

How to Write up a Quantitative Analysis

Prof. Jonas Bunte
Institute for International Political Economy
Vienna University of Economics and Business

In this section, you describe how you plan to test whether your proposed explanation has any empirical support. There are several steps to writing this section. First, you need to justify why your methodological approach is the appropriate approach for testing your hypothesis and comparing its explanatory strength to that of other explanations. Second, you then need to describe how you intend to implement this approach. Third, you need to present the findings.

Justify your Choice of Methodological Approach There are several basic types of research designs that might be appropriate to test your hypothesis and comparing its explanatory strength to that of other explanations. Examples include Large-n statistical analysis, Qualitative case studies, Game-theoretic models, and experimental studies, among others. In the first step, you should justify *why* a large-n statistical analyses is the most appropriate to test your hypothesis. Briefly discuss the various strengths and weaknesses of alternative methodological approaches, and explain why for your particular topic a statistical analysis is best suited.

Introduce your Research Design In a next step, you should explain how you plan to test your argument. Specifically, you must tell the reader your plan for conducting the statistical analysis. This section should include information on the following:

- **Sample:** Identify the sample that you will use for your statistical analysis. What is your unit of observation (countries? individuals? regions? firms? etc.). Second, which units will be included in your dataset (all countries in the world? only African countries? Only democracies? etc.). If you plan to exclude some units explain why they are excluded and how this exclusion might affect the results. Third, identify the time period under consideration. If you will examine only one year (month, day, decade, etc.), justify why you focus on this particular time period. If you will examine a longer time period (e.g., 1970-2010) so that your data has repeated measures of your variables, justify why this is the appropriate time period for your analysis.
- **Operationalization of dependent variable:** Discuss the measurement of your dependent variable (i.e., your outcome variable). Operationalization involves identification of empirical referents that measure the typically abstract concepts specified in your hypotheses. Discuss how the operational indicators of the abstract concept of your outcome are valid (measure the true variable as closely as possible) and reliable (yield consistent values across a series of observations). Subsequently describe the source of the variable: Will you have to compile the data yourself, or does it exist in some format already and is publicly available (if so, what is the source)?
- **Operationalization of independent variables:** Discuss the operationalization of your independent variables (that is, one variable for your hypothesis, but also

variables capturing the alternative explanations you identified in the literature review). Examine how the measure of your abstract concepts are valid (measure the true variable as closely as possible) and reliable (yield consistent values across a series of observations). Furthermore, discuss the sources of these variables.

- **Relevant control variables:** It is possible (and indeed likely) that variables other than your independent variables could also matter for the variation in your dependent variable. We need to account for these alternative factors while examining the explanatory strength of your independent variables. For this reason, you will need to include control variables in your analysis. In this section, list all control variables that you will include as well as their sources. However, no discussion of their operationalization is necessary.
- **Method:** Discuss the statistical methodology that you will use to analyze your data. Many different methods are possible. The appropriate method for your analysis depends on the type of your dependent variable (continuous variable? categorical variable? count? duration? etc.), the type of data (single cross-section? pooled cross-section? time-series? etc.), and the type of question you want to answer. Whichever method you choose, justify why it is the appropriate methodological approach for your research question. Following the justification, write up the full equation of your model. Using this equation, explain how the model will identify whether your argument receives empirical support.

Summary of analyses conducted After conducting your analyses, you should present and interpret your estimation results. This involves several steps:

- Explain how the coefficients estimated for each of your independent variables help you to differentiate which of the three hypotheses obtains most support?
- In this process, ensure that you accurately interpret p-values and/or confidence intervals
- Also, accurately interpret a null finding. Specifically, is it evidence that the true effect equals 0? (see Jeff Gill, 1999)
- Do you use causal language to describe a correlational finding? Do you specify the assumptions necessary to interpret their findings as causal?
- If you included interaction terms, make sure that you include all constituent terms. Moreover, to properly interpret the substantive effect, did you graph the marginal effect across the range of values of the interacting variable to test differences in relationships between subsamples (see Brambor, Clark, and Golder, 2006)?
- Do the control variables have the expected sign?
- What about model fit?
- In communicating your estimation results use graphs rather than tables and text. Avoid vagueness such as “this figure shows my main result.” Say something concrete, like, “I found that democratic countries are actually more likely to borrow from China than autocratic ones, as this graphs shows that ...”

In addition, you should conduct and report several robustness tests. The precise type and implementation of these robustness tests depend on your question, hypotheses, and model, thus it is difficult to give specific advice. However, as a general rule, your robustness tests should address the most obvious concerns that reviewers might have after reading your main results. Common robustness tests are the following:

- Variable definitions: Use different operationalization for your main hypotheses. Do the results still hold if you define your independent variable differently? If so, then you have shown that your findings do not depend on the specific variable you used.
- Control variables: Use a different set of control variables. Your original hypotheses might have excluded several variables that could also affect your dependent variable. Show that your findings are robust to changes in the set of control variables you used.
- Examine subpopulations: Split the sample to show that your theoretical argument operates in different circumstances. For example, you might want to show that the mechanism applies to both Democracies and Autocracies, etc.
- Include interaction effects: The sign or strength of your main independent variable x may depend on the value of a second variable z (i.e., if z is high, then x has a strong effect on y ; if z is low, then x has a weak effect on y). Theoretically anticipate these moderating forces, estimate an interaction model, and interpret it.