

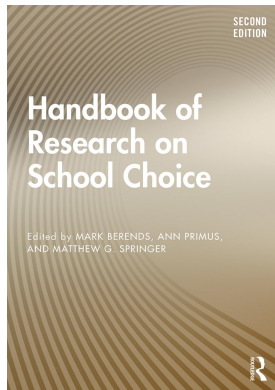
This article was downloaded by: 10.2.97.136

On: 03 Oct 2023

Access details: *subscription number*

Publisher: *Routledge*

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: 5 Howick Place, London SW1P 1WG, UK



## **Handbook of Research on School Choice**

Mark Berends, Ann Primus, Matthew G. Springer

### **Voucher Outcomes**

Publication details

<https://test.routledgehandbooks.com/doi/10.4324/9781351210447-17>

Karin Gegenheimer, Matthew G. Springer

**Published online on: 25 Jun 2019**

**How to cite :-** Karin Gegenheimer, Matthew G. Springer. 25 Jun 2019, *Voucher Outcomes from: Handbook of Research on School Choice* Routledge

Accessed on: 03 Oct 2023

<https://test.routledgehandbooks.com/doi/10.4324/9781351210447-17>

**PLEASE SCROLL DOWN FOR DOCUMENT**

Full terms and conditions of use: <https://test.routledgehandbooks.com/legal-notices/terms>

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

# VOUCHER OUTCOMES

*Karin Gegenheimer and Matthew G. Springer*

School vouchers have been the subject of much controversy since becoming a key player in the school choice movement. Vouchers offer families a fixed sum of money, often called scholarships, to send their children to private schools. Both publicly and privately funded voucher programs operate within the United States, but they tend to be relatively small in scale with respect to the number of students they admit. Other countries have implemented universal voucher programs, which serve as core components of their educational systems.

In practice, the evidence on the effects of vouchers is largely mixed—a finding (or lack thereof) that has contributed to the controversy surrounding these programs. Scholars—including Ladd (2002); Neal (2002); McEwan (2004); Gill, Timpane, Ross, and Brewer (2007); Levin (2008); Rouse and Barrow (2009); Figlio and Hart (2014); Epple, Romano, and Urquiola (2017); and Austin and Berends (2018)—have conducted a number of systematic reviews examining several issues related to voucher outcomes, such as student achievement, student sorting, and competitive effects. The purpose of this chapter is to provide a review of the empirical evidence on the effectiveness of school vouchers, both in the U.S. and abroad. In this review, we take a student-centered approach, focusing on the effects of school vouchers on the students who use them, with particular attention to academic outcomes.

## **Grades of Social Science Evidence**

To describe the empirical evidence on the effects of school vouchers, it is important to first consider the grades of social science research. Here we discuss the four main types, from most to least rigorous: 1) experimental, 2) quasi-experimental, 3) correlational, and 4) descriptive.

Experimental research provides the best method possible to examine a cause and effect situation. The most rigorous form is the randomized controlled trial, often referred to as the “gold standard” in education research. In a randomized controlled trial, researchers randomly assign individuals from the sample to either the treatment group (e.g., students who use a voucher) or control group (e.g., students who do not); this creates baseline equivalence between groups when properly implemented. That is, individuals in each group are on average similar in expectation across measures of observable and unobservable characteristics. The ability to establish baseline equivalency between the treatment and control groups is important because it allows researchers to attribute any change in the outcome to the treatment itself, and not to prior differences between the two groups.

To illustrate this point, imagine a study that examines the effect of an after-school tutoring program on student test scores. To do this, researchers simply compare the test scores of students in the tutoring program to those of students who are not in the tutoring program. They find that students who receive tutoring score higher than students who do not, and conclude that the tutoring program positively impacts test scores. What is the issue with this? Selection bias. Students who selected into the tutoring program could be systematically different from students who did not select into tutoring, such that their test scores would have been higher regardless of whether or not they participated in the program. Selection into treatment is informative of characteristics that may be correlated with the outcome of interest.

Imagine now that students are randomly assigned to the tutoring program. Some students who would have willingly participated are assigned to treatment (tutoring), and some are assigned to the control group (no tutoring). Assignment to treatment is no longer informative simply because it is random. In this case, researchers can claim that the observed differences in student test scores are the result of the tutoring intervention and not of pre-existing differences between the treatment and control group.

Another type of experiment, natural experiments, takes advantage of situations in which individuals are randomly assigned to treatment or control not by the researcher, but by existing systems or structures. For example, when voucher programs or charter schools use lottery systems to admit students because they are oversubscribed, the lottery itself is, in effect, randomly assigning students to treatment (attending a choice school) or control (not attending a choice school).

Unlike randomized controlled trials, natural experiments cannot control treatment implementation and, as a consequence, two main issues arise. First, these designs can lack external validity because results would only have inferences to oversubscribed programs or schools. Second, the researcher's inability to directly control treatment gives rise to issues of noncompliance. That is, the researcher is unable to ensure that individuals who are assigned to treatment actually receive that treatment. While there are ways to address this type of attrition and still estimate a treatment effect (e.g., calculate an intention-to-treat effect and a treatment-on-treated effect), this distinction is important because there are often systematic differences between those who were assigned treatment and received it, and those who were assigned treatment but did not receive it. Still, natural experiments are widely used in education research, as they are often more practical than a randomized controlled trial.

