

School vouchers: Recent findings and unanswered questions

Lisa Barrow and Cecilia Elena Rouse

Introduction and summary

Many people would argue that U.S. elementary and secondary public schools need to improve, and they would like to see U.S. students perform better in international comparisons. In addition, many people would like to do more to close the achievement gaps within the U.S. between lower-income and minority students and their counterparts. These concerns are shared by parents, employers, policymakers, and society more generally. There is less agreement on the root causes of the problems and how best to tackle them. For example, while some would argue that insufficient funding is the primary factor, others would say that this is not supported by the evidence, since per pupil spending has increased faster than achievement, as reflected in scores on the National Assessment of Educational Progress (NAEP).¹

One strategy for improving school performance that has received a lot of attention by all those interested in education policy is increased competition. The idea is that just as competition can enhance efficiency and value in the marketplace for goods more generally, it can do the same for education. Namely, if schools must compete for students, then they will take steps to ensure that the educational experiences they offer are valued by parents and students. The primary mechanism proposed by those who favor more competition in elementary and secondary education in the public sector is an education voucher—a coupon redeemable for a maximum dollar amount per child if spent to attend a private school. In this way, voucher programs remove the monopoly power of local public schools. Instead of having to attend a neighborhood public school, students can use the voucher to attend a private school.

In the 2007–08 school year, roughly 55,000 students in three states—Florida, Ohio, and Wisconsin—and the District of Columbia were using publicly

funded education vouchers to attend private schools (as well as other higher-performing public schools);² see table 1. Several other states have considered voucher programs, and some private organizations have helped create privately funded voucher opportunities. But are voucher programs effective? Do they improve the educational outcomes of the students who use them and do they improve the quality of the public schools? In this article, we review the existing empirical evidence on the impact of school vouchers on student achievement.³ After reviewing the research, we conclude that expectations about the ability of vouchers to drastically improve student achievement, at least as measured by test scores, should be tempered by the results of the studies to date. That said, many questions remain. For example, no studies have examined the longer-run impact of vouchers on outcomes such as graduation rates, college enrollment, and future wages. Similarly, the research designs for identifying the potential impacts on students who remain in the public schools are far from ideal. Finally, we have little understanding of whether vouchers would represent a cost-neutral alternative to our current system of public education provision at the elementary and secondary school levels.

Lisa Barrow is a senior economist at the Federal Reserve Bank of Chicago. Cecilia Elena Rouse is director of the Education Research Section, director of the Industrial Relations Section, and Theodore A. Wells '29 Professor of Economics and Public Affairs at Princeton University; she is also a research associate at the National Bureau of Economic Research. The authors thank Clive Belfield, David Figlio, Bhashkar Mazumder, Patricia Muller, Anna Paulson, and Jonathan Plucker for helpful comments and conversations. Emily Buchsbaum and Mitta Isley provided expert research assistance. They also thank the Ohio Department of Education and Washington Scholarship Fund for responding to their queries. Parts of this Economic Perspectives article are drawn from the authors' forthcoming article in the Annual Review of Economics.

In the next section, we discuss the theoretical reasons for why education vouchers should improve student achievement followed by a discussion of the empirical approaches used for identifying the effects of vouchers. We then review the best evidence from studies of publicly and privately financed school voucher programs on the short-run impact on student achievement, discuss the evidence on the potential impact on students who remain in the public schools, consider additional voucher outcomes, and conclude.

Why should competition improve our educational system?

The idea of injecting competition into the public school system is not new; for example, Milton Friedman (1962) argued in the early 1960s for separating the financing and provision of public schooling by issuing education vouchers. The rationale behind school vouchers is that competitive markets allocate resources more efficiently than do monopolistic ones and that public schools in the U.S. have “monopoly” power because children are assigned to attend their local neighborhood school. Parents can always choose to send their children to a private school, but that means paying for schooling twice—once through property taxes (for the public schooling they are not using) and again through private school tuition. If parents had more publicly funded options for their children’s schooling—once they had selected a residential location—then schools would have to compete for students. Further, more options may improve the match between the educational interests and needs of students and their schools. Importantly, schools in this model would have an incentive to improve in the areas valued by parents. Thus, if parents select schools based on their academic quality, then schools will compete for students by providing better academics; alternatively, if parents value religious education or sports, then one would expect to see schools compete to serve these interests.

