How Much Should You Trust Your Power Calculation Results?
Power Analysis as an Estimation Problem

Dr Shiyao Liu
Postdoctoral Associate
New York University Abu Dhabi
Co-author: Teppei Yamamoto

Abstract: With the surge of randomized experiments and the introduction of pre-analysis plans, today's political scientists routinely use power analysis when designing their empirical research. An often neglected fact about power analysis in practice, however, is that it requires knowledge about the true values of key parameters, such as the effect size. Since researchers rarely possess definitive knowledge of these parameter values, they often rely on auxiliary information to make their best guesses. For example, survey researchers commonly use pilot studies to explore alternative treatments and question formats, obtaining effect size estimates to be used in power calculations along the way. Field experimentalists often use evidence from similar studies in the past to calculate the minimum required sample size for their proposed experiment. Common across these practices is the hidden assumption that uncertainties about those often empirically obtained parameter values can safely be neglected for the purpose of power calculation.

Such assumptions are often consequential and sometimes dangerous. We propose that, given the practice in empirical research, power analysis should be viewed as an estimation problem, rather than as a matter of mere calculation as usually taught in statistics classes. From this perspective, we analyse three commonly used variants of power analysis – minimum required sample size, minimum detectable effects, and the power itself. How reliable these analyses can be under scenarios resembling typical empirical applications in political science would be discussed. Specifically, we analytically examine small sample properties of the plug-in estimators of these quantities and explore bias-correction techniques for the estimators. Monte Carlo simulations are used to investigate their performance in likely scenarios. The results reveal that these estimators are generally biased and that the biases are often large enough to render the entire analysis of little practical value. We illustrate our results with empirical examples drawn from subfields of political science while practical guidelines are offered to empirical researchers on when to (and not to) trust power analysis results.