Some forms of natural experiments can also fall under quasi-experimental designs (e.g., natural disasters unexpectedly change the circumstances under which schools operate or researchers leverage a strict cut-point based on some academic or social dimension that allocates units to receive an intervention and others to not); however, for expositional purposes in the general context of voucher research, we delineate between natural experiments based on oversubscription lotteries and other quasi-experimental approaches that harness different identifying assumptions, some of which we describe in more detail throughout this chapter.

The next grade of social science research methods is the quasi-experimental study. This consists of observational research that carefully attempts to isolate the effect of a treatment through means other than randomization. Quasi-experimental designs simulate experimental methods but, instead of establishing baseline equivalence through random assignment, they establish equality by creating synthetic treatment and control groups or by exploiting longitudinal data that allow comparison of outcomes of interests for students overtime switching, say, between a traditional public school and school of choice. Common methods within quasi-experimental research include regression discontinuity, difference-in-differences, instrumental variables, fixed effects, and propensity score matching, where each takes a different approach to creating a comparison condition. Quasi-experimental research often exhibits more external validity than experimental research; its methods rely on observational data that can be more comprehensive and representative of larger populations than what

is typically used in experimental settings (e.g., not restricted to oversubscribed schools). However, without random assignment or a true control group, causal claims become slightly more difficult to defend.

The final two grades of social science methods are correlational and descriptive studies. Unlike the first two types of social science research, correlational research moves away from estimating cause and effect and focuses on describing relationships among variables in a particular sample. Although correlational research cannot establish causality, it is often helpful in illuminating patterns that merit further attention from quasi-experimental or experimental methods. And lastly, descriptive research is largely exploratory in nature and focuses on describing characteristics of a single sample. Descriptive studies include both quantitative and qualitative work, and are often used to study phenomena on which little to no prior research exists.

For the purposes of this chapter, we limit our review to experimental and quasi-experimental studies to focus our evidence on the cause and effect of voucher programs. It goes without saying that restricting our review in this manner greatly limits the richness of available information. For example, we omit important voucher studies like Witte (2000) and Metcalf, West, Legan, Paul, and Boone (2003). (These studies are included in the reference list for readers.) Yet, given that our aim is to consider the effectiveness of school vouchers, it seems right to limit this discussion to studies that seek to identify the causal effect of voucher programs on student outcomes.

### **Academic Outcomes**

We review the evidence on the effects of school vouchers on students who use them, focusing here specifically on student test scores as the outcome of interest. Indeed, much of the literature on school vouchers uses test scores as the primary metric by which to assess their effectiveness. However, we acknowledge that test scores are not the only indicator of school performance and that they are in fact a rather narrow measure of student success. There is a host of outcomes that can shed light on the performance of voucher students relative to students in public schools. For example, researchers have examined the effects of vouchers on student satisfaction, parent satisfaction, and students' non-cognitive skills (e.g., Bettinger & Slonim, 2006; Howell & Peterson, 2006; Bettinger & Slonim, 2007). This research is without question important to understanding the effectiveness of voucher programs. However, we limit most of our review to the major pieces of work examining student test scores, as they are a good proxy for student learning—arguably the primary purpose of schooling—and because this is where the bulk of the research on vouchers exists. We also consider effects on educational attainment, following prior work highlighting the importance of attainment as a measure of school choice effectiveness (Hitt, Wolf, & McShane, 2018). Our review also differentiates between the two main types of voucher programs in the U.S., each of which we discuss in turn: 1) publicly funded voucher programs, and 2) privately funded voucher programs.

#### ***Publicly Funded Voucher Programs in the United States***

Publicly funded school vouchers offer students state-funded scholarships to attend private schools participating in the program. Although state statutes largely determine specific program components, the typical publicly funded school voucher program operates within a single district, city, or state—with the latest statewide programs in Indiana, Ohio, and Louisiana—and targets students on the basis of socioeconomic status or public school academic performance. In addition, state statutes often restrict the size of their voucher programs by limiting program parameters, such as the number of participating private schools, the number of vouchers offered each year, and the percentage of student enrollment allocated to voucher recipients. The evidence presented here, then, considers the effects of voucher programs operating largely within these specific constraints.

The first publicly funded urban voucher program in the U.S. is the Milwaukee Parental Choice Program (MPCP). MPCP launched in 1990 as a relatively small-scale program targeting public school students from low-income families. The program has since undergone several expansions that broaden its reach in terms of participating schools and students. Today, MPCP includes both secular and religiously affiliated K–12 private schools and provides vouchers to approximately 15,000 students. It remains the longest-operating voucher program in the U.S., having been reauthorized in 2006 under Wisconsin Act 125.

The evidence from Milwaukee is largely mixed. In an early study, Greene, Peterson, and Du (1999) examined the effects of voucher schools on math and reading scores, exploiting the lottery system through which voucher students were selected to compare voucher users to unsuccessful voucher applicants who remained in Milwaukee public schools (MPS). The authors found that in the third and fourth years of attending a voucher school, voucher users saw higher math and reading scores—with particularly large gains in math—in comparison to unsuccessful applicants. By restricting the comparison group to unsuccessful applicants in public schools, however, this study may have overestimated the effect of vouchers if a substantial number of unsuccessful applicants then exited the public school system and self-selected into the private sector.

