

LAB RESEARCH PROJECT INSTALLMENT 2: PROJECT PROPOSAL

Each group should turn in one collaboratively produced project proposal. The entire proposal should be written as one integrated, coherent, and consistently formatted document. It should consist of complete sentences assembled into paragraphs that are in turn organized in a logical way.

Your proposal should include (i) a brief description of the topic or issue you will be investigating, along with a more specific discussion of the particular questions or aspects of your topic that you plan to focus on, (ii) a review of previous research related to your topic, (iii) a description of the source or sources from which you plan to obtain your data, with an indication of which specific variables you plan to use from each source, (iv) a hypothetical spreadsheet illustrating the way you will organize your data for analysis, (v) two or three hypothetical tables, graphs or other figures that could be constructed from the data in your hypothetical spreadsheet and that might provide some insight into the questions you are investigating, and (vi) a conclusion. More details on each of these components are provided below.

Brief summary and more specific discussion. The brief description of the issue will typically consist of about three to five paragraphs in which you present your topic concretely and directly, with a minimum of jargon and no extraneous filler.

In the more specific discussion, you should describe the particular focus you plan to take on the general topic and/or specific questions you plan to investigate. You should begin formulating what you believe the “hook” of the project will be: What is the motivation for the project? Will your findings about the particular questions you plan to investigate have broader implications for an ongoing empirical controversy? For a particular contemporary public policy issue? For a theoretical issue debated by social scientists? Do you have particular questions about the general topic that you think are specially compelling, or that build upon previous research in interesting ways? Finding the particular focus or motivation for your project—the “hook”—is critical.

Previous research. For this proposal, you should find roughly six to eight previous studies to discuss in your literature review. Seeking out relevant previous studies is a task that you should work on all semester, and in future installments you will be asked to integrate discussion of more of the previous research into what you write. But for this installment, six to eight good sources should be appropriate. Remember, though, that to find six to eight good sources, you will probably need to read and consider more than six to eight studies.

For the most part, the sources you discuss should be based on original research of professional quality, conducted by an individual or group of individuals with appropriate expertise. In most cases, appropriate studies are found in academic journals, books, book chapters, or reports issued by government departments, multilateral agencies, or non-profit organizations. In some cases it may be appropriate to include items from the popular press or

non-academic web sites in your literature review, but remember that most of your sources should be original professional research.

Students (and professional researchers as well) often think of a literature review as a required but not very important or useful component of a research project—something that gets pasted into the paper somewhere, but that doesn't matter much. That view is incorrect. ***In fact, understanding the questions previous researchers have asked about your topic and what has been learned in previous studies is almost always the best way to focus your own thinking about what you want to do in your project, and to develop an effective “hook” to motivate your research.***

This point cannot be stressed too much, so I will elaborate: As you craft a topic and develop the ideas that will make up the hook for it, the most important thing you can do is to read a range of primary research on the topic, think about the questions the various studies have looked into, what the motivations for those studies were, and how the results of different studies are similar to or different from each other. Think concretely about the studies previous researchers have conducted, how they posed their questions and interpreted their results, and how the previous studies relate to one another—e.g., do they add up to a nice coherent set of findings about a topic? Are there major contradictions among them? Can you identify different “schools of thought” into which most previous studies can be categorized? Are there certain factors that almost all studies take into account? Are there any factors that have been left out of most studies but that seem important to you? As you read and reflect on previous research, what ideas do you have about further aspects of the topic that would be interesting to investigate?

To rephrase more briefly what the previous paragraph says: The best way to develop a good research topic of your own is to read and think carefully about previous studies on related topics. Sources like newspaper articles, magazines and websites can be helpful, but they are not a substitute for careful study of previously published research. Just trying to think up ideas yourself or brainstorming with your group can also be useful, but will not get you very far unless you also learn about the previous research.

In class I described in detail a process for working with the other members in your group to survey previous literature, discuss it with one another, with the hope that out of those discussions will grow the particular question or hook you focus your project on. And as I described in class, you should be reviewing literature and discussing it on an ongoing basis throughout the semester—this is something you can devote some time to during each of your group's weekly meetings. Doing this thoughtfully over a period of many weeks is the best way to develop a good hook for your project.

The next three paragraphs give are a few more specific guidelines for your literature review.

First, it is almost always better to organize your literature review by issue or topic than by source. Think about the issues raised by the variety of sources you have read, and organize those issues into a logically structured essay. Your references to specific previous sources should then be embedded in this structure: in the discussion of each issue you choose to highlight, mention

(using logically organized sentences and paragraphs rather than just a list) the sources you have found that are relevant to the specific topic, and explain what is interesting and relevant about those previous works (e.g., what their main conclusions were, what statistical or analytical methods were used, the data used in the previous studies, etc.) Simply discussing in sequence one article or book at a time is rarely a good way to organize a literature review.

