Term paper guidelines

Nicholas Rivers

Advanced Quantitative Research Methods for Public Policy, API6319A
Fall 2019

1 Overview

Your term paper will involve the preparation of a short empirical analysis on an issue related to public and international affairs of your choosing. You may choose a topic that is related to your MRP (if you know the topic of your MRP) so that you can use some of the material you produce in this course in your MRP.

In this paper, you are required to conduct *original* empirical analysis using the methods that we cover in class. The paper should be *roughly* 3,000-5,000 words, and should also include properly formatted figures and tables to support your conclusions. It should also include some reference to the literature, to provide context for your research. In the following sections, I provide some guidance about choosing a topic, the contents of your term paper, and writing your term paper.

2 Choosing a research question

In quantitative studies, a research question is typically cast as an inquiry into the relationship between two variables, or else as a hypothesis (to be tested against the data) about the expected relationship between two variables. As an example of these two ways of framing your research question, consider a study that seeks to evaluate the impact of the Quebec $7/day childcare program:

Research question What has been the impact of the Quebec childcare program on labour force participation rates of mothers of young children?

Null hypothesis The Quebec childcare program has had no impact on the labour force participation rates of young mothers.

I suggest casting your research question explicitly in one of these two forms in your paper.[1] Typically, the more focused/specific your research question, the more likely you will be able to answer it in a meaningful way. Ensuring focus in your research questions is difficult, and is one of the skills I hope that you will gain during this course.

A strong research question is one that is simultaneously:

interesting you should aim to choose a research question for which the answer is not known. When you tell a colleague about your research paper, their interest should be piqued.

important you should aim to address a research question that has important real-world consequences.

feasible you should be able to accomplish the research within the confines of your available time and skills. You should also be able to find the data required to conduct the analysis, and be able to think of a suitable research design for answering the question.

[1] Creswell (2009), p. 129-137 provides a very good introduction to the framing of quantitative research questions.

Choosing such a research question is an art, and it is difficult to succeed the first time round. As an example, it is relatively straightforward to come up with a research question that is feasible, but has limited importance. For example, it would be relatively straightforward to use Census data to answer a research question that asked: "is there a relationship between home ownership and age?" However, it is not clear that this is an interesting or important question (although you may be able to cast it in such a light). Conversely, many questions are interesting and important but difficult to answer: for example asking "Does foreign aid increase country GDP growth?" is difficult to answer because foreign aid is purposefully directed towards certain countries who may have characteristics that also determine GDP growth (this is and example of a "confounding variable"). Your goal in choosing a research question is to navigate towards a questions that is simultaneously interesting (we don't know the answer, but would like to), important (the answer is important), and feasible (you can find the answer).

In quantitative research, the focus is on *positive* research questions as opposed to *normative* research questions. That is, your research questions should ask "does X affect Y" or "how much does changing X cause changes in Y", rather than "should we change X?". So a question such as "Is Canada's foreign policy in the Middle East appropriate?" is not the kind that we can answer in this class.

Some investment of time at this stage is warranted, since failure to choose a compelling research topic makes it difficult to produce a strong paper. I would be very happy to provide some feedback on your choice of research question, and I will try to ensure that there is some class time devoted to "workshopping" research questions. In addition, see the sources below (particularly Booth et al.) and this website amongst many other web guides.

3 Data

For the purposes of the term paper that you will produce for this course, I suggest that you let your selection of research topic be guided substantially by what data is readily available to you, since you won't have time to spend a great deal of effort in sourcing data. Fortunately, there is an enormous range of data easily available for download. Here I provide a short and very incomplete annotated list of potential data sources to get you started. Feel free to come to talk to me about additional data sources if you are having trouble finding something that fits a question you are interested in. In general, I strongly recommend that your data set be a cross-sectional data set, since working with time-series data poses many statistical problems that require advanced skills. If possible, I also recommend that you avoid trying to infer cause-and-effect from country-level data sets (as we will discuss in class), because it is difficult to control for confounding variables completely.