Rouse (1998) took a similar approach to estimate the effects of vouchers on student achievement, but she expanded her sample to compare not only voucher users and unsuccessful applicants, but also voucher users and a random sample of MPS students. Comparing against public school students was intended to neutralize the estimated voucher effects by generating a set of results against which to compare the estimates from the quasi-experimental model. In addition, Rouse took a student fixed-effects approach to control for unobserved, time-invariant student heterogeneity. She found that, in math, voucher users outperformed unsuccessful voucher applicants and MPS students. In reading, on the other hand, effects were not as straightforward, with both negative and positive coefficients that were statistically indistinguishable from zero.

More recent research has compared voucher users' academic outcomes with those from a matched comparison group of MPS students. Findings from Witte and colleagues (Witte, Carlson, Cowen, Fleming, & Wolf, 2012; Witte, Wolf, Cowen, Carlson, & Fleming, 2014) showed positive effects in reading and math. Yet, the results were concentrated among specific subgroups of students, indicating that using a voucher to attend a private school may have been more impactful for some students than for others. For example, the effect of using a voucher on reading test scores was statistically significant only for students in the lower end of the initial achievement distribution (pre-voucher achievement). For math, the effects were largest among students in the upper end. The authors noted an important caveat to these findings, however. The positive test score gains appeared in the first year after the state implemented a test-based accountability policy. Witte et al. (2012, 2014) were thus hesitant to attribute the test score gains to voucher use, suggesting instead that they were more likely due to responses to the testing regime.

Another among the first of the large-scale voucher programs in the U.S. is the Cleveland Scholarship and Tutoring Program, a state-funded program implemented in 1996 that targets students from low-income families. In one of its most rigorous evaluations, Belfield (2006) compared voucher users to public school students whose voucher applications were rejected due to income ineligibility. The findings were rather inconclusive. Voucher students tended to underperform in math and outperform their public school counterparts in language, but the positive effects on language lost significance in models estimating the effects on fourth grade achievement. It is important to note that, although Belfield's research design addressed selection bias by comparing two groups of students who applied for vouchers, it did not account for the fact that the treatment and comparison groups were systematically different in terms of family income. Thus, the estimates may have been biased; one group could have had higher or lower levels of achievement due to factors associated

with income, not factors related to using or not using a voucher. The results should therefore be interpreted carefully.

Although the initial work on voucher outcomes generally tended to indicate mixed results, evaluations of more recently implemented school voucher programs have suggested that vouchers have a negative effect on student achievement. The District of Columbia School Choice Incentive Act of 2003 authorized the first federally funded school voucher program, the D.C. Opportunity Scholarship Program (OSP). In 2011, the program was reauthorized and expanded as part of the Scholarships for Opportunity and Results Act. The OSP remains the only federally funded voucher program to date.

Wolf et al. (2010, 2013) conducted official evaluations of the OSP, comparing voucher users to unsuccessful applicants who were not selected through the random lottery. The authors found no effects on math or reading until year four, at which point effects on reading became positive and significant. Yet, this effect was concentrated among female students and students who had not been in so-called “schools in need of improvement” (SINI) (i.e., students who were not flagged as priority applicants based on family income level). Dynarski, Rui, Webber, and Gutmann (2017) re-evaluated OSP after its first year of reauthorization and expansion and found that using a voucher actually harmed student achievement. After one year in a private school, voucher users experienced negative growth in math and no significant changes in reading relative to unsuccessful applicants in public schools. Again, though, these results varied by student subgroup. Among K–5 and non-SINI students, the impact of using a voucher was negative in both math and reading.

In Ohio, Figlio and Karbownik (2016) compared a matched sample of EdChoice Scholarship voucher users and public school students to estimate the effects of program participation on student test scores. They found that students who moved to private schools using a voucher performed worse on reading and math state tests than did students in the matched comparison group. The authors acknowledged, however, that Ohio private schools do not face the same state testing accountability mechanisms that public schools do, which may have accounted for voucher students’ relative achievement deficit. That is, instead of devoting instructional resources to state test preparation, private schools may use their curricular flexibility to create alternative instructional priorities. Voucher students’ underperformance on state tests relative to students in public schools, then, does not necessarily indicate that using a voucher leads to lower overall academic performance.

Waddington and Berends (2018) found similarly negative effects from the Indiana Choice Scholarship Program. This program is unique in that its eligibility requirements target both low- and middle-income families, and the amount of tuition that students receive from the voucher depends on their family’s income level. In their study, Waddington and Berends focused on full voucher students (i.e., students from low-income families who receive either 90 percent or full tuition). The authors created a matched sample of students who received a voucher to attend a private school and students who did not receive a voucher and remained in a public school. The treatment group was restricted to students who switched from a public to private school upon voucher receipt. Voucher students performed worse in math compared to both their prior performance in a public school and to the performance of their public school counterparts. There were no differences in English/Language Arts. The authors concluded that voucher students performed similarly to public school students in English/Language Arts, but tended to face losses in mathematics that persisted up to four years that students received a voucher.