There are two hypothesized ways by which increased school choice would improve student educational outcomes. The first is a “direct” effect for those students who actually exercise choice. Assuming that students would only choose to attend a school other than their neighborhood school if that school were academically *better* (or a better match), then the academic achievement of students who opt for a different school should improve relative to what their performance would have been had they stayed in the neighborhood public school. In addition, there is an “indirect,” or “general equilibrium,” effect on students remaining in the public schools. Competition should induce the

public schools to improve in an effort to attract (or retain) students. Thus, not only should the achievement of those who choose to attend private schools increase, but so should the achievement of those who do *not* choose as well. In other words, competition should increase the efficiency of public schools. Of course, expansion of the private sector is a critical component of increasing competition. Without new school entries and/or increases in the size of current private schools, vouchers would have limited ability to increase choice.

Many empirical studies find that students in private schools have higher educational achievement levels than those in public schools.⁴ The findings from such studies are presented by voucher advocates as *prima facie* evidence that vouchers would improve student achievement for all. Namely, voucher advocates argue that private schools outperform public schools because their existence depends on providing a good product. Educational vouchers are intended to make public schools compete in this same way; so, only schools (either public or private) providing a good product would survive. However, this literature is not conclusive because of the difficulty (described later) in identifying the impact of schools on student achievement. Not surprisingly, critics argue that the observed superiority of private schools in these studies arises because the students who attend private schools differ from the students who attend public schools rather than because private schools are more effective than public schools.⁵ If the observed relative superiority of the private sector is due more to the particular background characteristics of its students than the greater effectiveness of its schools, the achievement of current public school students would not necessarily improve in private schools.

While the debate continues on whether private schools, in general, are better at educating children than are public schools, researchers have since turned to more direct evidence on the impact of vouchers by studying actual school voucher programs. We begin with an overview of the challenges of testing whether vouchers improve student outcomes before reviewing the evidence to date.

Empirical approaches to studying school vouchers

To study whether educational outcomes in the presence of vouchers are better than educational outcomes in the absence of vouchers, ideally, one would begin with a group of students and educate them for a period of time under the current public school system. At the end of the period, one would assess various

TABLE 1

Description of publicly funded voucher programs in 2007–08

Voucher program	Start of program	Number of scholarship students in 2007–08	Other information
EdChoice Scholarship Program (Ohio)	2006	6,580	<p>Students are eligible to apply if they currently attend or will be entering an EdChoice-designated public school in the upcoming year or if they currently attend a charter school but would otherwise be assigned by the school district to an EdChoice-designated school.</p> <p>There is currently a maximum of 14,000 scholarships that can be awarded across the state, and there are no income requirements to receive a scholarship. If more than 14,000 students apply, then students who are renewing their scholarships followed by students who are at or below 200 percent of the poverty level receive priority.</p> <p>The program pays the minimum of a school's tuition or \$4,375 per student from kindergarten through eighth grade and \$5,150 per student from ninth through 12th grades.</p>
Cleveland Scholarship and Tutoring Program (Ohio)	1996–97	6,017	<p>Any student living within the boundaries of the Cleveland Metropolitan School District (CMSD) and entering kindergarten through eighth grade is eligible to apply. Low-income students are given priority, and scholarships are awarded by lottery drawings. A student is not eligible to apply once he or she has entered high school, although scholarships have been made available to program participants once they reach high school.</p> <p>All children currently attending Cleveland Metropolitan Schools in kindergarten through 12th grade are eligible for the tutoring program.</p> <p>The scholarship program pays either 75 percent or 90 percent of a school's tuition (depending on family income) not exceeding \$3,450 for the 2007–08 school year.</p>
A+ Opportunity Scholarship Program (Florida)	2000	1,305 in public schools (see next column)	<p>Students are eligible for the A+ Opportunity Scholarship Program through the highest grade of their public school if their school is currently "failing" (graded an F) for its second year in a four-year period. The scholarships could continue into high school if the high school was assigned a grade below C.</p> <p>The program made an average scholarship payment of \$4,206 per student in 2005–06.</p> <p>These scholarships (vouchers) for private schools were declared unconstitutional by the Florida Supreme Court in January 2006. Thereafter students could no longer use these vouchers to attend participating private schools; they are, however, still able to use the vouchers to attend higher-graded public schools.</p>

outcomes for the students and administer a test, the results of which would perfectly reflect what the students know. Then one would turn back time so that the same group of students was reverted to their conditions at the beginning of the experiment. That is, the students would be the same age, have the same living conditions, and so on. This time, one would educate the students in a system with education vouchers. At the end of the same period of time, one would again assess the students' outcomes and what they know. The difference between the students' outcomes under the current and voucher systems would isolate the impact of vouchers because vouchers would be the only difference at the beginning of each experimental period.⁶ If implemented on a small scale, this experiment

would allow one to estimate the direct effect of vouchers; if implemented on a large scale, this would uncover the potential effects on all students, including those who remain in the public schools.

