# ECON 762:  Information about the Project

L. Magee                                                                                             January, 2013

———————————————————————

## 1  Topics, Data and Project Proposal

- A wide range of topics is acceptable. It must involve some econometric estimation or related statistical procedure. Check for the availability of data early on.

- Some links to data sets can be accessed from the *Useful Links* on our Economics Department's web page.  DLI at McMaster may be of interest to those wishing to use publicly available Statistics Canada data sets.

- A proposal is due February 15.  This is one or two pages describing the research question, data sources, and econometric method that you anticipate using. You still can change your project topic after this if necessary.

## 2  Content of Research Paper

- Although advanced econometric techniques can be used, it is not necessary to do so. Sometimes OLS is sufficient.  If a well-known econometric method is used, it is not necessary to describe the theoretical econometric details (e.g. the variance formula) in the paper.

- The emphasis of the project can vary. Here are some possibilities.

    - A critical summary of the existing work on a topic, followed by your own econometric results.

    - A report based closely on research reported elsewhere.  Compare their results with yours and try to explain the difference. Your data may be from a different country, or cover a different time period, for example.  You might even use the same data set, but use a different econometric approach.

    - If a lot of effort or ingenuity was required to compile the data set, this could be the emphasis of the paper.

    - Rather than an empirical study, the paper could report a simulation study, for example one that compares the performance of different estimators or tests.

# 3    Format

You can get an idea of the usual format of an academic-style research paper by browsing through economics journals such as *The Review of Economics and Statistics*. It is usually 15 to 30 pages double-spaced. It can be longer if the extra length is to accommodate tables, graphs or appendices. The paper is divided into sections, often something like this.

**Abstract** A few sentences to summarize the research question and your main result.

**1 Introduction** The economic issue is stated in words. The motivation is given here for your own contribution. At the end of this section, briefly summarize the contents of the rest of the paper, section-by-section.

**2 The economic model and/or literature review** If there is an economic model that underlies the econometric model, it is presented here in more detail, both in words and in mathematical form if appropriate. Other people's work is referenced, their findings are noted, and your own approach is presented.

**3 Data and econometric approach** Describe the data, definition of variables, any data problems, how they were collected, etc. It is good to present summary statistics of the data, e.g. plots or tables of time series, means, st.devs, mins and maxes of variables in large data sets. The econometric approach, along with any estimation and model specification issues, are discussed.

**4 Results** Coefficient estimates, specification tests, etc., are presented and interpreted in light of the issues mentioned in the introduction.

**5 Conclusion** A summary of what you did and the major findings. You could also suggest directions for further research coming from your work.

**References** A list of the papers and books that you have referred to in the paper, ordered alphabetically by first author's last name.

**Tables, Figures and Appendices** If there are any, they can be put at the end. Appendices are handy if you have some background information or results that might be of interest to some readers, but are not central enough to belong in the main part of the paper.

# 4    Other Comments

These comments are grouped according to the sections suggested above.

Abstract, Introduction and Conclusion

Try to make these sections as clear and informative as possible. In many situations, readers will draw their main impressions about a paper from these sections, and may not look very closely at the other sections.

The economic model and/or literature review

- When reviewing other related papers, organize your comments around a small number of questions or issues. Do not describe aspects of other papers are irrelevant for your own work. One way to check for other work on your topic is to search using Google Scholar.

- Motivate the econometric study that will be reported later on. State a question or questions you wish to answer using the econometric estimates. The results given later should then be of interest, even if the estimates of key coefficients are not statistically significant.

- Throughout, be clear about which ideas or results are yours and which are taken from other research, giving references for the latter. When presenting mathematical results taken from other work, indicate where it has come from and include the reference.

Data and econometric approach

- When working with time series variables measured in currency units, make sure the units are consistent across variables unless there is a special reason to do otherwise. For example, it can be confusing to measure some variables in constant dollars and others in current dollars. Often it is a good idea to measure everything in the same currency units, e.g. constant 1991 Canadian dollars.

- When constructing price or output indexes that require splicing together two or more separate time series from different sources, make sure they have the same base year, or adjust one of the series if necessary. As long as the time periods of the series overlap, this can be done. More generally, if data are missing for part of a series, it sometimes is acceptable to fill in the missing data. See me if you have questions about this.

