As in statistics, we will concentrate on problems of estimation and hypothesis testing. Unlike statistics, the primary focus is not on means of distributions, but rather on something much more interesting--relationships between variables. Most social science is about relationships between variables, and we will develop some statistical techniques to estimate the direction and magnitude of these relationships, and test hypothesis about them.

We will first address these topics with a number of simplifying assumptions, and discuss some very nice statistical properties of the estimators we develop--unbiasedness and efficiency (minimum variance). We will then ask what happens to these properties when the simplifying assumptions do not hold, and ask whether these properties can be revived by adjustments to the estimating technique.

My approach to teaching undergraduate econometrics is based on learning by doing. The text reflects my belief that both an intuitive understanding of the theory, and hands on training are important to develop an understanding and appreciation of econometrics as a powerful, marketable, and fun tool to use. You will learn how an uncountable number of interesting (and even uninteresting) questions can be explored with the use of econometrics.

Learning econometrics requires developing and implementing many skills simultaneously. These skills are math, statistics, economic theory, computer, writing, presentation and common sense. Please be prepared for an exhilarating and rigorous experience. We will likely cover the portions of the text outlined below. Those of you who feel a bit rusty on your algebra and basic differential calculus should brush up starting now. I have put a few books on reserve for your math reference.

Regression analysis (econometrics) is done by computer programs. You will use STATA for the research paper and some of the problem sets. STATA for both PCs and real computers (Apples) are available on the BC computer system, and for purchase through the BC Economics Department. The link for Stata Grad Plan orders is <http://www.stata.com/order/new/edu/gradplans/gp-campus.html>

**Software:** The homework assignments will require you to become familiar with Stata, a general-purpose statistical package in wide use across social science disciplines. A short tutorial on the use of Stata will be provided in class by Ms. Phillips. Stata has the same look and feel on all platforms on which it runs: Mac OS X, Windows, Linux, and Unix. There are extensive web-based tutorials on the use of Stata for regression analysis accessible via the course home page. There is also extensive on-line help within the program, and links from Stata's "findit" command to Internet-accessible resources as well. You may submit any questions on Stata use to me via email, which I read and answer seven days a week.

You may access Stata, on or off campus, from your own computer using the BC Applications Server (<http://apps.bc.edu>). Off-campus use may require the use of VPN on your machine (see the Help Center). An adequate number of copies of Stata for both Windows and Mac OS X platforms have been installed in the CTRC in O'Neill. If you would like to acquire a copy of Small Stata (the student version) for Windows, Mac OS X or Linux for your own use, you may order it through the Stata GradPlan (<http://www.stata.com/order/new/edu/gradplans/gp-campus.html>). Materials ordered by phone or fax through the GradPlan are ready for pickup in 1-2 business days from Mrs. Tubman in the Economics office, LCOB 412B.

The student version, Small Stata, is limited in the number of observations and variables which may be analyzed, but otherwise is a full version of Stata. Stata/IC is more appropriate if you are thinking of using the software for a senior thesis project. You will need access to the Internet to work with the datasets that will be analyzed in homework assignments.

## **Text:**

### **Required:**

*Introductory Econometrics* by Jeffrey Wooldridge, 4<sup>th</sup> ed., South-Western Publishing Company. The text is available at the Boston College Bookstore.

Recommended text: C.F. Baum, *An Introduction to Modern Econometrics Using Stata*, Stata Press (<http://stata-press.com/books/imeus.html>), 2006.1 On reserve at O'Neill Library.

### **Supplementary Readings:**

Any major newspaper. Suggestions: Boston Globe, New York Times, Wall Street Journal. There will also be several fascinating econometric articles assigned during the semester.

**Course Requirements:** The requirements for the course and grade weights are:

Grading will be based upon:

Two Midterms (20% each): 1) Thursday, October 1. 2) Thursday, November 12.  
Several Problem Sets, Quizzes, and Class Participation (10%).  
Final Exam (30%). Exam Week.  
Research paper, due Thursday December 10<sup>th</sup>. 20% (Details below.)

Presentation of Research paper: TBA

If school happens to be canceled (e.g. snow day) on the day of a scheduled exam, the exam will be given the next class meeting. There are no make up exams. If there is an excused absence from an exam, i.e. documentation from the dean, the following or previous exam will be weighted more heavily.

### **Academic Integrity:**

"Academic integrity is central to the mission of higher education. Please observe the highest standards of academic integrity in this course. Please review the standards and procedures that are published in the university catalog and on the web, at: <http://www.bc.edu/offices/stserv/academic/resources/policy/#integrity> Make sure that the work you submit is in accordance with university policies. If you have any questions, please consult with me. Violations will be reported to the Deans' Office and reviewed by the College's Committee on Academic Integrity. This could result in failure in the course or even more severe sanctions."

