

Guidelines for Students Doing Independent Research/Study with Professor Holmes
(6/16/2009)

Your objectives in doing independent research should be to learn how to criticize and analyze research done by others, how to do your own research and, hopefully, to have it published. This is a very difficult task. However, the benefits are considerable. Hopefully, You will change from being an academic tourist to a participant/producer. I believe that the skills, self-discipline and attitude required are transferable to other parts of your life. In particular, making personal decisions and career decisions (I doubt if many of you will become college professors but I believe the skills you learn doing independent research should be applicable to your career, whatever that may be. To me this is what good education is about. A good research project should result in a paper that is of publishable quality and this will increase the probability you will (want to go to and) be accepted at a good graduate program or be employed to do interesting economic/statistical analysis in a good job.

The simplest way I have found for either a graduate or undergraduate student to learn these skills is to try to replicate the results reported in a published paper. This is the best way to both learn statistics/econometrics and how to think critically/analytically.

There are three stages or parts to your independent study; 1) how to pick an empirical study to attempt to replicate 2) how to replicate the results published in such an article. 3) how to extend, correct, revise, and create your own contribution (hopefully, joint with me).

1) How to Pick an Empirical Study to Replicate

You should only consider articles from peer-reviewed journals. Papers and articles that are not peer-reviewed, e.g. some easily found on the web, commonly have stupid arguments and boneheaded evidence. Verifying this rigorously will not be surprising and therefore not of much interest. An article published in a peer-reviewed journal has been read and approved by at least two people that have some expertise in the area and probably quite a few more. It is very likely that such an article may be the basis of research subsequently published by others and, therefore, of considerable interest to some people.

You should pick a topic that you personally find interesting because you are going to spend a lot of time on it and that should be as much fun as possible. It will not be fun if you find a topic boring to start with. I suggest that you start by perusing some general interest journals, which are not highly technical, such as *The American Economist*, which is written for undergraduates, *The Journal of Economic Perspectives*, which is written for a general audience, and perhaps *The Journal of Economic Literature*, which sometimes publishes surveys about broad issues or topics. You should

not pick a topic that is too complex or sophisticated. If you don't understand it, how are you going to be able to criticize and extend it? The test of whether or not you understand something is whether you can clearly explain it to someone else. Obviously, there will be some things you don't know and you don't know how to do, but you can learn these (perhaps with my help). However, the central question addressed, the approach taken, and the issues raised should make sense to you and you should find these interesting. You can find examples of what other undergraduates have published in the *Undergraduate Economic Review*, which only publishes such papers.

In order to find a specific article to focus your efforts upon, I suggest a top-down approach. Starting with a general interest article, then research the articles referenced and peruse them. These may lead you to a set of journals that specialize in a particular area and you can then skim through articles in them. Alternatively, having found a general topic/article you can search minor journals that are not top-ranked using Google or Google scholar. (There are several rankings of journals that you can research online.) You should keep notes of your research pilgrimage, so that if the first article you choose turns out not to be productive you have some alternatives. I find it helpful to keep notes of why I take a certain direction and what I find there. However, bad luck happens. Similar to other aspects of your life, e.g. love and investing, I advise that you don't stick with a losing strategy.

2) How to Replicate the Results Published in An Article You Choose

After you have selected an article to research, you must read and analyze it intensely. I personally have spent between 20 and 60 hours reading, outlining, and understanding an article that I wish to focus upon. When you have reached the point at which you can present the argument made orally and/or in writing without reference to the original article - as if you were the author - then you're ready for the next step; to replicate the results.

There are two parts to this: (1) obtaining the data, and (2) replicating the statistical estimation and tests. It is usually very difficult to obtain the identical data used by an author [So much for the claim that economics is a science.] Some journals require the authors to submit their data and it is available to whoever asks for it. Usually this is not the case. In either event, you need to research the data as if you were doing it from scratch. You need to read the detailed description of how the data were gathered and reported in the study you are trying to replicate. THE DEVILS IN THE DETAILS!!!! Think of yourself as a detective. You need to know what alternative data are available and might have been used. Your discovery and thinking about this may provide a natural extension, revision or correction of the study/article.

Let me give you two examples from my own research on the relationship of employment and wages of production workers in manufacturing. Economic theory treats these variables as occurring in

continuous time. However, the US Department of Labor reports that they collect the data based upon a survey of only some firms for the week that includes the 12th day of each month. Think about how closely this corresponds to the concepts in economic theory. How does this affect the econometric issue of whether or not a variable is measured with error? Do you think that such a fact should influence how a dynamic theory is estimated and tested? [More than 3/4 of the possible observations are not reported!!] My second example should be equally illuminating. Both inflation and its possible effect on the purchasing power of earnings are commonly measured by the CPI. However, there are several measures of the CPI. The most popular is for all urban consumers. An alternative is the subset of urban wage and clerical workers. Clearly, the latter is the more relevant measure for the purchasing power of the earnings of production workers in manufacturing. (In addition, about 40% of both of these are made up of the cost of housing – not relevant for many purposes, e.g. The Expenditures of teenagers.) It is very likely that you will find reasonable alternatives for the measurement of many of the variables used in the article you are researching, and that the author will not clearly state which is employed. Many researchers do not even think about such issues. You want to be the exception, and on your way towards being a scientist. (By doing this you have an easy and natural way of extending or redoing the empirical analysis in the paper you are researching using more appropriate measures of the variables.)

Having obtained the possible data employed in the study (possibly in an Excel spreadsheet/format), it should be an easy task to import this data into a statistical software program and replicate the statistics reported in the article. It should be fairly easy to learn this on your own, as these programs are pretty easy to master and use. This should be the least daunting of your tasks.