- Do not <u>test</u> for multicollinearity (which is the presence of high correlations among regressors which can result in the near-singularity of $X'X$). Multicollinearity is the same sort of problem as is having a small sample size – it is a lack of information. It is not like autocorrelated errors, for example.

- Explain why each explanatory variable is used. This often includes a brief discussion of what sign is expected for each coefficient.

- With time series data, it is often informative to show plots of each variable against time.

- It is not necessary to explain the formulas or econometric theory underlying the estimation technique that you use. It may be interesting, though, to discuss why you chose one estimation approach over another, or to compare the results from two approaches.

Results

- Try to make sense of the results. Avoid simply reporting the numerical results in the text. For example, avoid using a lot of sentences like "The coefficient estimate for the variable $X1$ was 0.62". If a coefficient has an unexpected sign, suggest a possible reason, especially if it is an important coefficient. It is common to obtain coefficient estimates that have the opposite sign to what you expected. If you find yourself in this situation, you may be interested in the paper "Oh No! I Got the Wrong Sign! What Should I Do?" by Peter Kennedy at *http://www.sfu.ca/economics/research/discussion/dp02-3.pdf.*)

- State what econometric strategy, estimation method, and software was used.

- In models with several explanatory variables, the coefficient estimates must be interpreted while keeping in mind that other variables also are being controlled for. For example, if a variable $x_i^2$ is included as a regressor along with $x_i$, the coefficient on $x_i$ cannot be interpreted in isolation from the coefficient on $x_i^2$.

- Give precise variable descriptions for all variables used, and state where you got the variables. Often this information is put in a table in an appendix.

- For linear regressions, it is not necessary to write out the equation variable-by-variable in the text. It is sufficient to give a table of coefficient estimates and $t$-statistics, standard errors, or $P$-values. From the table it is clear what the model is as long as you state the type of model (e.g. OLS, probit, etc.) and the name of the dependent variable, and all of the independent variables are listed. Sometimes there are too many independent variables to list them all in the table. The omitted ones can be listed in a footnote.

- Usually it is valuable to include results from more than one specification (e.g. more than one set of regressor variables), to give the reader an idea of how sensitive the results are to the way the model is specified. This kind of specification or sensitivity analysis can be applied in other directions too. For example, see if the results change when the time period covered by the data set is changed, or if only people living in a certain region or with certain characteristics are included.

- If a variable has a coefficient estimate that is very significant, e.g. a $t$-stat of $\pm 5$ or more, the model may be improved by including other variables that are nonlinear functions of it. For example, if a continuous variable is very significant, try including its squared value, its cubed

4

value, and so on, or try regression splines. If a dummy variable is very significant, create new variables that interact it with (multiply it into) other key regressors. In effect, this allows the coefficient of these other regressors to depend on whether the dummy variable equals 0 or 1.

References

Before handing in the paper, check to ensure that each article or book that you refer to in the paper is included in the list of references at the end.

Tables and Figures

- Make tables and graphs as understandable as possible on their own. A casual reader may not want to have to carefully read the paper in order to understand the main points of the tables and figures.

- In tables of results, give at least three significant digits. More than four digits usually is not necessary.

- Scaling variables beforehand can prevent estimates from being awkwardly small or large. If necessary, use scientific notation, e.g., 1.36E-04 or 1.3610-4 instead of .0001. Really small or big coefficients can be avoided by re-scaling the independent variables. For example, defining income in thousands of dollars instead of dollars will make the coefficient on income one thousand times larger.

## 4.1 Things to check before handing it in

- Proofread the paper for spelling and grammar.

- Number the pages, and number the sections.

- Make and keep a copy of the paper before handing it in. I prefer that you hand in a paper copy, but if that is not convenient, you can send it electronically.

## 4.2 Marking Criteria

Considerations for grading the paper include: literature review and presentation, appropriateness of econometric technique, interpretation of results, connection of your results to the existing results and/or economic theory, effort required for data collection, and exposition. No fixed weights are pre-assigned to these categories.