Advanced Econometrics

Additional notes on Carrying Out an Empirical Project

Institute of Economic Studies
Faculty of Social Sciences
Charles University in Prague

Useful Reading

Carrying out an Empirical Project

- J.M. Wooldridge, Introductory Econometrics, South-Western College Publishing
- Chapter 19

Choosing a Topic

- Start with a general area or set of questions.
- Without stating a goal of the analysis, we can not even know where to begin.
- Make sure you are interested in the topic - very important motivationally.
- Use online services to investigate past work on the topic - IES offers most of the publishing houses access online, you can download the papers from the building or via Proxy Server: ezproxy.is.cuni.cz
- Past work is the start for you.
- Narrow the topic to a specific question to be investigated after the review of the papers.

Choosing a Topic cont.

- In the term papers/empirical projects, no-one expects you to develop a whole new theory.
- Even re-estimating a model on the new, updated data can bring quite a contribution.
- You can use Czech, regional data which are interesting.
- The aim of the empirical project assignments is to see that students are able to work with the data and manage complex problem of empirical estimation.
- Preparation of the term paper will help you a lot in organizing your ideas.
- Think about it as gaining additional skills which are valuable not only to academia - you will have to persuade also your boss with sound analysis!

Literature Review

- All papers should contain a short review of relevant literature.
- It is important to show how your work relates to the current literature.
- In a short term paper, literature review is included in the introduction.
- In your bachelors theses, you will probably want to devote a whole section to the review of the literature.
- Generally, it belongs to the motivation of your work, why is your work interesting.
- It is important to cite correctly! Always remember to give a credit to people who invented the model and clearly state their contribution.

Choosing the Data

- You need to think about what kind of data may help you answering your question.
- What kind of data you need to collect: cross-sectional, time series, panel data sets?
- This will be important for stating the model.
- It is helpful to look at other past studies, what kind of data they used, where they have the data from?
- The biggest problem: availability!
- There are several ways of obtaining the data, the first step is to check all possible online data bases.
- Authors are often willing to share the data sources if you ask politely.
- Even the data collection might be frustrating, do not give up! Interesting data will make your work contributive.

Using the Data

- Once you have data available, you need to put it to usable form.
- Most of the statistical softwares allow for raw text (ASCII) file.
- Still, other formats may be valid – check with software you use.
- It is crucial to understand the data organization before importing.
- It is also crucial to import data properly – mainly panel data.

Using the Data cont.

- It is extremely important to first look at the data.
- Even simply by opening the dataset, plotting, etc. one can learn lot of things about the data.
- It is important to find missing observations, errors, outliers, etc.
- Missing variables are often treated in multiple ways, as 0, or "N/A", or 999... All the ways will bias your results.
- By computing minimum, maximum, sample moments we can quickly find problems in the data.

Using the Data cont.

- You have to understand the nature of variables, are the data binary? What is the unit of measurement? Otherwise you may use wrong model, and/or misinterpret the results.
- Sometimes it is natural to transform the data, use logarithms if model requires it.
- Before transforming the data, it is good to make sure about ordering.
- Often data may be ordered for example from 2011 up to 1990, and if we compute the differences (changes) in an opposite way, the analysis will be useless (we need to reorder from 1990 to 2011).
- We can create variables important for analysis, like dummy variables, categorical variables, trends,...

Econometric Analysis

- You need to choose a model appropriate for the datasets.
- It is good to start with the model clearly based in theory.
- From the review of the past work, it is often clear what model is appropriate.
- Still, it is good to start with simple model (parsimony).
- When estimating a model, you have to be able to show that crucial assumptions are met by the data.
- Otherwise the analysis can not be trusted!
- If there is heteroskedasticity or serial correlation, correct for it.
- Describe clearly all the steps needed for estimation.
- Test significance of variables.

Econometric Analysis cont.

- It is often good to include sensitivity analysis:
- Estimate the original model and modify it in a reasonable way to show that results are not sensitive to changes and your conclusions are sound.
- If you have for example data for several countries and estimate some macroeconomic model, does all the countries behave in the similar way?
- Is there for example a country which would be totally different and change the results?
- It is often good to use more estimation techniques in order to show that conclusions are sound.
- This is mainly valid in the panel data analysis which needs mor careful estimation.

Writing an Empirical Paper

- A successful paper combines careful and convincing data analysis with good explanation.
- In other words, you have to put the interesting "story".
- Correct topic and data specification, correct analysis is only one step (although most important).
- Second very important step is presentation of the findings.
- Most of the researchers spent many years learning how to write a paper in convincing form, so do not be discouraged at the first attempts.
- This is a "learning by doing concept", the earlier you start the better for you!
- That is why I want you to write an empirical paper at the end of this course!

The Structure of the Empirical Paper

Structure most of the times follows:

- **Abstract** - a very short text describing what you do, why you do it and what is the main contribution. Remember, that title and abstract sells the work.

- **Introduction** - states the basic objectives, includes motivation why it is important to study the topic, literature review on what have been done and how it connects to your work, important questions you would like to answer, organization of the work.

- Introduction has to attract reader to read the whole work. Do not miss your chance!

- **Theoretical framework** - short description of the theoretical framework you are going to use, discussion of methods.

The Structure of the Empirical Paper cont.

- There is no need to write a full economic theory, just describe the main concepts.
- Use references, cite the work of others.
- **Econometric Model** - description of the model you are going to use, motivation for variables choice, description of estimation methods.
- You can provide reasons for using the estimation method, why to use GLS? Discussion about IV, etc.
- **Data description** - full description of the data used, simple summary statistics (min, max, mean, variance), number of observations.
- You should include data sources, units of measurement.

The Structure of the Empirical Paper cont.

- **Empirical Results** - should definitely include the estimates of the model.
- Show that the analysis is correctly done, specify and test all the hypotheses.
- Even when using simple OLS, you need to show that the estimator works! I.e. the residuals meet assumptions.
- Most important is to discuss interpretation after correct estimation.
- Does coefficients has the same signs/magnitude as you expected?
- When using more methods, compare carefully the differences and draw conclusions.
- **Conclusion** - summary of what have you studied and what have you found, can also include venues for further analysis.

Finally

- Remember that presentation is important.
- Take care about producing appealing figures, tables.
- Carefully describe everything you put into the tables, significance levels, p-values so the results will be clear.
- Think about appealing title which will attract everybody to read your analysis.
- Keep the analysis clear, logical, do not distract the reader.
- If there are any additional results which may support the analysis, put them to the Appendix.
- Be careful about plagiarism! It is serious problem of author's rights violation.
- Nowadays it is not a problem to find % of the text matches: everything is online.
- Be original, do not steal!