Results from the Louisiana Scholarship Program painted perhaps the least optimistic picture of voucher efficacy. Abdulkadiroğlu, Pathak, and Walters (2015) and Mills and Wolf (2017) studied voucher impacts after the first and second years of the program’s statewide expansion. Similar to Milwaukee’s voucher program, Louisiana awards vouchers by a random lottery in schools with more eligible applicants than available seats. Abdulkadiroğlu et al. (2015) compared students awarded vouchers through the lottery to unsuccessful voucher applicants, limiting their analysis



to students who applied for vouchers in the first year after statewide expansion, and found that voucher users scored lower than public school students in both math and reading. These negative effects on achievement decreased in the second year after statewide expansion but remained statistically significant (Mills & Wolf, 2017). Explaining the large negative effects, Abdulkadiroğlu et al. (2015) suggested that school selection into the statewide voucher program could have biased the estimates; if private schools that applied to be voucher schools were low-performing, they may have wanted voucher recipients to increase enrollment. Alternatively, Mills and Wolf (2017) posited that misalignment between private school curriculum and the Louisiana state tests could have played a role in the program's seemingly inefficacy, as private schools only recently began to implement curriculum aligned to the state tests.

The literature examining the impact of publicly funded voucher programs on student achievement overwhelmingly has shown mixed results. It is worth noting, however, that among more recently implemented programs, the evidence seems to suggest—with added certainty—that vouchers negatively influence student achievement. Although the research discussed here is among the most rigorous in terms of employing designs that come close to estimating causal effects, these studies cannot account for all sources of bias. For example, as we mention above, studies that exploited random lotteries to generate instrumental variables estimates of voucher effects are subject to bias if the comparison group excluded unsuccessful applicants who selected out of the public school sector. Studies that took advantage of matching strategies to create synthetic comparison groups attempted to limit selection bias by matching the treatment and comparison groups on a set of observable characteristics. But this technique cannot fully eliminate the sources of selection bias that stem from unobservable characteristics. So again, it is important to interpret the evidence with such considerations in mind.

### ***Privately Funded Voucher Programs***

There are several privately funded voucher programs operating throughout the U.S. Like publicly funded programs, private programs offer students tuition scholarships, often on the basis of socioeconomic status. The size of these privately funded programs can vary a great deal. Perhaps the largest one is the Children's Scholarship Fund (CSF), founded by the Walton Family Foundation. The CSF collaborates with various partner programs to manage voucher programs in multiple cities across the U.S. Aside from its relatively large size, researchers have given quite a bit of attention to the CSF because its design mimics a randomized field trial. The CSF selects voucher recipients through a random lottery, where all applicants face an equal probability of selection. This process creates an experimental setting wherein study participants are randomly assigned to treatment, which researchers have used to estimate the impact of being offered a voucher on various student outcomes.

Much of the available evidence on the CSF comes from four sites: New York City; Washington, D.C.; Dayton, Ohio; and Charlotte, North Carolina. Data collection at each of the sites followed the same procedure. Applicants were required to submit baseline test scores as part of their eligibility requirements. In each subsequent year of the study, both successful and unsuccessful applicants were invited to participate in follow-up testing sessions.

Howell and Peterson (2002, 2006) and Howell, Wolf, Campbell, and Peterson (2002) conducted CSF impact evaluations in New York, D.C., and Dayton by comparing the year over year growth in test scores of voucher recipients to that of unsuccessful applicants. These studies presented intent-to-treat estimates, as they employ an instrumental variables approach where being offered a voucher is an instrument for private school attendance. The authors found little evidence that being offered a CSF voucher improves test scores for all students.

These studies did, however, find positive effects on achievement for Black students. In New York City, Black students experienced test score gains in each of the three years of the study, a finding

that has been replicated in other studies that singularly have evaluated the CSF in New York City (e.g., Mayer, Peterson, Myers, Tuttle, & Howell, 2002; Barnard, Frangakis, Hill, & Rubin, 2003). The results in D.C. and Dayton were less consistent from year to year; nonetheless, the evidence has suggested that among Black students, voucher recipients outperform unsuccessful applicants.

Using evidence from New York City, Krueger and Zhu (2004) disputed the claim that the CSF has improved achievement for Black students. They demonstrated that the positive effects among Black students were sensitive to how student race was defined. Whereas Howell and Peterson's previous studies (2002, 2006) identified a student as Black if the student's mother is Black, Krueger and Zhu (2004) showed that if a student is identified as Black based on either the mother's or the father's race, the results for Black students decreased in magnitude and lost significance. Their results regained significance, however, when they controlled for baseline test scores. While the authors do not hypothesize why using a broader definition of Black students nullifies the voucher effects, they do conclude that the sensitivity of these findings suggests that the CSF program in New York City has trivial effects on participating Black students' achievement. Responding to Krueger and Zhu, Peterson and Howell (2004) acknowledged that their results were not robust to alternative specifications of race, but maintained that the overall evidence suggested positive effects for Black students.

