Chapter 3
Using Propensity Score Matching to Improve Validity in Public Administration Research

Michael Howell-Moroney
The University of Memphis, USA

ABSTRACT

Randomized clinical trials have a longstanding status as the gold standard in detecting causal effects. In the social sciences, randomized clinical trials are rare because of their attendant logistical and cost burdens. Most social science research makes use of observational data. The empirical challenge posed by observational data is that treatment assignment is no longer random. This challenge continues to spur innovation across many disciplines toward more sophisticated techniques for estimating causal relationships. Scholars have developed a common theoretical framework for estimating causal effects, often called the potential outcomes or counterfactual framework. This chapter demonstrates the propensity score matching methodology as a way to estimate causal effects using observational data. Throughout, an example from public administration research, the effect of government employment on volunteerism, is used to illustrate the concepts. Empirical estimates of the treatment effects show that there may be a causal effect of government employment on volunteerism.

INTRODUCTION

Scholars in public administration have called for increased use of experiments in public administration research (Bozeman and Scott, 1992; Perry, 2012). Advocates of the behavioral public administration approach have vigorously argued for the use of experimental methods to advance theory and establish causation (Hassan and Wright, 2020). Although experimental approaches are a welcome addition to public administration research, it is still not clear whether these methods will enjoy widespread adoption. As Perry (2012) writes, "It is fair to observe that not much progress has been made in the intervening years in the application of experimental methods to our field" (p.480).

DOI: 10.4018/978-1-7998-8243-5.ch003

There are a few reasons why experimental methods are underused in public administration. First there is disciplinary inertia that limits usage of experimental methods. Disciplines that use experimental methods like psychology and sociology have a long history of doing so. Changing disciplinary inertia would mean a large shift toward methods that are at best, uncommon among public administration scholars. Second, experimental methods are not a common component in the training of public administration doctoral students (Bozeman and Scott, 1992). The lack of doctoral training in experimental methods creates an additional source of reinforcement for disciplinary inertia. If future scholars do not receive the requisite training, then it is unlikely that they will use experimental methods in their own research. Third, experimental designs are not well-suited for research that uses an aggregate unit of analysis, such as organizations or cities. The reason is that aggregate groups cannot generally be randomized into treatment and control groups, which is necessary for an experiment. Fourth, the nature of the study itself can preclude an experimental design. Sometimes the variable of interest cannot be randomized. Consider a study looking at the effect of gender on job performance. It would not be possible to randomly assign subjects to different gender groups.

Because randomized clinical trials are often not feasible, social scientists mostly rely on observational data. Observational data differs from the data in a randomized clinical trial because assignment to treatment is not based upon a random process. This challenge has been addressed by scholars in numerous disciplines using a variety of quasi-experimental methods. The goal of these methods is to estimate an elusive quantity, the *average treatment effect*. This chapter discusses the counterfactual framework as a method of estimating treatment effects using observational data.

The randomized clinical trial has long been recognized as the gold standard in establishing causation (Hariton and Locascio, 2018). At the first stage, a randomized clinical trial randomly assigns subjects to a treatment or control group. The treatment is applied to the treatment group and then the group differences are compared, often using a standard difference-of-means test. The randomization of assignment to the treatment or control group eliminates the problem of self-selection bias. That is, subjects do not control whether they are assigned to the treatment or control. In addition, randomization ensures a balance on other variables relevant to the study, including unobservable characteristics. Because the treatment assignment mechanism is random, that effectively hedges against other intervening variables confounding the treatment effect. Through randomization, the randomized clinical trial produces an unbiased measure of the difference between groups (Hariton and Locascio, 2018).

Instead of data generated by randomized clinical trials, much social science research uses observational data where the treatment is not random. Here "treatment" takes on a much broader meaning than just a medical procedure. Treatment can be thought of as the cause for a particular effect. If the objective were the study of a of voluntary job training program, the treatment would be participation in the job training program and the effect some labor market outcome. In a study of how state death penalty policies affect violent crime, the treatment would be the state policy on the death penalty and the outcome some measure of crime.

In these examples, the treatment is not randomly assigned, but is the result of choice (self-selection bias). In cases where treatment is nonrandom, researchers must take care in attributing causal effects. Consider the voluntary job training program, where individuals self-select. It would be important to know if those that participated in the program are somehow different from those that did not. State death penalty statutes are also not a random occurrence. Some states choose to have the death penalty, while others do not. When there is a self-selection mechanism that determines treatment, there can often be substantial bias in estimates of treatment effects.