Syllabus for AEM 7100: Econometrics I
Spring 2014

Lecture Time: Mon Weds 2:30PM - 4:00PM
Class Location: B02 Warren Hall
Professor: Shanjun Li
     405 Warren Hall; SL2448@cornell.edu; Phone: 607-255-1832
     Will use blackboard site to distribute homework and materials
     Office Hours: Mon 4:05-5:05PM at 405 Warren Hall
TA: Leah Bevis leb99@cornell.edu
    Office hour: Thursdays 8:30-10:00 (with a few exceptions) at 434 Warren

Summary: This course provides an introduction to empirical methods that are commonly employed in applied microeconomic research. Together with AEM 7110, it provides a graduate sequence in econometrics. The course is applied in nature: our focus will be on learning the tools necessary for carrying out empirical work through hands-on data work and analysis; less efforts will be spent on proving theories and deriving properties of the estimators. My goal is to provide you with an understanding of some commonly used empirical methods that will allow you to make informed methodological decisions when conducting your own empirical work. ***The main leaning method of this course is learning by doing so I expect you to spend a lot of time on problem sets, programming, and doing research project. Coming to classes alone is very far from being enough.***

Textbooks: The required text for the course is *Econometric Analysis* by William Greene 7th Edition (4th, 5th or 6th should be ok too), Prentice Hall. A hardcopy is on reserve at the Mann library. Another required text is *Discrete Choice Methods with Simulations* by Kenneth Train. Download at http://elsa.berkeley.edu/books/choice2.html.

Other texts that might be useful are: *Econometrics of Cross-sectional and Panel Data* by Jeffrey Wooldridge and *Microeconometrics: Methods and Applications* by Colin Cameron and Pravin Trivedi.

Prerequisites: Matrix algebra and statistical methods courses at level of AEM 6940, ILRST 3110 or ECON 6190. Multivariate calculus, matrix algebra, probability and distribution theory, statistical inference, and an introduction to the multiple linear regression model. Appendices A and B in *Greene (editions 6 or 7)* are assumed.

Software: You should learn to use both Stata and Matlab. Stata is a powerful statistical program with a broad set of pre-programmed econometric and statistical tools. It is quite popular among economists, and is continuously being updated with new methods. It is an excellent package for most econometric analysis, but is limited when you want to use new or less-common econometric methods which have not yet been programed. Matlab is a high-level matrix programming language with a wide variety of built-in statistical functions. Many econometric methods have been programed in these languages and are available on the web. The advantage of using Matlab is that you are in complete control of your analysis, and it is easier to program new methods than in Statat. Some

disadvantages are that you have to do much of the programming yourself, programming complicated procedures takes significant time, and programming errors are hard to prevent and difficult to detect and eliminate.

In short, Stata could be very handy for data cleaning and routine analysis and Matlab is preferred for more sophisticated methods (not yet canned in Stata routines). I switch back and forth between the two depending on the specific project and in many cases, I use both. If you strongly prefer to stick with only one software, that is fine. But you will find that for more involved methods such as simulated MLE, Matlab will be a better tool than Stata. Moreover, Matlab has built in parallel processing in recent versions, which are very helpful for computationally intensive work.

For many exercises, I will ask you to code things up in Stata and Matlab. Data sets needed for the exercises will be distributed to the class via blackboard course website.

Evaluation: Grades for the course will be based on:
• Six or seven problem sets (total 30%)
• Midterm examination (20%)
• Group/individual empirical project (50%)

Group/individual Empirical Project: I expect each of you to work in groups of 2-3 students (if you strongly prefer work by yourself, that is fine as well). You should form your groups no later than March 1st and start to think about the project as soon as possible. If you could not find a group to join by March 1st, email my TA Leah so that she can help you. Ideally, you should find your own topic and data (this will be rewarded in the evaluation) and I would be happy to guide you through the process. But if you have trouble with this, talk to me no later than April 1st, and I can provide a set of topics and potential data sources. Another way to fulfill this requirement is to replicate an existing study. Many journals now require authors to post their data set and you can find an article to replicate from these journals (American Economic Review, Journal of Political Economy, American Economic Journals, Journal of Applied Econometrics etc.).

The paper should follow the standard format of a research paper in economics and should be 15-30 pages (no including references but all else) in font 12 and double spacing. **Due: 5PM on May 15th.** I need both an electronic copy and a hard copy to Leah (cc-ing me).

Topics:
1. OLS mechanics
2. Finite sample and large sample properties of OLS
3. Hypothesis testing
4. Regression extensions
5. Instrumental variable
6. Panel Data
7. GMM
8. System of Equations
9. MLE
10. Discrete choice model