Obviously, such an evaluation is not possible. So, to study the direct effect of vouchers, researchers must rely on comparing the achievement (or other outcomes) of students who were offered a voucher (or actually used a voucher to attend a different school) with the outcomes of students who were denied a voucher, were ineligible for a voucher, or remained in the public schools for other reasons.⁷ The empirical challenge is that the outcomes of the nonvoucher students may not provide a valid approximation of what

TABLE 1 (CONTINUED)

Description of publicly funded voucher programs in 2007–08

Voucher program	Start of program	Number of scholarship students in 2007–08	Other information
McKay Scholarships for Students with Disabilities Program (Florida)	1999–2000	19,439	To be eligible for this program, students must have been part of the Florida public school system for at least one year, have been counted in the prior school year’s October and February enrollment surveys, and have an individual education plan. For the 19,439 students currently enrolled, there has been a total payment amount of \$99,212,622.04 so far for the 2007–08 school year. There was a total payment amount of \$119,092,631.54 across the state for the 2006–07 school year, assisting 18,273 students. Scholarship amounts are awarded on an individual basis.
Milwaukee Parental Choice Program (Wisconsin)	1990	18,882	Milwaukee Parental Choice Program students have family incomes at or below 175 percent of the federal poverty level (\$35,843 for a family of four in 2007–08). Once a student is in the program, family income may rise to 220 percent of the federal poverty level (\$45,057 for a family of four in 2007–08). The voucher was worth a maximum of \$6,501 in 2007–08.
DC Opportunity Scholarship Program (District of Columbia)	2004	1,903	Eligibility requirements give preference to students in kindergarten through 12th grade who are eligible for free or reduced lunch and who are enrolled in public schools that fail to make sufficient yearly progress as defined by the No Child Left Behind Act. The program offers scholarships for up to \$7,500 to cover the costs of tuition, fees, and transportation for students to attend participating DC private schools. Scholarships are renewable for up to five years as long as students remain eligible and in good academic standing at their private schools.

Sources: Ohio Department of Education, <http://EdChoice.ohio.gov> and www.ode.state.oh.us/gd/gd.aspx?page=2&TopicRelationID=672; Florida Department of Education, www.floridaschoolchoice.org; Wisconsin Legislative Fiscal Bureau, www.legis.state.wi.us/lfb/Informationalpapers/29.pdf; Washington Scholarship Fund, www.washingtonscholarshipfund.org; and Wolf et al. (2007).

would have happened to the voucher students had they not been offered or used a voucher. The voucher students may not have done as well (or as poorly!) as the nonvoucher students in the absence of the voucher program because their individual characteristics differ substantially. For example, the students with parents who are very educationally focused and motivated may be more likely to apply for a school voucher, and yet these students may have done better than their nonvoucher classmates even in the absence of the voucher program because of their higher level of parental support and encouragement. Unless the research design can take these individual characteristics fully into account, the estimated impact of vouchers will not likely generate the true impact of vouchers on student achievement (in statistical terms the estimated impact will be “biased”). As a result, researchers have relied on analytical strategies that attempt to control for all differences between the two groups of students, observed and unobserved.

The first strategy is to control for students’ achievement (typically a test score) prior to using the school voucher in order to adjust for nonschool factors (such

as having educationally focused parents or low family income) that affect this achievement and that might also be correlated with a student’s likelihood of applying for a voucher. This strategy amounts to comparing the change in achievement of students before and after participation in the voucher program with the change in achievement of students who did not participate in the program. The assumption that must hold for this estimate to generate the true (unbiased) effect of vouchers is that there are no other differences that would explain changes in the test scores between the two groups of students except for the use of a voucher. Although this approach has some appeal, one might reasonably be concerned that students who were not doing well (or were doing very well) in the public schools were more interested in a different schooling experience and that prior test scores do not perfectly reflect academic “ability,” achievement, or motivation.