Second, if you mention a previous study in your literature review, you should say enough about it to let the reader know why you have included it and what its significance is. At a bare minimum, that usually means stating what the questions addressed in the research were, as well as what the main findings were. Just saying that certain researchers did a study on a certain topic, without saying what they learned, usually does not add much. If you don't have more than that to say about a previous study, you should probably not include it in your essay.

Third, in addition to saying something substantive about the questions asked and the main results of the studies you discuss in your literature review, you should say something specific about how your awareness of that study will affect your project: what did you learn from the study that will help you define and/or investigate your own topic. Simply noting that a previous study is on a topic that is the same as or similar to the one you plan to work on is not sufficient.

Data. In the data section of your proposal, you should identify the source or sources of the data you plan use to for your project. For standard or prominent data sources that you can reasonably expect your reader to be familiar with (e.g., the World Bank's *World Development Indicators Online*, or the US census), simply naming the source might be sufficient. But for less well-known sources, you should give a general description of the data and the source from which they were obtained.

In the case of less well-known data sources, deciding what your description of the data source should consist of requires some thought and judgment; there is no general formula or checklist to follow. A general (though not very precise) principle is that you should discuss anything about the data that a reader would need to know to understand what you plan to do for your study and how you will interpret the data. Examples of the kinds of things you might discuss include: who collected the data, whether the data were collected as part of an ongoing project that collects similar data on a regular (e.g., annual) basis, or how the population from which the data were collected was defined. Was the data collected in a survey? From government statistics? From observations of behavior? News reports? Coding of legal documents? Imagine that you want to tell someone about the dataset, and then think about what you would need to tell that person about it to give her/him a clear and concrete idea of what it consists of.

For every data source you plan to use, one essential piece of information is what the unit of observation is (i.e., what kind of individual or item does each row of the spreadsheet represent?). Be sure to state clearly what the unit of observation is for every data source you mention.

For every data source you mention, you should also indicate specifically which variables you believe will be useful for your project. You should specify the particular variables you plan

to use in enough detail to allow another researcher (or another student in this class) to go to the original source of the data and find exactly the variables you are referring to. The ideal thing is simply to give the names of the variables as they appear in the original source. If for some reason that is not possible, find some other way to describe exactly which variables you have in mind.

To the greatest extent possible, you should use data obtained from “primary sources.” That is, you should obtain your data from the source or sources that originally collected them, in the format in which they were originally stored. For instance, you might find that a certain think-tank has used data collected by the US Bureau of Justice Statistics to create an Excel file with data on rates of violent and non-violent crimes in the 50 states of the US and Washington, DC, from 1984 through 2009. Even if this Excel file is conveniently posted on the think-tank’s website, you should not use that version of the data. Instead, track down the original data in the format in which they were first reported by the Bureau of Justice Statistics, and then extract the parts of the original data files that you will need for your project. (We will be seeing how to do this during lab over the next few weeks.)

For each data source, you should describe the kind of documentation that is available for it. Something called a “codebook”? or a “Users’ Guide”? or a web interface that lets you search for variables and find their definitions? How can you find out what variables are contained in the data set and what the unit of observation is? Is any other special documentation available or necessary for understanding what is in the data file?

Remember that you should be able to gather all of the information about your data and data sources that you need for this proposal without actually looking at the contents of any of the data files. You should be able to get the general information you need about the study that collected the data, as well as specific information about the variables contained in the dataset, by studying the documentation that accompanies the data files.

Instructions continued next page.

Hypothetical spreadsheet. Construct a hypothetical spreadsheet that illustrates the structure of the dataset that you hope to create (and then analyze) for your project. You may construct this spreadsheet either by drawing it by hand, or by creating an Excel spreadsheet and pasting it in to the document you print. Here is an example of a hypothetical spreadsheet, with examples of some appropriate explanatory comments:

Country	Year	GDP_PC	Openness	Procedures
Argentina	2004	#	#	#
...
Argentina	2009	#	#	#
Belgium	2004	#	#	#
...
Belgium	2009	#	#	#
...
Zimbabwe	2004	#	#	#
...
Zimbabwe	2009	#	#	#

Our unit of analysis will be country/years. Our data represent over 108 countries, for the years 2004-2009. The three variables we are interested in are *GDP_PC*, *Openness*, and *Procedures*.

GDP_PC and *Openness* are taken from Heston, Summers and Aten (2011).

GDP_PC is per capita GDP, measured in international purchasing power parity dollars, not adjusted for inflation. In the Heston, Summers and Aten dataset, this variable is called "rgdpch."