- Domestic/public policy:
  - The library maintains a list of data sources that are nicely cataloged by keyword and topic. Most of these data sources are micro data (that is, they are at the individual level, rather than the country or province level) and the data cover both surveys conducted by Statistics Canada, as well as opinion polls conducted by Gallup. A few (of many thousands) interesting surveys available here include:
    - The General Social Survey is conducted annually, and each year asks topical questions, ranging from time use, to social and community networks, ageing, victimization, risk, etc.
    - The Canadian Census of Population is conducted every five years and includes information on ethnicity, age, birthplace, family size, dwelling size and type, employment, education, income, employment, pension, language, religion, and other information.
  - Statistics Canada's CANSIM database contains mostly time-series data on many economic and social indicators of Canadian well-being.
- International affairs/policy:

- The OECD (rich-country think tank) collects data on a number of primarily economic variables.
- The World Bank data set includes country-by-country observations on over 8,000 variables, including things like development indicators, landmine casualties, education, demographics, health, and migration (among many others).
- The World Health Organization has data on child nutrition, immunization, disease, life expectancy, HIV/AIDS, maternal health, etc.
- The UN Environment Programme (UNEP) data base contains information on carbon dioxide emissions, energy consumption, land use, pesticide use, protected areas, etc.
- The main UN data base has an enormous amount of data on all sorts of things: fertility, energy consumption, population, official development assistance (foreign aid), information and communication technology, and all sorts of other things. Plus, the database provides links to other main data sources.

You may want or need to use data from more than one data source in your paper. For example, you might want to use life expectancy data from the World Health Organization in conjunction with environmental quality data from the UN Environment Program, to estimate the impact of reduced environmental quality on health. In this type of case, you will need to merge the two data sets. This can be accomplished using the `inner_join()` tidyverse function. I can provide additional help if you encounter this situation.

When choosing your research topic and data, you should generally ensure that your data set is appropriate for the statistical methods we cover during class. Appropriate data sets are:

- Cross section or panel data. We did not cover time series methods in any depth in the class, and analyzing time series data appropriately requires some techniques with which you are not familiar. As a result, you should normally aim for cross-sectional data.
- Big enough. The statistical methods that we cover are aimed at summarizing a data set with a statistic, such as the mean. If the data set is very small, your estimates of these statistics will be very imprecise, and it will be difficult to draw conclusions from your analysis. As a rule of thumb, you should aim for a data set with at least 1000 observations. Bigger data sets generally allow you to draw more precise conclusions.

4 What should be in the paper

Your paper should include at least the following:

- At least one well-formatted and properly labeled figure that you produce from your data, to demonstrate that you have learned how to produce a compelling figure and to use the figure to illustrate a point in your paper. Your paper should not only include the figure, but should use the figure to make a point (i.e., don't just throw the figure into your paper, but actively refer to it).
- At least one table summarizing your data in a useful way (as above, your paper should refer to the table).
- At least one statistical test of a hypothesis. The results of this statistical test should be referred to in the paper, and the statistical test should help to inform the main conclusions that you draw from the paper.
- The code you used to conduct your quantitative analysis (in an appendix).
- Some reference to the literature to contextualize your research (I expect you to cite at least 3 or 4 studies that are comparable to the work that you have done).

Note that these are *minimums*. A convincing empirical paper will typically require much more evidence than above, and you should seek to provide a convincing paper.

5 Writing the paper

Unlike a qualitative paper, a quantitative paper typically follows a quite specific template. There are many guides to writing an empirical paper (see below for a smattering). I summarize what I think should be the main sections of your paper, as well as what should be included in each section, here:

Introduction Your introduction should provide the context for your problem, should provide a brief summary of the available literature (I expect you to cite at least 3 or 4 relevant studies here), should state the research question or hypothesis clearly, should summarize the main findings of the paper, and should inform the reader of what is coming next. If you are having trouble getting started, think of the introduction as just five or six paragraphs, each of which should contain a distinct idea (you may want to expand any of these into more than one paragraph). A 'recipe' for writing an introduction is:

1st paragraph Very general introduction; should grab the reader's interest. Tell them why the paper is important. This paragraph should tell the reader what your research question is in precise terms.

2nd paragraph A brief summary of the current state of knowledge. This is where you make reference to the existing literature.

3rd-5th paragraphs This is where you say what your study does, how it does it, and what you find.. It can be really straightforward: "In this study, I use data on ... to estimate the relationship between ... and .... I find ...". Here you should also say how your study fits into the previous literature.