Barnard et al. (2003) and Jin, Barnard, and Rubin (2010) re-analyzed CSF effects in New York City, employing particular methodological approaches to account for study attrition and missing data. Using a principal stratification approach, Barnard et al. (2003) produced results consistent with Howell and Peterson (2002, 2006)—that is, that vouchers have a positive and significant effect on achievement for Black students. Jin et al. (2010) used a modified general location model, and while they did not conduct subgroup analyses by race, they did conduct such analyses based on students' prior public school performance, which generated a set of interesting results. In math, voucher recipients who came from low-performing public schools experienced achievement gains relative to those students who came from high-performing public schools. The effects on reading were less straightforward. In grade four, the results were positive for students originally from low-performing public schools; in grade one, the positive effects only held for students from high-performing public schools.

Further exploring the differential effects of CSF vouchers by prior achievement levels in New York City, Bitler, Domina, and Penner (2013) found results consistent with Jin et al. (2010). Specifically, Bitler and colleagues demonstrated that, although being offered a voucher did not impact math achievement, students accepted to CSF scored higher in reading compared to unsuccessful applicants, and these effects were greater for students at the lower quintiles of the prior achievement distribution. These results provide preliminary evidence that vouchers may reduce educational disparities, as the positive effects seem to be concentrated among lower-performing students. In sum, the results from CSF are mixed. There is evidence to suggest that CSF vouchers do benefit specific subgroups of students, though this finding is apparent mainly in New York City.

The results from the CSF in Charlotte, however, tell a different story. Greene (2000) compared voucher users to students who did not receive vouchers to students who received vouchers but did not use them. His results thus represent both intent-to-treat and treatment-on-the-treated estimates. Greene found that voucher users outperformed both comparison groups in math and reading. Cowen (2008) conducted a similar analysis, but found somewhat different results. He employed a complier average causal effect model to detect the effect of attending a private school for students who were offered a voucher and used it. His estimates showed a positive effect in reading, but the results for math were statistically indistinguishable from zero. In all, the results from Charlotte further highlight the concern that estimates of voucher efficacy are sensitive to model specifications. Still, the bulk of the work on CSF (across multiple sites) indicates that using vouchers to attend private schools is beneficial for some, but not all, students.



## College Attendance and Persistence

In addition to student test scores, school vouchers may impact student achievement by influencing college attendance and persistence. Despite the degree to which vouchers influence more immediate measures of student achievement (i.e., test scores), one could still be inclined to support school vouchers if they were effective in encouraging students' college-going behaviors. Because large-scale voucher programs are a relatively modern phenomenon in the educational landscape, it has only recently become possible to study the relationship between school vouchers and long-term student outcomes like college attendance. Evidence suggests that vouchers do indeed have a positive effect in this domain.

Among some of the first work on long-term voucher effects, Cowen, Fleming, Witte, Wolf, and Kisida (2013) studied the impact of Milwaukee's voucher program on high school graduation and college enrollment and persistence. Using an intent-to-treat approach, the authors compared MPCP students to a matched sample of MPS students, following both groups for a total of six years, from ninth grade to what would be two years after on-time high school graduation. They found that students who were exposed to MPCP were more likely to attend and persist in a four-year institution (where college persistence is defined as being enrolled in the same institution in the first and second year after graduating high school). While the results for high school graduation were not significant at a conventional level, the findings regarding college attendance imply encouraging results. That is, for students who do attend college, exposure to the MPCP has a positive effect on on-time high school graduation, as students who are enrolled in college during the study's time frame would have had to graduate from high school in four years.

Evidence of positive effects on college enrollment also comes from New York City's CSF program. Chingos and Peterson (2015) found that being awarded a voucher increased the probability of college enrollment and degree attainment, but only for minority students and students with U.S.-born mothers. The authors explained that the positive results among the sub-sample of minority students align with prior work on NYC's voucher program showing different effects on student achievement by student race—namely, that voucher use is associated with achievement gains for Black students only (e.g., Mayer et al., 2002; Barnard et al., 2003). That the study found positive effects among students with U.S.-born mothers, however, is more difficult to interpret. The authors speculated that information asymmetry between U.S.-resident families and non-resident families could have driven these results. Perhaps voucher students with U.S.-born mothers selected higher-performing private schools because their families, having lived in the U.S., were more familiar with the educational landscape. This explanation is plausible, given the complexity of NYC's K–12 education system and the high level of immigrant residents in the city.

Wolf et al. (2010, 2013) found similarly positive effects in Washington, D.C. In the first round of the OSP, these studies found that both students who were offered and students who were awarded a voucher were more likely to graduate from high school. However, Wolf et al. (2013) noted that these effects were concentrated among female students, students on the higher end of the initial achievement distribution, and low-income students. In addition, recall that these evaluations largely found null effects on math and reading achievement, which led the authors to posit that private schools that participate in the voucher program may place more emphasis on teaching soft skills that are necessary for high school graduation than on standardized testing.

The small but growing body of literature on voucher effects on college attendance and persistence is growing, but initial results present an encouraging perspective. If a primary purpose of K–12 education is to prepare students for postsecondary education, then it seems as though voucher programs offer a step in the right direction.