I expect all students to do only their own work on exams and quizzes, and to make a serious individual effort on the problem sets and research paper. Any breach of academic integrity (e.g. cheating on a quiz or plagiarism) will result in a failure for the *entire class*. We will discuss collaboration on the problem sets and the paper in class.

## **Research Paper:**

One very important component of this course is a significant research paper, done by groups of three students. This will be a practical application of econometrics to test some hypothesis that your group finds interesting. The topic is up to the group and can come from economics or from any other discipline.

## **Timetable:**

Students usually find this research paper to be one of the most rewarding and useful parts of the course. Students also find it one of the most time consuming. Below is a timetable designed to spread this work over the semester. A draft of the paper is due *before* the end of the semester (Thursday, April 11) allowing some time for some quick extensions and revisions.

Tuesday September 8 <sup>th</sup>	Opening Day
-2 weeks-	
Tuesday September 22:	Submit a brief description (abstract) of your group's topic--Carefully thought out. (<=1 page)
-3 weeks-	
Tuesday October 13 <sup>th</sup>	Submit the first draft. <i>Substantial</i> progress on every section of the paper. Includes <i>basic</i> data analysis.
-3.5 weeks-	
Thursday November 5 <sup>th</sup>	Submit the second draft. <i>Significant</i> improvement from <u>first draft</u> . Includes <i>advanced</i> data analysis.
-5 weeks-	
Thursday December 10 <sup>th</sup>	Final paper due.

## **Here is a possible outline for the research paper:**

### I. Introduction

- What are the hypothesis being tested?
- Why are they interesting?
- What does theory (economic? other?) predict about the relationships being tested?
- What are your methods? (Briefly)
- What are your results? (Briefly)

### II. Brief Literature Review

- What other work has been done on these issues?
- What has been found?
- How does your paper differ from the others?

### III. Data

- What data would you like to have?
- What data have you been able to obtain?
- What special data problems did you encounter? (Appendix only)

### IV. Empirical Work

- Regression Analysis
- Interpretation of the results

### V. Conclusions and Summary

- What have you learned?
- What are the policy implications?
- Are there suggestions for future research?

### **Examples of Topics:**

The topic of the paper is up to the group, although you should clear it with me. Pick an area where data are readily available, and avoid topics requiring new surveys (although this has been done) or involving simultaneous equations (for example supply and demand). Experience shows that cross-sectional studies work better than time-series for this assignment, although the later has been done. The Census volumes, for example provide excellent and abundant data on U.S. states and cities. Topics from previous years include:

- State by state variations in divorce rates      -Voter behavior
- Crime rates in U.S. cities      -Wage, earnings or poverty differentials by state
- The determinants of teen-age pregnancies by state
- Catholic Church scandal influence on donations
- State by state alcohol consumption      -Baseball attendance across cities
- Salary determination in major league sports -Determination of MCAS results
- Determinants of U.S. strike activity      -Determination of welfare participation
- State by state variations in suicide rates      -State by state variations in traffic fatalities
- State by state variations in fertility rates      -Voter turnout in Presidential election years

## Course Outline:

### **Part I**      **The Basic Regression Model--Cross Sectional Data**

1. Introduction--An overview of Regression Analysis (Chapter 1, 19, Appendices A, B, C)

The big picture, What is econometrics and regression analysis, The setting, Examples.

2. Ordinary Least Squares (Chapters 2, 3)

Estimation--Single independent models, Multivariate regression, The classical assumptions, Gauss-Markov Theorem, Omitted variables, Irrelevant variables, Multicollinearity.

3. Multivariate Regression: Inference (Chapter 4)

t-tests, F-tests, Economic vs. Statistical significance.

4. OLS Asymptotics, and Further Issues (Chapter 5, 6)

Consistency, Scaling, Functional forms, etc.

5. Qualitative Information (Chapter 7)

Dummy Variables.

6. Heteroskedasticity (Chapter 8)

Consequences, Robust Inference, Testing for heteroskedasticity.

7. Specification and Data Problems (Chapter 9)

Lagged dependent variables, measurement error in dependent and explanatory variables.