A second, more compelling strategy to generate estimates of the effect of vouchers on student outcomes is to use a random assignment design. In this “experimental” research design, students are randomly assigned to either a “treatment group” that is offered a school

voucher or a “control group” that is not. In this case, there are no differences in the observed or unobserved individual characteristics, on average, between the two groups because the offer of a voucher was determined not by one’s family’s income or motivation, but purely by the flip of a coin. As a result, one need not control for other student characteristics. Properly implemented, such a strategy is viewed as the “gold standard” for estimating a causal relationship between vouchers and student outcomes. In practice, however, nonrandom differences can emerge between the treatment and control groups to the extent that the researchers conducting the study are not able to adequately follow up with both groups of students.

A simple comparison of the outcomes of students in the treatment group (those offered vouchers) and those in the control group (those not offered vouchers) will generate a true (unbiased) estimate of the impact of being offered a voucher on student outcomes—a parameter known as the “intention to treat” in the research literature. This impact reflects two factors that are important for proper evaluation of a voucher program: the rate at which students who are offered a voucher actually use one and the relative achievement of students in private schools. As such, the intention to treat has two appealing properties: It is the only unambiguously true (unbiased) estimate that one can obtain using typical statistical methods, such as ordinary least squares regression, and it reflects the overall potential gains from offering the vouchers as a policy (since those who are offered vouchers cannot be compelled to use them).

Many are also interested in whether students who actually use vouchers experience academic gains as a result—an effect known as the “treatment on the treated” among researchers. Because actual use of a voucher is not randomly determined, analysts must resort to nonexperimental methods to generate consistent estimates of the treatment-on-the-treated gains by those who actually use vouchers to attend private school. A common approach is to use an instrumental variables strategy: Whether a student was randomly offered a voucher is used as an instrumental variable for the student attending a private school. This type of analysis would generate a consistent estimate of whether the schools that the voucher students attended were more, less, or equally as effective as the schools that the nonvoucher students attended.

Evidence on the direct impact of school vouchers on students

In the U.S., two types of school voucher programs have been studied—those financed by the government (publicly funded school vouchers) and those provided

by the private sector (privately funded school vouchers). From a public policy perspective, the evidence from publicly funded programs is most relevant as these programs incorporate some of the design features that might be built into a larger school voucher program, such as limitations on which students are eligible to receive a voucher and the provision or reimbursement of transportation (if any at all). That said, some of the most compelling evidence (from a methodological perspective) comes from the privately funded vouchers, so we review that evidence here as well. We begin with evidence from publicly funded programs.

We translate the estimated impacts for all of the programs into “standardized effect sizes” (σ) in order to compare estimates across studies. In particular, the estimated difference in test scores between voucher and comparison (nonvoucher) students has been divided by the standard deviation of the test score from a national sample of students. The reason for dividing by the standard deviation is to account for the fact that studies have used different tests to assess the students. The problem is that an assessment of whether a gain is “big” or “small” depends on the shape of the underlying distribution of the test. Thus, for example, a five-point gain using a test that has a narrow bell shape (a small standard deviation) implies a larger gain in student learning than does an eight-point gain using a test that has a wide bell shape (a large standard deviation). Thus, researchers “standardize” the test score gain by the spread of the distribution to account for its underlying shape and often report “effect sizes” in standard deviation units.

Once standardized, however, one must still judge whether an estimated effect size is large or small. Recently, Hill et al. (2007) attempted to review effect sizes from many studies of educational interventions. While they caution that it is only valid to compare effect sizes when using comparable populations, contexts, and interventions, as well as the outcomes being measured, they report that effect size estimates from randomized studies average 0.33σ for elementary schools, the typical grade level for the studies of voucher programs we review here.

Table 1 (pp. 4–5) briefly describes publicly funded voucher programs in the U.S. Since the launch of the Milwaukee Parental Choice Program in Wisconsin in the early 1990s, several other publicly financed voucher programs have been started, including one in Washington, DC, in which the vouchers are allocated on a randomized basis. Not only is the Milwaukee Parental Choice Program one of the oldest publicly funded voucher programs in the U.S., it has also been subject to numerous studies. The program is open to

low-income students who may use a voucher to attend any participating school (including religious schools) worth approximately \$6,501 in the 2007–08 academic year. Nearly 19,000 students and 120 schools participated that academic year.