6th paragraph This tells the reader what is coming up in the rest of the paper. "The following section explains the sources of data that were used to conduct the analysis. Section 3 describes the methodology, ..."

Methods and data Here, you should be very clear about exactly what you did, with the aim of making your research reproducible. You should explain exactly what data source(s) you used, how you transformed any variables, etc. You should also explain what methods you chose to conduct your analysis, and what assumptions underlie those methods. Be up front about any shortcomings in your method, and do your best to address them.

Results Your results section should make an overall claim (e.g., "Female labour supply increased by 2% as a result of the implementation of the universal subsidized childcare program.") It should be supported by the empirical evidence that you bring to bear in your analysis. In the results section, you should also provide a descriptive summary of the data set that you will be using (typically prior to conducting statistical analysis).

Conclusion Think about how your reader's mind should be changed by reading your research: in this section, remind them of this, and tell them what your work implies, especially for the broader public policy/international affairs context. In addition, tell the reader about any deficiencies in your research: were there data problems you encountered? Assumptions you required? What would be useful research to be undertaken in the future, given your findings?

Appendix In the appendix, you should include a printout of the R `.R` file that you used to conduct your analysis, as well as any other notes related to your data that you think could be relevant.

In the approximately 10-15 page article that you will be writing, I suggest the following (rough) breakdown for the article: Introduction: 2-3 pages; Methods and data - 2-3 pages; Results: 4-6 pages; Conclusion: 2 pages. I strongly suggest that you use a writing guide to produce your term paper (see below).

6 Useful resources

There are a number of resources that you might find useful in producing your research paper, in addition to the Statistics textbook and course notes which will support your quantitative analysis. A sampling follows:

- Weimer, D. and Vining, A., 2010, Policy analysis: concepts and practice, Longman, 5th edition.
  - This is a graduate textbook in policy analysis that is an excellent resource. Chapters 1 and 9 provide very well written examples of policy analyses, and chapters 14 and 15 provide an excellent guide to developing and writing your own policy analysis. If you don't have this book already, you might find it useful to take it out of the library and skim these chapters.
- Wooldridge, J.M, 2009, Introductory Econometrics: A Modern Approach, Southwestern, 4th edition.
  - This is an undergraduate text for economists, and covers some of the material we covered in this class (and much more). Chapter 19 is especially useful, as it focuses on carrying out an empirical project. This book should be in the library.
- Booth, W.C., Colomb, G.G, and Williams, J.M., 2008, The craft of research, University of Chicago Press, 3rd edition.
  - This text focuses on choosing and conducting a research project. It contains special focus on the writing process. Chapters 3-5 present a useful discussion on choosing and narrowing down a research topic.
- Creswell, J.W., 2009, Research Design: Qualitative, Quantitative, and Mixed Method Approaches, Sage Publishers, 3rd edition.
  - This is a general on research design appropriate for graduate students, and offers a very useful chapter (ch. 7) on framing quantitative research.
- Papers by public policy think tanks
  - A number of think tanks produce the types of research papers that are appropriate for this class. Some Canadian examples include the CD Howe Institute, the Institute for Research in Public Policy, and the MacDonald-Laurier institute (among many others). Other prominent examples include the Brookings Institute, the World Bank, and many others.
- Zuercher, C., Notten, G., Vucetic, S., 2011, Guidelines for writing papers in GSPIA, University of Ottawa.
  - On the course website.
- Research as a second language (blog)
  - A really useful blog on the writing process. See especially the '40-paragraphs' post linked above.

7 Potential research topics

Here, I provide a few research ideas to get you started, which I think pass the test of being interesting, important, and feasible. Note, however, that these are just a few ideas to get you started. Please come to see me about developing your own research topic.

- Do immigrants to Canada receive lower incomes than native-born Canadians?
- Does being bilingual in Canada reduce the likelihood of unemployment?

- Do societies with lower levels of income inequality have higher or lower rates of crime?
- Does introduction of water metering promote water conservation?
- Does a long commute reduce people's well-being?
- Does air pollution cause people to miss work?
- Does a higher minimum wage increase youth unemployment?

If you're stuck, one potential source of ideas is in the published literature. For this course, it would be acceptable to try to replicate findings from a prior study, using a different data set. For example, you may find a US study that relates income to education levels, and try to use the same approach to understanding the relationship using Canadian data.