**Part II and III      Time Series and Advanced Topics**

8. Regression with Time Series Data (Chapter 10)
  
9. Panel Data Methods (Chapter 13)  
Chow Test, Policy Analysis.
  
10. Advanced Panel Data Methods (Chapter 14)  
Fixed Effects, Random Effects.
  
11. Instrumental Variables Estimation (Chapter 15)  
Two Stage Least Squares, Endogeneity
  
12. Simultaneous Equation Models (Chapter 16)  
Identification

One component of this course is an econometrics paper. This will be a practical application of regression theory to test some hypotheses that your group finds interesting. The topic is entirely up to you, and can come from economics or some other discipline. These papers should be done by groups of three students.

Doing a research paper is the best way to learn how to apply econometric techniques to a practical problem. But the work is time consuming so it should be spread out through the course of the semester.

### **The Topic**

The topic of the paper is up to you, though you should clear it with me. There are literally thousands of interesting data sets that are easily accessible via the Internet. Large household surveys often make for interesting data sets, because they contain lots of variables for large samples of people. Topics that have been done in econometrics courses in the past include:

- The impact of race on the severity of sentencing for murder
- Determinants of teen-age pregnancies
- The determinants of annual strike activity in the U.S.
- Voter turnout
- The impact of plane crashes on airline stock returns
- The determinants of wage rates over the lifetimes of workers
- The impact of family background on earnings
- The determinants of price variation in collectable comic books

### **The Big Four: Narrow, Interesting, Original, Feasible**

Consider some bad topics. First, is there life on other planets? A hell of an interesting question, but it's just not feasible to try to answer it. Next one—"Economic performance in the 20<sup>th</sup> century." Much too broad, not specific enough. How about "The effect of education on earnings?" That's narrow and feasible, but not original or interesting, since it's been done a thousand times.

A good topic hits on all four of the attributes listed above. When you think of a topic, rate it from 1 to 10 on each of these four criteria. That will help you decide whether to pursue it.

### **Data Sets from the Web**

The World Wide Web is a great tool for accessing data. Here are just a few leads, but realize that the possibilities are unlimited.

1. <http://FMWWW.bc.edu/ec-p/data/ecfindata.html>

A compendium of economic and financial data sets available at Boston College, including data dealing with stock returns, exchange rates, the macroeconomy and international comparisons of economic activity.

2. <http://www.worldbank.org/lsms/>

Home of the World Bank's Living Standards Measurement Surveys, household-based survey data for a variety of countries.

3. <http://www.umich.edu/~hrswww/>

The Health and Retirement Study, a state-of-the-art survey of households approaching retirement, with detailed information about health, income and wealth. Useful for investigating many policy issues related to Social Security reform, saving and health.

4. [http://www.nber.org/data\\_index.html](http://www.nber.org/data_index.html)

A data set archive maintained by the National Bureau of Economic Research.

5. <http://fmwww.bc.edu/ec-p/data/scf98/>

The most recent Survey of Consumer Finances data set, collected by the Federal Reserve Board.

6. <http://www.icpsr.umich.edu/archive1.html>

The Inter-University Consortium, kept at the University of Michigan (contains dozens of data sets of all kinds).

An example of how to grab data and information from the Web—a fast and easy way to get very useful data

I downloaded the 1998 Survey of Consumer Finances data set. It is a large survey of households conducted by the Federal Reserve Board. Looking through the documentation for this data set (i.e., the listing and description of the data set's variables) I found several interesting questions. Here's one:

X7100 When making major decisions about credit or borrowing, some people shop around for the very best terms while others don't. (What number would you be on the scale?/What number would your family be on the scale?)

1. \*ALMOST NO SHOPPING
- 2.
3. \*MODERATE SHOPPING
- 4.

## 5. \*A GREAT DEAL OF SHOPPING

I think this is potentially interesting. Why are some people picky about borrowing rates and others not so picky? Shopping around is likely a time intensive activity, so maybe the propensity to shop around has something to do with wages. People with kids are probably strapped for time, so maybe they don't shop around so much.

The main point is that it is easy to access high-quality data from the web, and it took all of about 30 seconds to find an interesting question.

### **Putting the paper together**

Always work from a clear outline in crafting your paper. Here is a suggested outline:

#### 1. Introduction

- What is the hypothesis to be tested?
- Why is the hypothesis important and interesting?
- What does theory predict about the relationships being studied?

#### 2. Brief Literature Review

- What other work has been done on this issue?
- What has been found?
- How does your paper differ from the others? How is your paper better?

#### 3. Data

- What data would be ideal for testing your hypothesis?
- What data have you been able to obtain?
- What special data problems did you encounter?

#### 4. Empirical Work

- Regressions
- Interpretation

#### 5. Conclusion and Summary

- What have you learned?
- What are the policy implications?

- What are the directions for future research?