## Voucher Programs Abroad

Several other countries have implemented school voucher programs. Two of the largest programs are Chile's nationwide voucher program, and Colombia's Programa de Ampliación de Cobertura de la Educación Secundaria (PACES). The evidence from Chile is informative because it speaks to the effectiveness of a voucher program that operates at the national level and as a primary component of the nation's education system. However, its design also presents challenges for program evaluation. Chile offers vouchers to all students, which creates issues of selection bias. Specifically, student selection into voucher use is not random; therefore, it is difficult to know whether any change in student outcomes is due to a voucher effect or due to issues related to non-random selection into treatment. For this reason, we focus the bulk of our review on the evidence from Colombia's PACES program, while also offering recent evidence from a small state-specific voucher program in India. Additionally, we encourage readers to see chapters in this volume by Ladd and Fiske (Chapter 7) and Levin (Chapter 16).

The national government of Colombia established the PACES program in 1991 as part of an effort to decentralize its education system. Although PACES is government-funded, it is not universal. Rather, it targets students in the two lowest socioeconomic strata based on residential location. Unlike many voucher programs in the U.S., PACES is only open to students entering secondary school, and applicants must have attended a public primary school to be eligible to apply. Before applying, students must have been admitted to a participating private school (i.e., one that would accept the voucher as tuition payment). After initial voucher receipt, students could renew the voucher each year through grade 11 as long as they met grade promotion requirements. Students who were retained could no longer participate in the voucher program (Angrist, Bettinger, Bloom, King, & Kremer, 2002).

Angrist et al. (2002) and Angrist, Bettinger, and Kremer (2006) conducted the two primary studies on PACES. The 2002 study focused on short-term academic outcomes, such as grade repetition, middle school completion, test scores, and a host of non-academic outcomes. Here, we focus on the results regarding differences in voucher users' test scores relative to unsuccessful applicants. The authors randomly sampled students from the 1995 applicant cohort for testing. Testing occurred three years after the initial date of application. Using an instrumental variables model, where voucher offer is an instrument for attending a private school, the authors found that lottery winners had higher test scores than lottery losers, the magnitude of which is equivalent to the achievement gains associated with one additional school year. Although the estimates for the entire sample were significant and substantial, the positive effects increased in magnitude when the sample was limited to only female applicants.

In their 2006 study, Angrist et al. extended their evaluation to include long-term outcomes such as on-time high school graduation and student achievement measured by performance on Colombia's college entrance exam. Again, the authors used a similar instrumental variables model to estimate the voucher effects. Although measuring student performance on the college entrance exam invites room for bias—as lottery winners are more likely to take the exam compared to unsuccessful applicants—the authors addressed this concern by using a modified Tobit model that imputes artificial test score data. They found that being offered a voucher increased test scores, an observed effect concentrated among students at the lower end of the initial achievement distribution.

Lamarche (2010) revisited PACES to examine both short- and long-term student outcomes. He used quantile regression to explore the effects of being offered a voucher on various student subgroups. Lamarche replicated the Angrist et al. (2002) finding that vouchers have a positive effect on student test scores three years after application, but noted that the effect was most apparent among students whose prior achievement was on the lower end of the distribution. Alternatively, long-term effects—measured by student performance on the college entrance exam—were positive and

relatively consistent across the prior achievement distribution. Lamarche speculated that, because voucher renewal is contingent on strong academic progress, the voucher effect may disproportionately influence students on the lower end of the distribution who face a greater incentive to improve performance in order to keep their vouchers.

The most recent evidence on PACES comes from Bettinger et al. (2019). Using up to 12 years of data since program implementation, this study focused on long-term voucher impacts and found that students who used a voucher were more likely to graduate high school in four years or later and enroll in higher education. Additionally, among students who entered vocational post-secondary schools, voucher users completed more years of school and were more likely to graduate on time compared to nonvoucher users. Overall fertility is not affected, but the timing is. Voucher winners are less likely than others to have children during their teenage years. Looking even further out, findings demonstrate that students who attended a voucher school saw increased earning—particularly voucher users who attended vocational post-secondary schools. A broad interpretation of these results suggests that PACES improves high school attainment and employability.

Though not as large and long-standing as PACES, we observe another example of an international voucher program in India. In 2008, India implemented a voucher program in the state of Andhra Pradesh. The program was implemented as a randomized control trial, wherein researchers randomized voucher offers at both the village and the household level. Out of the 180 villages in Andhra Pradesh that have at least one private school, 90 were selected to be treatment villages and the remaining 90 control villages. Within the treatment villages, households were randomly selected to receive a voucher application. In the spirit of traditional voucher programs, applications were targeted to disadvantaged students currently attending public schools. Once submitting a voucher application, students were awarded vouchers through a random lottery.

Four years following the start of the program, Muralidharan and Sundararaman (2015) conducted the first evaluation of voucher effects. As their primary measures of student achievement, the authors used student test score data obtained from tests that the research team administered at the end of the second and fourth year of implementation. The study found no significant achievement differences in math and Telegu (the native language) between voucher winners and public school students. However, students who were awarded vouchers performed significantly better in social studies, science, and Hindi. The authors attributed these findings to differences in instructional minutes in voucher schools compared to government public schools. Public schools spend more instructional time on math and Telegu than do voucher schools. Voucher schools allocate these extra instructional minutes to other subjects like social studies and science. Because voucher students performed at the same level as public school students in math and Telegu despite receiving fewer instructional minutes, the authors concluded that the voucher program improves student outcomes.