Most of the studies regarding potential achievement impacts of the Milwaukee program were conducted when the program had only been in operation for about four years and vouchers could only be used at nonreligious schools. At that time, about 12 schools and 800 students participated. Because the schools participating in the program were required to take all students who applied and to randomly select among applicants in the event of oversubscription, researchers had two potential comparison groups available—unsuccessful applicants and a random sample of low-income students from the Milwaukee Public Schools. Using both comparison groups, Rouse (1998) reports mixed results of the “direct” effect of the program: She estimates intent-to-treat effect sizes ranging from 0.06σ to 0.11σ in math and from -0.03σ to 0.03σ in reading, although the impacts in reading are never statistically different from zero, meaning that the difference may have arisen by chance.⁸

Evidence from the Cleveland Scholarship and Tutoring Program (CSTP) suggests even smaller impacts on student outcomes. The voucher program is open to all students living within the boundaries of the Cleveland Metropolitan School District, with preference given to students in low-income families.⁹ Students are permitted to use the vouchers at both nonreligious and religious schools. (The tutoring program provides tutors to interested students from kindergarten through 12th grade.) The CSTP data allow researchers to identify three groups of applicants: voucher recipients who use the voucher, voucher recipients who do not use the voucher, and nonrecipients.¹⁰ Additionally, Cleveland Metropolitan School District and test score data are available for a (non-random) sample of public school students.¹¹

Analyzing data from the cohort of students who entered kindergarten in 1997, Belfield (2007) compares voucher winners and rejected applicants with the available sample of Cleveland public school students. He also estimates the effect of attending private school by comparing voucher users with the rejected applicants. In the third year of the program (when the cohort is in second grade), he finds that voucher winners scored significantly lower in math (-0.08σ) and lower in reading (-0.05σ) than those in the public school sample. Further, he finds that voucher users scored significantly lower in both math and reading (-0.11σ and -0.13σ , respectively) than the rejected applicants.¹²

In the fifth year of the program (when the cohort was in fourth grade), the results are more mixed with estimated effects ranging from -0.08σ in math to 0.07σ in reading for the effect of using a voucher, but neither estimate is statistically different from zero.¹³

While the studies of both Milwaukee and Cleveland attempt to construct valid comparison groups and thereby identify causal impacts of the voucher programs on student outcomes, all of them rely on observational data and therefore may not fully account for pre-existing differences between the voucher and comparison groups. This leads to biased estimates of the impact of vouchers. In the case of Milwaukee, the bias could be either positive (in that the students who participated in the voucher program were more motivated) or negative (in that the random sample of low-income students in the public schools was *too* advantaged relative to the voucher participants). While Rouse (1998) attempts to determine the extent of any such bias (and concludes it is likely minimal), it remains an untestable assumption. Belfield (2007) is subject to the same general research design concern.¹⁴

This methodological concern could, in theory, be addressed in the study of the relatively new DC Opportunity Scholarship Program (DC OSP) in Washington, DC, which is being evaluated using a random assignment program design.¹⁵ In the first two years of the program (2004 and 2005), 2,038 eligible public school students participated in the lotteries; 1,387 of them were awarded scholarships, and the remaining 921 students became the “control group.” Wolf et al. (2007) estimate that after one year, intent-to-treat effect sizes for the first two cohorts of students ranged from -0.01σ to 0.08σ in math and from -0.01σ to 0.03σ in reading. After two years, Wolf et al. (2008) report that the impacts ranged from -0.02σ to 0.01σ in math and from 0.05σ to 0.08σ in reading. Not only do these ranges include negative impacts but none of them are statistically different from zero at the 5 percent level.

Thus far, the evidence from the publicly funded voucher programs suggests, at best, mixed improvement among either students who were selected for a voucher (the intent to treat) or students who used a voucher (the treatment on the treated). The largest estimates, from the Milwaukee Parental Choice Program, suggest potential (intent-to-treat) gains of 0.11σ in math and gains of 0.14σ for those who actually use a voucher to attend a private school; most of the other estimates are much smaller or even negative. However, with the exception of the studies on the DC OSP, the studies suffer from potentially unsatisfactory comparison groups. As such, we now turn to evidence from the privately funded programs.

Although a fairly recent U.S. General Accounting Office (2002) report found 78 privately funded voucher programs to review, only a handful have been subject to any evaluation.¹⁶ Three privately funded voucher programs—based in New York City; Dayton, Ohio; and Washington, DC—had randomized study designs, making them the best suited for rigorous evaluation. As in the publicly funded DC OSP, the privately funded programs in each city had greater numbers of applicants than vouchers available; therefore, applicants were randomly selected to receive or not to receive a voucher offer. For the New York City program (School Choice Scholarships Foundation), for example, the number of applicants was so large that the “control group” is made up of a random sample of applicants not selected to receive a voucher. We briefly describe each of the privately funded voucher programs with a randomized study design in table 2.