Target your completion date to be a week before the actual due date. There are two reasons for the early target date: (1) to leave enough time for rewriting the paper--rewriting usually improves the paper a lot--and (2) to guard against unforeseen difficulties at the end of the project.

### **A Few Miscellaneous Pointers**

We will be discussing how to do a research paper throughout the semester. Here is a partial list of the most important do's and don'ts:

- A good research topic is always narrowly focused. An example of a poorly focused, broad topic: "An Investigation of the Overall Competitiveness of U.S. Industry from a Historical Perspective." ("Competitiveness" is a vague term--we have no idea how it is linked to measurable variables. It would probably take a large book to cover this broad topic.) An example of a more narrowly focused topic: "Does the 55 Mile Per Hour Speed Limit Reduce Highway Fatalities?" A good check for whether your topic is narrow enough--can you say what you want to do in one clear sentence?

- Never make an unsubstantiated statement. For example, never write something like "Policymakers have become increasingly concerned with the rising cost of health care" and leave it at that. You must back up every statement with specific facts and examples.

- If you quote or paraphrase another source you must cite the source in a footnote. Any direct quote longer than a few words should be footnoted.

- Please put footnotes on the same page rather than at the end. It saves time since the reader does not have to flip from text to end and back.

- Any sources used in researching the paper should be cited in the bibliography.

- Use the *American Economic Review* (AER) as a style guide. That is, photocopy any article out of the AER and do the footnotes and bibliography (references) the way they are done there.

Attached is a copy of Boston College's statement on academic honesty. Please read it carefully and feel free to discuss it with me if relevant issues arise.

### **Academic Integrity**

The College expects all students to adhere strictly to the accepted norms of intellectual honesty in their academic work. Any form of cheating, plagiarism, or dishonesty, or collusion in another's dishonesty is a fundamental violation of these norms.

--*Cheating* is the use or attempted use of unauthorized aids in any exam or other academic exercise submitted for evaluation. This includes data falsification; the fabrication of data; deceitful alteration of collected data included in a report; copying from another student's work; unauthorized cooperation in doing assignments or during an examination; the use of purchased essays or term papers, or preparatory research for such papers; submission of the

same written work in more than one course without prior approval from the instructor(s) involved; and dishonesty in the requests for either extensions on papers or make-up examinations.

--*Plagiarism* is the deliberate act of taking the words, ideas, data, illustrative material, or statements of someone else, without full and proper acknowledgement, and presenting them as one's own.

--*Collusion* is assisting or attempting to assist another student in an act of academic dishonesty.

As part of your scholarly development, you must learn how to work cooperatively in a community of scholars and how to make fruitful use of the work of others without violating the norms of intellectual honesty. You have a responsibility to learn the parameters of collaboration and the proper norms for quoting, summarizing and paraphrasing. Your faculty instructor can give you additional information and instruction in this area.

--*Sanctions and Procedures*: Faculty members who detect any form of academic dishonesty have the responsibility to take appropriate action. They also have the responsibility to report the incident and penalty to the department chairperson and the appropriate class dean. This report will remain in the offending student's file for five years beyond the date of his/her graduation. In cases of multiple or very serious offenses the Dean of the College may take further action up to and including exclusion of the student from the University.

If the gravity of the offense seems to warrant it, or if the faculty member prefers that another academic authority decide the matter, he or she may refer the case directly to a dean.

On the other hand, if the student feels that a faculty member's decision is unfair or excessive, he or she may appeal the matter to an associate dean or to an Administrative Board. Further details about this process may be found in the College of Arts and Sciences Sections of the Undergraduate Catalog.

This handout is designed to help you with that decision and to give you a better idea of what the termpaper project is all about.

*What makes for a good termpaper topic?*

A good termpaper topic has four basic attributes. It is

- Σ Narrow
- Σ Interesting
- Σ Original
- Σ Feasible.

Let's consider these one at a time.

*What do you mean, "narrow"?*

Narrowness is the most important attribute of a good termpaper. It is easiest to see the virtue of a narrow topic by first looking at its opposite, a topic that is too broad. Here's an example of an overly broad topic:

"How do environmental factors influence a person's behavior?"

This topic is so broad and diffuse that it hardly qualifies as a topic at all. It's more like a general topic area. The main reason why it is such a bad topic for a 15 page termpaper is that, in that short a space, you could hardly scratch the surface of such a vast issue. Jared Diamond had to devote the equivalent of 1,000 termpaper pages to this vast problem. If you tried to address such a huge subject in only 15 pages, you would wind up with something that was impossibly watered down, uninformative, and dull.