Openness is defined as $[(\text{dollar value of exports} + \text{dollar value of imports}) \div (\text{dollar value of GDP})] \times 100$. In the Heston, Summers and Aten dataset, this variable is called "openk."

Procedures is taken from World Bank/International Finance Corporation (2012). It is defined as the average number of legal procedures required to resolve a contract dispute, as found in surveys of businesses conducted annually in all the sample countries. This variable also goes by the name *Procedures* in the original data set; it is included in a category of measurements grouped under the heading "Enforcing Contracts." Djankov, La Porta, Lopez de Silanes and Shleifer (2003) provide details on the method used to construct this measure...

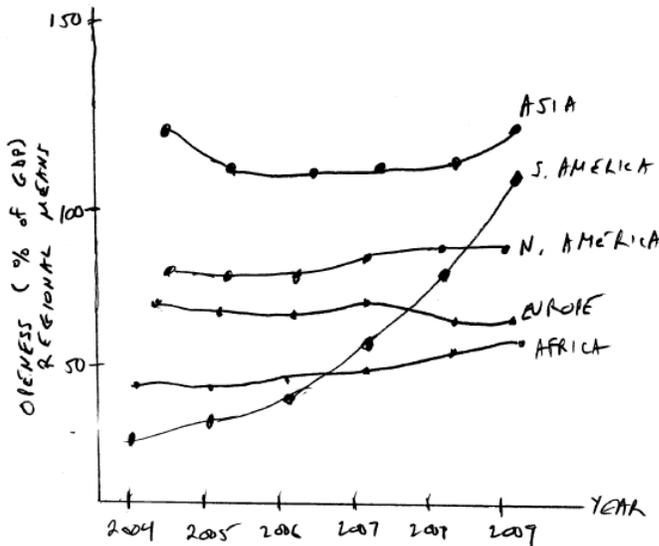
As illustrated in the above example, you should explain what the unit of analysis will be in the dataset you construct, what variables will be included, how each variable is defined, and the source from which each variable will be obtained.

Instructions continued next page.

Hypothetical graphs or tables. Think of several hypothetical graphs or tables that you could construct with the data in your hypothetical spreadsheet that would provide some preliminary information on the topic you have chosen or some preliminary answers to the questions you have posed. For each of the figures you think of, describe what the figure would show, and describe what you could learn from it related to your project—i.e., what are some of the possible patterns you could find in the figure, and what would the different possible patterns imply about the issues you are investigating? Then draw by hand a sketch of what each figure might look like.

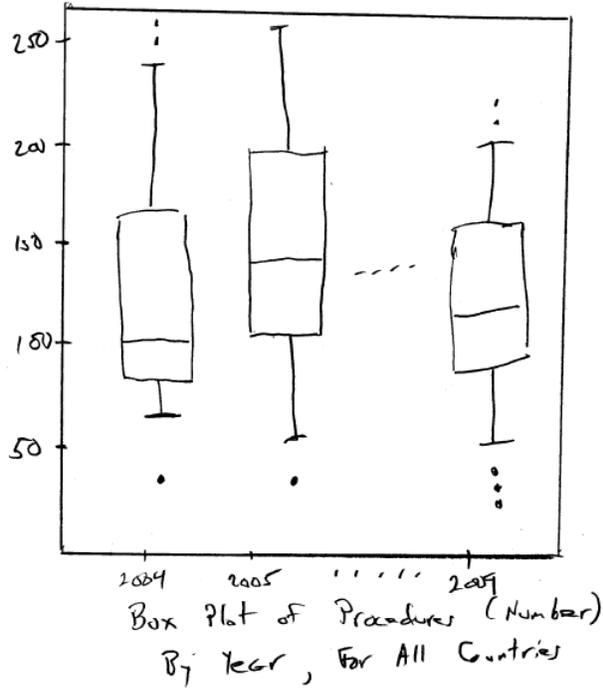
For example, the following hypothetical graphs could be constructed from the data in the hypothetical spreadsheet discussed above:

One figure we are interested in would be a scatter plot showing 5 different time series. On the horizontal axis we would have years (from 2004-2009), and against that we would plot the average value of *Openess* for each year among countries located in each of 5 continents, Africa, Asia, Europe, North America, South America. For each series of points (i.e., for the series of points representing any continent), the vertical height of the points would represent the average value of *Openess* for all countries in the continent, in each year.



Instructions continued next page.

Another figure we are interested in is side-by-side box plots of all values of *Procedures* for every year. Each box plot would represent the values of *Procedures* for all the countries in the sample, for a given year.



Conclusion. Conclude your paper with a few paragraphs that summarize the most important points you have raised in the proposal and a forecast of how you plan to proceed.

Turning in this assignment: This installment is due on Friday, March 6. You should bring a printed copy of your proposal to turn in at the beginning of class that day.