

Political Science 501, Baumgartner
Short paper topics/discussion points
Week 4: Sampling

Readings: Nachmias, Ch. 8, 10, NES codebook introduction, lobby project sampling discussion, Hojnacki and Kimball

Review the following and be prepared to discuss:

Populations, sampling frames and the problems of establishing good sampling frames in some circumstances. How would you establish a sampling frame for the study of wars, for example? Also note where sampling frames are automatic and no problem, as in studies of fixed and known populations.

What are various types of non-probability sampling designs and why would one ever do those? What is a convenience sample? In Prothro and Grigg's article from last week, were Tallahassee and Ann Arbor strongly justified sampling decisions, or is it reasonable that all work has to be done somewhere, so the question is how well you sample within the parameters that you set for your research?

Do you trust quota samples? What would increase your trust?

Any good reasons in particular circumstances to use a non-probability sample?

Compare simple random samples with various complications to them such as periodic samples, stratified samples (potentially with different weights to the various strata), cluster samples, and other complex sampling designs. Do these complications, such as multi-stage clustering, increase or decrease the efficiency of the sample? If they decrease the efficiency, why do people do them?

Be prepared to calculate sampling errors and to estimate required sample sizes in order to achieve a desired level of accuracy. Note that it depends on two things: the square root of the sample size, and the variance in the population. Get familiar with that equation and think about its implications for how you design a research project. Do you want to study things with high or low variance? Can you figure out ahead of time exactly how big a sample will be big enough?

Know the difference between standard error and a confidence interval and how to calculate them. If we want the confidence interval to be ± 3 points, then what does the standard error have to be? If the sample variance is 1000, then what does the sample size have to be?

Review various survey methods in Ch. 10, with a focus on the in-person interview. What can go wrong besides sampling error? What response rates do you see in published research? What do we know about the impact of low response rates?

Review one of the three examples from the literature in detail with a focus on the sampling strategy, the response rates, and the efforts of the authors to solve these sampling problems.

Paper topic:

When would it ever be more accurate to do a sample rather than a census? Review Morgenstern from last week on the costs and errors associated with gathering statistics from large populations such as all Americans, all businesses, all wars, all countries, or some other large group. Consider a hypothetical large population one might want to study with either a census or a sample in order to estimate some parameter. Assume that the cost of identifying, finding, and gathering the data on each individual costs X for the first 50% of the population, but then costs rise to $1.5X$, $2X$, $5X$, and $10X$ as you cross the thresholds of 50, 75, 80, and 90 percent of the population. Be as specific as you can in your answer. What would it depend on? How would you figure it out. At the end of your paper, give an example of a population where this would definitely not be worth it (e.g., a census would definitely be better), and one where it may indeed be better to do a sample, if you find that is possible.