So let's consider an example of a narrow topic then, to see why it is so much better:

"How does the output of jazz recordings by male jazz artists vary with their age? And how does the pattern for men compare with that of women?"

It's easy to see that this question is much more specific than the first. Basically, it boils down to plotting one or two simple graphs, or maybe a spreadsheet. A person working on this topic might, for example, get the names of the top-selling jazz albums produced in the past couple of years. He or she could use a source, such as *The Essential Jazz Companion*, to get the information about each artist's age and sex for each recording. Then it's a matter of writing down the information on a spreadsheet. For example:

<u>Album</u>	<u>Artist</u>	<u>Sex</u>	<u>Age</u>
1. Groovy Sounds	Bill Wackencracker	Male	24
2. Super Vibes	Mary Warbler	Female	60
3. Funky Fusion	Matt Patheny	Male	47

etc.

You might, for example, have quite a large sample of recordings. You might also have to make some decisions about how to treat bands. Maybe you would ignore them, and just concentrate on single artists. Maybe you could just write down the information for the person fronting the band.

This leads to a potentially interesting set of data, which you could then display in various forms, to give some idea of what's going on. You might find, for example, that most jazz albums are recorded by males in their early 20's. You might find that male jazz artists produce most of their work early in life, and then produce almost nothing later. Your information would probably show that men dominate the jazz field, and that women artists produce albums much more consistently throughout their careers, instead of bunching their output early in their lives.

This evidence is interesting, because it relates to Geoffrey Miller's sexual display hypothesis. Miller's idea is that artistic displays are used to attract mates. And, since it is mainly males who need to display, it is no surprise that they would tend to produce the most. In addition, the age pattern is also consistent with Miller's hypothesis—displays pay off the most during the prime mating years.

In fact, this example comes from something very similar to a study that Geoffrey Miller actually did. And what he found was similar to the patterns described above.

*What's interesting?*

The example above shows that narrowness and interestingness can be tied together, but not always. The guiding principle is this: Something narrow is interesting if it provides a window through which we can get insight into the "big picture." In the case of the jazz musician study, the big picture was Miller's sexual display hypothesis. That big picture is important, because if it is really true that males have an incentive to show off their fitness in various ways, it has several implications about how we interpret peoples' behavior—everything from binge drinking to the pursuit of science and engineering.

So, while the question itself is narrow—involving, for example, something as specialized as counting jazz albums—that question should fit into the broader scheme of things.

*Just because something is narrow, does that mean that it is automatically good?*

No. Here's an example. Suppose someone starts off with a very broad topic, like "The Impact of Environment on Behavior." To make it more narrow, they try an artificial

approach: they might focus just on men, or just on people living in Boston, or just on male students living in Boston in the 1960's.

True, this topic is narrower than before, but it is still hopelessly vague. Why do we care about male students in Boston in the '60's? No reason is given. Maybe there is one, but it needs to be provided. We need to know how that narrow topic fits into the broader scheme of things, in order to see why it is important.

The trouble with the title, despite the apparent narrowness, is that the words "environment" and "behavior" are far too vague. Here's how the situation can be improved. Substitute something specific for "environment." One critical part of the environment of male students during the 1960's was the draft. Men were subject to be drafted into the armed forces to fight in Vietnam. One critical part of their "behavior" was going to school. So one narrow topic is, "What Impact Did the Draft Have on Men's Decisions to Go to College."

*How are we supposed to come up with something original? Hasn't everything that matters already been researched?*

You would be amazed at how little we know about so many topics, especially those related to what we are discussing in this course. There are new sources of information coming out all the time: mostly in the form of data sets that you can pull off the Internet. And there are old sources of information, like the jazz encyclopedia used by Miller, that can be put to new and interesting uses.

*But what if I'm exploring a question and I can't come up with any useful information about it?*

This can happen, but it's easy to spot the danger signs so you can avoid this problem. The problem has to do with the *feasibility* of your topic. Before exploring ways to make sure your topic is feasible, let's look at one that flunks the feasibility test even though it scores well in other ways.

Consider the question, "Is There Life on Other Planets?" It's definitely narrow, original, and is probably about the most interesting and important question there is. But it's just not feasible. Carl Sagan's graduate students wasted a better part of their careers listening for radio signals from other planets when they could have been spending their research time more productively.

You can avoid this problem by checking in advance that the information you need is indeed available. For example, if your topic has to do with how attitudes toward drugs have changed over time, you should make sure that such information is available, and the best place to start is the web.

So early in the process of investigating a research topic you should be able to *prove* that the required information for settling the question is indeed available.

