

ECON 435, Term paper: Your own empirical project

November 8, 2015

The assignment

The goal for this assignment is to replicate the empirical results of a paper that you select.¹ I expect you to select a paper, obtain the data, and replicate the findings in the selected paper. I also need you to summarize the main idea behind the paper in your own words, focussing on the empirical strategy and the interpretation of the results.

Organization

1. **Deadline** to have your paper **approved**: November 18, 16:59. That is not the deadline for submitting it, but for me to **approve** it.
2. **Deadline** to **hand in** your assignment: start of the final exam.
3. Hand in, by email, a .zip or other **compressed** format with the same ingredients as the problem sets, and a pdf of the paper you are replicating.

Plagiarism

- Do not write down anything you do not understand
- Anything that you copy or paraphrased **must** be attributed to the original source. I take this very, very seriously. Exceptions: our textbook or my notes.
- Refer to **other people's code** as well, if you use it. This could be other students, an online source, or ...
- You are not allowed to use paid assistance, like tutors or proofreaders.
- When in doubt, contact me!

¹If you are passionate about conducting your own research instead of replicating the research done by others, you are welcome to come and pitch your research idea to me.

Conditions on the selected paper

The paper you are going to replicate must be approved by me by the deadline above. When I decide whether the paper you chose is OK, I take into account the minimal requirements below. I will also use my forecast of the feasibility of the replication (I may judge that it is too easy or too difficult: a possible outcome is that you replicate two papers, or only half a paper.)

- The methods must be aimed at uncovering a **causal** effect rather than a correlation. In this course, we do not care about correlations, nor about predictions.
- The data set must have at least **50** cross-section units. A time series analysis is not admissible as a term paper.
- You must convince me that the data is **available**.
- The methods used in the paper must be **more advanced** than OLS or a comparison of means. The panel data and instrumental variables methods discussed in the course are OK, and anything more advanced is also OK. Check with me if you are unsure.

If you decide to conduct your own research instead of doing a replication, a different set of requirements applies. Talk to me.

Topic

A great way to go is to find existing research that you are interested in, and then go from there. A replication is painful, and that pain will be reduced if you care about the subject matter. The paper you choose has to be empirical in nature, and needs to satisfy the constraints in the previous section. Where do you go and find replicable paper? The following list provides you some starting points:

- Look at the websites of high-quality economic journals that publish empirical papers and require the authors to make the data available
 - A great source of ideas is the “American Economic Journal: Applied”. This journal publishes high-quality papers that often use techniques related to the techniques in this course. Almost all papers have data available. See: <http://www.aeaweb.org/aej/app/index.php>.
 - A similar great source is American Economic Journal: Policy
 - As well as the American Economic Review
 - Journal of Applied Econometrics
 - Review of Economics and Statistics
 - etc.
- Look at the websites of prominent economists.
 - Daron Acemoglu’s website, see problem sets;
 - Angrist’s data archive at [harvard]
- Data repositories can provide a good starting point, for example:

- Harvard’s dataverse
 - For an overview, see this [overview] at the Open Access Directory (Economics is under Social Sciences)
 - Another good starting point is [here], although tracking down a paper that uses a given data set may be challenging
 - Microdata for development economics: [devecondata]
- Some popular books on economics may provide a useful starting point
 - Freakonomics, and Superfreakonomics (or on their Wikipedia pages, where the questions are summarized): e.g. relationship between abortion and crime; parenting and education;
 - Supercrunchers (by Ian Ayres): e.g. how to explain wine prices (suitable for ideas, not for replicable papers);
 - Poor Economics
 - “Mostly Harmless Econometrics” and “Mastering Metrics” (a lot of the data used there is available online, see e.g. Angrist’s website)
 - newspapers and magazines (The Economist, Significance magazine)
 - blogs;
 - an unresolved issue from a previous or concurrent course;

Format

What I care about is receiving a quality Markdown document with interesting content. There is no minimum/maximum length, but if you want some guidance: **5 pages**² excluding references, figures, and tables. Explain everything in your own words. Do not write down anything that you do not understand.

1. **introduction:** 0.5 page. State the question; convince me that it is important/interesting; **5 points**
2. **data description:** 1 page. Describe how you obtained the data; report descriptive statistics using tables and figures. Minimum is a table with means, standard deviations, minima, maxima, and number of observations for the variables used in the subsequent analysis, and one paragraph interpreting that table. Pictures are very welcome. **10 points**
3. **models and results:** 3 pages. Motivate and describe different methods used (state models, assumptions, estimators); interpret your estimates and relate them back to the original research question; attempt to make a strong case that some of your estimates reflect causal effects³. **30 points**
4. **conclusion:** 0.5 page. Summarize approach and findings. Discuss strengths and weaknesses. Suggest further research, or implications for policy..**5 points**

²This may not seem like a lot, but the data gathering, processing, and trying of several models, takes a lot of time and only results in a couple of tables.

³If you are critical: try to convince me even if you do not believe it. If it makes you feel better, you can destroy your argument in a subsequent paragraph.

You can earn additional points by extending the analysis in the paper in some way, if that analysis is accompanied by a motivation, and an interpretation of the findings. You will be penalized if your R code does not work. You will be penalized heavily if your analysis and discussion is not centered around causal effects.