Taken together, the results from Columbia and India indicate that vouchers can have a positive impact on student achievement, in both the short and the long term. Of course, the national context is important in interpreting results from any education program, and it is true that the educational landscape in Colombia looks much different than it does in the U.S. However, the work discussed here provides an optimistic perspective on voucher efficacy.

## Conclusion

As this chapter shows, determining the overall effect of vouchers on student achievement is a difficult task. Indeed, the lack of conclusive evidence on voucher efficacy is a main source of the controversy surrounding school vouchers. Much of the evidence from publicly funded programs in the U.S. suggests that using a voucher harms student achievement, whereas results from the privately managed Children's Scholarship Fund paint a more positive picture, at least for specific student subgroups. As researchers continue to add to our understanding of how vouchers influence student

outcomes—whether they are more immediate measures like standardized test scores, or long-term outcomes like high school graduation and college attendance—an area ripe for further inquiry is the degree to which program design, location, or funding source moderates voucher effects. Although the literature moves back and forth on whether or not school vouchers *overall* improve student performance, there is less discussion around how specific characteristics of voucher programs may lead to better or worse outcomes. And as school vouchers increasingly become a central part of our education system, this line of research is critical to understanding how policymakers can design voucher programs that help all students succeed.

## References

- Abdulkadiroğlu, A., Pathak, P.A., & Walters, C.R. (2015). *School vouchers and student achievement: First-year evidence from the Louisiana Scholarship Program* (Working Paper No. 21839). Cambridge, MA: National Bureau of Economic Research.
- Angrist, J., Bettinger, E., Bloom, E., King, E., & Kremer, M. (2002). Vouchers for private schooling in Colombia: Evidence from a randomized natural experiment. *The American Economic Review*, 92(5), 1535–1558.
- Angrist, J., Bettinger, E., & Kremer, M. (2006). Long-term educational consequences of secondary school vouchers: Evidence from administrative records in Colombia. *American Economic Association*, 96(3), 847–862.
- Austin, M., & Berends, M. (2018). School choice and learning opportunities. In B. Schneider (Ed.), *Handbook of the sociology of education in the 21st century* (pp. 221–250). New York, NY: Springer.
- Barnard, J., Frangakis, C.E., Hill, J.L., & Rubin, D.B. (2003). Principal stratification approach to broken randomized experiments: A case study of school choice vouchers in New York City. *Journal of the American Statistical Association*, 98(462), 299–311.
- Belfield, C. (2006). *The evidence on education vouchers: An application to the Cleveland Scholarship and Tutoring Program* (Working paper). New York, NY: National Center for the Study of Privatization in Education.
- Bettinger, E., & Slonim, R. (2006). Using experimental economics to measure the effects of a natural educational experiment on altruism. *Journal of Public Economics*, 91(1), 1645–1648.
- Bettinger, E., & Slonim, R. (2007). Patience among children. *Journal of Public Economics*, 90(1), 239–255.
- Bettinger, E., Kremer, M., Kugler, M., Medina, C., Posso, C., & Saavedra, J.E. (2019). *School vouchers, labor markets and vocational education* (Mimeo). Palo Alto, CA: Stanford University.
- Bitler, M.P., Domina, T., & Penner, E.K. (2013). Distributional effects of a school voucher program: Evidence from New York City (Working Paper No. 19271). Cambridge, MA: National Bureau of Economic Research.
- Chingos, M.M., & Peterson, P.E. (2015). Experimentally estimated impacts of school vouchers on college enrollment and degree attainment. *Journal of Public Economics*, 122, 1–12.
- Cowen, J.M. (2008). School choice as a latent variable: Estimating the “complier average causal effect” of vouchers in Charlotte. *Policy Studies Journal*, 36(2), 301–315.
- Cowen, J.M., Fleming, D.J., Witte, J.F., Wolf, P.J., & Kisida, B. (2013). School vouchers and student attainment: Evidence from a state-mandated study of Milwaukee’s Parental Choice Program. *The Policy Studies Journal*, 41(1), 147–168.
- Dynarski, M., Rui, N., Webber, A., & Gutmann, B. (2017). *Evaluation of the D.C. Opportunity Scholarship Program: Impacts after one year* (NCEE 2017–4022). Washington, D.C.: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.
- Epple, D., Romano, R.E., & Urquiola, M. (2017). School vouchers: A survey of the economics literature. *Journal of Economic Literature*, 55(2), 441–492.
- Figlio, D., & Hart, C.M.D. (2014). Competitive effects of means-tested school vouchers. *American Economic Journal: Applied Economics*, 6(1), 133–56.
- Figlio, D., and Karbownik, K. (2016). *Evaluation of Ohio’s EdChoice Scholarship Program: Selection, competition, and performance effects*. Columbus, OH: Thomas B. Fordham Institute.
- Gill, B.P., Timpane, M., Ross, K.E., & Brewer, D.J. (2007). *Rhetoric versus reality: What we know and what we need to know about vouchers and charter schools* (2nd ed.). Santa Monica, CA: RAND Corporation.
- Greene, J.P. (2000). *The effect of school choice: An evaluation of the Charlotte Children’s Scholarship Fund Program*. New York, NY: Manhattan Institute.
- Greene, J.P., Peterson, P.E., & Du, J. (1999). Effectiveness of school choice: The Milwaukee experiment. *Education and Urban Society*, 31(2), 190–213.
- Hitt, C., Wolf, P.J., & McShane, M.Q. (2018). Achievement versus attainment: Are school choice evaluators looking for impacts in the wrong places? In M. Berends, R.J. Waddington, & J. Schoenig (Eds.) *School choice at the crossroads: Research perspectives* (pp. 173–185). New York, NY: Routledge.

