



Why Randomized Trials May Not Adequately Measure the Impact of School Choice Programs

Christopher Lubienski, Indiana University-Bloomington

Figuring out what works in school reform is vitally important – especially learning whether school choice programs actually lead to educational improvements. Debates over which research methods to use have been central, for example, in efforts to pass an extension of the Opportunity Scholarship Program in the District of Columbia. Should evaluations of this voucher program rely exclusively on a randomized trial research method, where researchers compare outcomes for students whose families won vouchers in a lottery to outcomes for students who did not win the vouchers (the “control group”)? Similar discussions rage following the reauthorization of the U.S. Elementary and Secondary Education Act. What types of research designs should be used to evaluate school reforms in future years?

The randomized trial approach has been called the “gold standard” in educational evaluation research, because it resembles the approach medical researchers use to evaluate the effects of new drugs or treatments on a randomly selected “treatment group” versus the otherwise similar “control group.” However, an exclusive reliance on this approach has met some resistance, and policymakers are also considering additional research designs – including quasi-experimental studies and correlational studies, which are also seen as useful.

Although in some ways arcane, discussions about how best to gather and analyze evidence about the impact of educational reforms are important because policymakers need to determine which programs work best and assess whether they can work equally well for broad populations and various subsets of schoolchildren and their families.

Assessing School Choice Programs

Careful evaluation of controversial school choice programs is of particular interest. Proponents argue that these programs can improve student outcomes and broaden access to quality schools. Skeptics want to see evidence that these programs actually deliver promised benefits and avoid detrimental side-effects, such as blocking access to good schools for certain children. This debate was front and center during the reauthorization of the DC voucher program. In the past, the program has been evaluated by comparing outcomes for students randomly selected for a voucher against those randomly rejected. Yet final results of the federal evaluation in 2010 found “no conclusive evidence that the program affected student achievement,” although there was evidence that some students offered a voucher were more likely to graduate from high school. As the evaluator explained in Congressional testimony, the vouchers seemed to make a difference “for some subgroups but not necessarily others, overall in some years but not others, and in one educational domain (reading) and not the other (math).” This result, he said, is “roughly consistent with what most other rigorous private school choice evaluations have found.” As Princeton economist Cecilia Rouse and Lisa Barrow sum up, “research to date finds relatively small achievement gains for students offered vouchers, most of which are not statistically different from zero.”

In the wake of inconsistent and insignificant results from randomized voucher studies, this research method has now turned into a bit of a political football. Supporters of vouchers originally embraced this method as likely to prove that vouchers are effective tools to improve education, but now voucher opponents call for randomized studies as a way to prove that these programs have little or no impact.

The Limits of Randomized Trials

Randomized control trials resonate with policymakers not only because they have a good reputation among experts, but also because the simplicity and beauty of this research design is easily understood and quite

compelling. The best way to tell whether vouchers work, presumably, is to compare a random group of students who got them to another similar group of students who did not win them in a lottery. But there are also real limitations to this approach.

The results of a particular randomized trial cannot be generalized. For example, after a 2010 study found that DC students offered a voucher graduated from high school at a higher rate than students randomly rejected, the head of the new organization administering vouchers proclaimed that expanding the program would boost graduation rates for all students in DC Public Schools. But that may not be true, because graduation requirements may differ – for example, between public schools and some of the private schools of questionable quality that have emerged to attract voucher students. What is more, the families that applied for vouchers are more motivated to help their children’s educational prospects.

Proponents of voucher programs often claim that they give poor families the same choices about schools that affluent families already have. But we cannot draw this broad conclusion because there are actually three categories to consider: children from affluent families that can choose private as well as public schools; children from less affluent families that seek alternatives such as vouchers; and children from families that do not even apply. Randomized control trials can only tell us whether vouchers improve education for the small minority of students whose families actually applied – less than 3% of all DC public school children in 2014-15.

Randomized studies in education have other limitations. Unlike in medical trials, there is no placebo. Families and students know if they have been selected or rejected to get vouchers – and this can affect their later behavior and attitudes. Additionally, these studies fail to consider the social and economic mix of students in various schools, even though scholars already know that peer groups affect student learning.

The Bottom Line

As they assess the impacts of school choice or other educational interventions, policymakers would do well to avoid relying on any single study or particular research design. Randomized control trials are helpful, but do not reveal the whole picture. Multiple methodologies can provide different kinds of insights. In addition, we also need to consider whether a particular study has been independently peer-reviewed, whether it has been replicated, and how its results fit into larger bodies of research about how schools work. Policymakers should proceed with caution, and understand the limitations of any single study or method of research.