Empirical Methodology

This handout will guide you in how to organize your empirical methodology section. Your paper can be broken down into the following outline:

I.	Introduction	IV.	Empirical Methodology
II.	Literature Review	V.	Results
III.	Economic Model	VI.	Conclusions/Extensions

It is important to keep sections III and IV distinct, but you can combine the Economic Model section into the Literature Review and/or the Empirical Methodology if it makes sense for your individual paper.

The economic model section is designed to take an abstract economic theory (such as demand and supply for housing, marginal revenue product, education production functions, or the efficient markets hypothesis) and use this model to indicate how the variable of interest (housing prices, athlete salaries, student performance, or stock prices) will behave in different circumstances. It would be useful to the reader (and you) to apply the model in an abstract sense in a couple of case studies, focusing on the explanatory variable(s) of interest. For instance, use the supply and demand model to illustrate how an increase in the interest rate will affect the demand for housing and the equilibrium house price.

The organization of empirical methodology section is given below. This may vary for individual projects; don't hesitate to talk with me if you are unclear on an alternate organization of the methodology section.

A. Data
B. Stylized facts from the data
C. Regression equation

A. Data

In Section III, you will have discussed specific explanatory variables that you are interested in and their predicted effects on your dependent variable. You want to begin this section with a description of your data sources, coupled with a description of the data. Is it a cross section, such as, individual NBA player salaries in 2002? Or a time series, such as daily S&P 500 stock prices between 1929 and 2003? This part does not require any analysis, you are simply conveying to the reader what data you are using.

B. Stylized facts

It's useful for the reader to get a sense for what the data looks like. This subsection should include graphs of the key aspects of the data (if dealing with time series) and/or summary statistics (mean, standard deviation, median, etc.) with some cursory observations to guide the reader through this information. What was the median home price in Sacramento in 2002 versus 1992? You could show this using a time series plot. What is the average student test score in your sample? How much does it vary? These questions could be answered looking at summary statistics. For cross section studies, the distribution could yield some interesting observations – this can be illustrated using a histogram. You are the judge of how to present this information to the reader – you need to be familiar with the data yourself before making these judgments. Spend some time looking at different sub-samples of the data (top 10% of salaries or test scores, etc., the behavior of stock prices pre and post 1945, etc.).

C. Regression Equation

Now, you're ready to remind the reader of your particular test and how you are going to go about using regression to test it. This section should include a regression equation, a discussion justifying this equation, and a description of the expected signs on the coefficients for each of the explanatory variables (spending more time on those that are of particular interest for your study). Remember, the regression coefficient measure the marginal effects of the explanatory variable on the dependent variable (holding the other variables constant, ceteris paribus). Please review previous assignments where we worked on regression equations if you don't feel comfortable interpreting them. *When justifying your regression equation and discussing the expected signs for the coefficients, you should make some clear connections back to your theory section.*