The more challenging part involves replicating the transformations that the researcher may have performed upon the data. Is the data transformed into logs? If the data is say monthly, then is it first differenced on a monthly basis or perhaps on a 12-month basis? Such transformations for time series data are usually required to make the series stationary and to avoid spurious regression results. An alternative to differencing the data is to fit a deterministic trend and then to use the residuals, which by definition are detrended, as the input data. My own experience is that the majority of time series studies of the minimum wage failed to address this issue at all. Another issue is seasonal adjustment. Some data series are available already seasonally adjusted and some are not. Seasonal adjustment can be a major transformation of the data. For example a bank's holding of vault cash normally varies more between the days of the week, e.g. Monday through Friday, than it does between one year and the next for the same day of the week. If such data was not seasonally adjusted and the day of the week the variable was measured was different for different observations, then the results could be tremendously different than if the data had been seasonally adjusted.

At this point, it should be obvious that the difficulty in replicating the results from an article is that you must give it a lot of thought. However, that is exactly what you want to learn how to do. You want to learn how to think critically and analyze an empirical proposition or question. This brings me to the third point.

3) How to Extend, Correct, Revise, And Create Your Own Contribution

or

THERE IS MORE THAN ONE CORRECT WAY TO DO A THING!

Let me suggest what your attitude should be in reading an article and replicating its results. Most people, and especially students, are intimidated by the printed word. We believe that because it is printed, it has to be correct and is true. I recommend you take the opposite attitude. The author has made some stupid mistakes and your job is to find out what they are. You are a detective (whose true abilities have perhaps not yet been recognized). You should believe that the author of the article you are researching has made little mistakes and perhaps big conceptual ones. You just have to discover them. You should note or highlight spelling and punctuation errors (excluding mine as that is hopeless), errors in mathematics or typos (and ascertain whether they are typos or actual errors), you should look for errors in the references, including miss quoting an author or misinterpreting them. The possibility of data errors and, in particular, using an inappropriate measure and assumptions for theoretical variable I have already dealt with above.

Because I believe we comprehend complex, abstract issues primarily by the use of paradigms, I encourage you to consciously endeavor to learn to do this. (Just as I have encouraged you in my classes to apply the critical and analytical skills you develop to the decisions you make in your personal life and career area). As an example, I will give you another paradigm of the inappropriate choice of variables. Let us say that the issue is the estimation of the demand for money in the United States. One determinant should be the nominal expenditures of demanders. Presumably, a relevant measure is disposable personal income per capita. How should per capita be measured? Total population is commonly used but that includes very young children and the institutionalized part of our population, neither of which have a demand for money. Alternatively, one could consider the number of households and the number of people in the labor force. The difference in the results from using one of these versus the other can be quite substantial. I believe without a great deal of thought you can justify the choice of one of these over the other without a great deal of difficulty, if you asked the question; how would you, and those people you know, behave?

I believe that the availability of very user friendly, sophisticated statistical and econometric software has led many researchers to produce poor research. What is in short supply is thinking about a problem. You should ask a series of questions. What is the fundamental hypothesis or question being addressed/investigated? What are the null and alternative hypotheses the author is

investigating? It is likely that they have not articulated this very well and perhaps not thought about it. You want to do so. The correct statistical/econometric tests depend upon what the alternative hypothesis is. It is likely that you can extend the article you are researching by simply correctly testing the null hypothesis against the alternative.

You should question whether the author has posed a testable hypothesis and/or are there other tests that should be relevant? For example, correct statistical/econometric tests depend upon whether the null hypothesis is nested in the alternative hypothesis (is a proper subset of) or whether it is non-nested. Perhaps the author did the wrong test. As a separate issue, the validity of these tests rests upon a set of assumptions that are routinely tested as output from standard software programs. Are these assumptions plausibly satisfied in the study you replicate? If they are not, this may suggest a way to redo or extend the study.

You should not think that you can or should address all of the possible statistical or econometric defects and problems in the paper you are researching. You should be very happy (and I will be very pleased) if you can find and address just one of them. It doesn't matter which one. Once you do one successfully, you will gain the confidence to do another one. You will find, hopefully, that doing research is fun once you start finding mistakes and seeing how you could possibly improve the results that were published.

Big conceptual errors are obviously much more difficult to discover/recognize. My advice is to try to think about the framework the author is using and try to conceive of alternatives. They exist. My subtitle for this section is "THERE IS MORE THAN ONE CORRECT WAY TO DO A THING!", So think of alternatives.

I also suggest you try to outline the structure and direction of causality. This should be the most important and critical part of the framework of an article. You may find an author is unclear or confused about this most basic concept. David Hume's classical definition "If A causes B, then when A occurs then B occurs", implies that $B=f(A)$. If the assumptions an author makes implies or includes the possibility that $B \neq f(A)$, then A cannot cause B. (This basically is what statistical hypothesis test is about.) At least three of my early theoretical publications involve demonstrating this mistake. I do not believe there is any shortcut to learning to think in this manner, but I do believe that we all improve with practice.

I hope these guidelines will help you to produce really successful independent research with me. Please feel free to e-mail me or call me at home with any problems or questions you have or to arrange a meeting at my office.

I recommend the very old, readable, clear, undergraduate exposition of econometrics by G.S. Maddala at the undergraduate level, and I use William Greene's graduate text *Econometric Analysis, fifth edition* as a reference for more sophisticated topics/issues.

Good Luck and Have fun!!!