- Howell, W.G., & Peterson, P.E. (2002). *The education gap: Vouchers and urban schools*. Washington, D.C.: Brookings Institution.
- Howell, W.G., & Peterson, P.E. (2006). *The education gap: Vouchers and urban schools* (rev. ed.). Washington, D.C.: Brookings Institution.
- Howell, W.G., Wolf, P., Campbell, D., & Peterson, P.E. (2002). School vouchers and academic performance: Results from three randomized field trials. *Journal of Policy Analysis and Management*, 21(2), 191–218.
- Jin, H., Barnard, J., & Rubin, D.B. (2010). A modified general location model for noncompliance with missing data: Revisiting the New York City School Choice Scholarship Program using principal stratification. *Journal of Educational and Behavioral Statistics*, 35(2), 154–173.
- Krueger, A.B., & Zhu, P. (2004). Another look at the New York City school voucher experiment. *American Behavioral Scientist*, 47(5), 658–698.
- Ladd, H.F. (2002). School Vouchers: A Critical View. *Journal of Economic Perspectives*, 16(4), 3–24.
- Lamarche, C. (2010). Measuring the incentives to learn in Colombia using new quantile regression approaches. *Journal of Development Economics*, 96, 278–288.
- Levin, H.M. (2008). Issues in educational privatization. In H.F. Ladd and E.B. Fiske (Eds.), *Handbook of research in education finance and policy* (pp. 391–404). New York, NY: Routledge.
- Mayer, D.P., Peterson, P.E., Myers, D.E., Tuttle, C.C., & Howell, W.G. (2002). *School choice in New York City after three years: An evaluation of the School Choice Scholarships Program: Final report*. Princeton, NJ: Mathematica Policy Research.
- McEwan, P.J. (2004). The potential impact of vouchers. *Peabody Journal of Education*, 79(3), 57–80.
- Metcalfe, K.K., West, S.D., Legan, N., Paul, K., & Boone, W.J. (2003, March). *Evaluation of the Cleveland Scholarship and Tutoring Program* (Technical report 1998–2001). Bloomington, IN: Indiana Center for Evaluation.
- Mills, J.N., & Wolf, P.J. (2017). Vouchers in the Bayou: The effects of the Louisiana Scholarship Program on student achievement after two years. *Educational Evaluation and Policy Analysis*, 39(3), 464–484.
- Muralidharan, K. & Sundararaman, V. (2015). The aggregate effect of school choice: Evidence from a two-stage experiment in India. *The Quarterly Journal of Economics*, 130(5), 1011–1066.
- Neal, D. (2002). How vouchers could change the market for education. *Journal of Economic Perspectives*, 16(4), 25–44.
- Peterson, P.E., & Howell, W.G. (2004). Efficiency, bias, and classification schemes. *American Behavioral Scientist*, 47(5), 699–717.
- Rouse, C.E. (1998). Private school vouchers and student achievement: An evaluation of the Milwaukee parental choice program. *The Quarterly Journal of Economics*, 113(2), 553–602.
- Rouse, C.E., & Barrow, L. (2009, September). School vouchers and student achievement: Recent evidence and remaining questions. *Annual Review of Economics*, 1, 17–42.
- Waddington, R.J., & Berends, M. (2018). Impact of the Indiana Scholarship Program: Achievement effects for students in upper elementary and middle school. *Journal of Policy Analysis and Management*, 37(4), 783–808.
- Witte, J.F. (2000). *The market approach to education*. Princeton, NJ: Princeton University.
- Witte, J.F., Carlson, D., Cowen, J.M., Fleming, D.J., & Wolf, P.J. (2012, February). *MPCP longitudinal educational growth study: Fifth-year report* (Report 29). Fayetteville, AK: University of Arkansas, Department of Education Reform.
- Witte, J.F., Wolf, P.J., Cowen, J.M., Carlson, D.E., & Fleming, D.J. (2014). High-stakes choice: Achievement and accountability in the nation's oldest urban voucher program. *Educational Evaluation and Policy Analysis*, 36(4), 437–456.
- Wolf, P.J., Gutmann, B., Puma, M., Kisida, B., Rizzo, L., Eissa, N., & Carr, M. (2010). *Evaluation of the D.C. Opportunity Scholarship Program: Final report* (NCEE 2010–4018). Washington, D.C.: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.
- Wolf, P.J., Kisida, B., Gutmann, B., Puma, M., Eissa, N., & Rizzo, L. (2013). School vouchers and student outcomes: Experimental evidence from Washington, D.C. *Journal of Policy Analysis and Management*, 32(2), 246–270.