The first objective of the final project is to **replicate** the relevant tables and figures from one empirical economics paper that interests you. You do not need to replicate every table in the paper. Choose the tables and figures that constitute the main points of the paper. Replication can be an illuminating experience, especially if you choose a paper that really interests you.

Assignments 3-6 will encouraged you to work steadily throughout the quarter on this project and these first steps should also be included in the final project. The second major objective of the final project remains: your **original extension** and write-up of the paper to be submitted by noon on Friday, **March 23, 2019**.

Key reading on how to write a great paper that starts with replication: King, Gary. "Publication, Publication," PS: Political Science and Politics, Vol. XXXIX, No. 1 (January, 2006), 119-125.

You are also encouraged to get inspiration on how to write a great replication paper from the prominent replication papers listed in the '*Replications*' file on the course Canvas site.

Some suggestions for how to approach your final project:

You need not address all of these questions, and you should write a well-argued final project rather than a bullet-pointed response to these items. I will not accept the submission of raw output from statistical software packages, but I welcome creative and elegant visual displays of results or data.

Explain the economic problem or research question that the original paper addresses. Can you explain the paper in plain English to a non-economist? What is the dependent variable? What is the key independent variable? What is the proposed causal effect and its mechanism? What are some confounding factors (e.g. omitted variables, measurement error, reverse causality, or self-selection) that the original paper addresses? Is the estimated effect economically significant?

What is the unit of observation (e.g. countries, calendar quarters for a single country, country-years for a panel of countries, sub-national regions such as states of the United States, firms, households, or individuals)? What are other relevant dimensions of the data (e.g. frequency of the data or time or geographic span)?

How did you gather the data, for example, downloaded the cleaned data from the author's or journal's website or reconstructed the data from the description in the paper? Note that the latter empirical strategy will require more effort in terms of data work, but also widens the range of original extensions you can conduct.

Compare your attempt at pure replication of the results (e.g. summary statistics and regression output) in the original paper. Do your results differ in sign, size, or significance from those of the original paper? If so, can you explain the difference (e.g. data source, data construction, estimation procedure)?

Explain your extension of the original paper. You might consider overlooked confounding factors, shortcomings in the estimation procedure, outliers and leverage points, or limited external validity of the model. Compare the sign, size, and significance of the results in the extension to those in the original paper.

Guidelines on the final paper to be submitted:

The paper you hand in should be in PDF format. You are strongly encouraged to use LaTeX to write up your paper.

You should aim to have the paper in the format and style used in economics journals. Organize the paper into sections. Your paper should at a minimum include:

1. **Introduction**. Clearly state the question being asked and why it is of interest: *What do we want to know and why should we care*? It is usually helpful to indicate where this research fits into the literature. The last paragraph should briefly summarize the main empirical findings.
2. **Empirical Strategy**. Clearly state which empirical strategy and econometric method(s) are chosen to answer the (causal) question posed? Is it the most suitable empirical strategy for the research question? Provide a formal description of the empirical strategy in the paper.
3. **Data**. Describe the data; i.e. which sources it stems from, variable definitions, and descriptive statistics. Describe variable construction and data sources, such that your work can be replicated by any qualified person. Describe measurement and how it maps into the relevant theoretical constructs. It is helpful to be conscious of the "ideal" data set and discuss the relevant ways in which the data depart from this "ideal" and, if appropriate, how the econometric techniques are modified to accommodate this.
4. **Empirical Results**. A general discussion of the characteristics of the results: Provide tables of estimates and your interpretation of your findings. The accompanying tables/figures of results should be presented in a professional manner and be sufficiently self-contained. The reader should be able to ascertain all relevant information from the tables/figures without having to consult the main text.
5. **Conclusion**. Describe what is learned from the paper. Summarize the question you set out to answer and what the empirical evidence suggests is the answer. Provide any further caveats and discuss what would be fruitful avenues of future research to extend this work; e.g. what further analysis or data collection might be suggested by your results.

Additional sections are possible, but are not required. For more advice regarding the write-up, read these "Writing Tips for PhD Students" by John Cochrane.

Papers should be 1.5 spaced with 1 inch margins and should have no more than 15 pages of written text. That is, excluding tables and figures.

You should also submit the **data** and the clearly commented **code** that generated the results in the final project, such that we can readily match the code with the results in the paper. Best practice is to keep the raw data in its original form, and use one Stata (or SAS or R) script that can generate *all* your results. For more advice regarding data management, read this guide for RAs by Matt Gentzkow and Jesse Shapiro.

Grading criteria will include the originality of your extension of the replicated paper, the correct implementation of econometric methodology learned in class, the quality of the writing, clarity of presentation of analysis and results, critical discussion of the replicated paper's and your own analysis and suggestions for improvement. Formatting of the paper and readability of the code will also have a small positive weight. Your ability to connect the analysis to the material we have covered in class will be given considerable weight: you should include a discussion of the assumptions required for the econometric analysis to give a credible answer to the research question, and a detailed discussion of whether you believe these assumptions to be satisfied. Remember that it is fine if you believe they are not satisfied, as long as you very clearly explain why you think this is the case, and how you might improve upon the current analysis.