Across all three cities, Howell, Peterson, et al. (2002) find that using a voucher has no overall impact on student test scores. Indeed, after three years the estimated impact of attending a private school is only 0.02σ . Similarly, both Mayer et al. (2002) and Krueger and Zhu (2004) report very small impacts (at most a 0.06σ impact for using a voucher) in any year for the program in New York City, and none of the estimates are statistically distinguishable from zero.

Nevertheless, a widely publicized result from these privately funded programs is that there may have been differences across subgroups of students. Indeed, Howell, Peterson, et al. (2002) and Mayer et al. (2002) report statistically significant positive effects of private school attendance on test scores for African American students alone. Also, after three years, those African American students who used a voucher are estimated to have experienced a 0.23σ gain in achievement across the three cities; those African American students who used a voucher from the New York City program are estimated to have gained 0.26σ .¹⁷

However, these results are not robust. In their re-analysis of the data from the New York City program, Krueger and Zhu (2004) report that the results by race are particularly sensitive to two analytical decisions. First, Krueger and Zhu include all students, whereas Mayer et al. (2002) include baseline test scores in all of their specifications, which leads them to exclude the students who were missing baseline test score information; most of the excluded are first grade students who were not administered a baseline test. Because students were randomly chosen to receive or not to receive a voucher, baseline characteristics such as test scores should have been identical for the two groups (on average). The primary reason for including baseline

characteristics would be to improve the precision of the estimates. However, Krueger and Zhu find very little difference in the precision of the estimated impact of vouchers when using the larger sample without baseline test scores. As a result, they argue that the gain in terms of statistical precision is not great enough to warrant the cost in terms of not generating estimates that are representative of the original target population.

The second substantive difference between the studies is how the researchers define a student’s race. Mayer et al. (2002) identifies a student as African American if the mother’s race is reported as African American and non-Hispanic (*irrespective* of the race or ethnicity of the father). Krueger and Zhu (2004) use alternative identifications. They identify a student as African American if either parent is African American and non-Hispanic; also, in their definition of African American, they include the group of students whose parents responded “other” to the survey but indicated that they (the parents) were “black” in the open-ended response. With the larger sample and the broadest identification of students as African American, Krueger and Zhu report that the estimated impact of being offered a voucher (intent-to-treat impact) for African American students falls to 0.05σ after three years and the estimated impact of using a voucher (treatment on the treated) falls to 0.03σ ; neither estimate is statistically different from zero.

In sum, there is little evidence of overall improvement in test score outcomes for students offered an education voucher from privately funded voucher programs. Although there may be evidence that some subgroups of students benefit from being offered a voucher, the evidence is not robust to sensible alternative ways of constructing the analysis sample. In addition, the results of these experiments may not be valid for thinking about the average benefits of offering vouchers to all students. Namely, all participants in the experiment—both voucher recipients and non-recipients—had expressed an interest in vouchers.

Evidence of public school response to competitive pressure

As we have emphasized, the studies discussed previously are based on relatively small voucher programs such that there was likely little competitive pressure to which the public sector would have responded. As such, the estimates are primarily of the direct effect of vouchers for those who use them. However, the true prize of a voucher system—or any significant increase in the competitive pressure experienced by the public schools—is overall improvement in the

TABLE 2

Description of privately funded voucher programs

Voucher program	Evaluation period	Number of scholarship students in program evaluation	Other information
School Choice Scholarships Foundation (New York City, NY)	1997–99	1,200	<p>Students were eligible to apply if they were entering first through fifth grades, currently attending a public school, and qualified for the National School Lunch Program.</p> <p>The program began in 1997, paid up to \$1,400 annually, and initially guaranteed three years of receipt. The scholarships were later extended beyond the initial three years.</p>
Parents Advancing Choice in Education, or PACE (Dayton, OH)	1998–99	530	<p>Students in kindergarten through 12th grade whose family income was less than 200 percent of the federal poverty line were eligible. (Students currently enrolled in private school were eligible for the program but not included in the evaluation. The evaluation focuses on students in the first through seventh grades at baseline.)</p> <p>At most, the scholarship was worth \$1,200 or 60 percent of tuition, whichever was less. However, voucher awards were smaller for higher-income families. The program began in 1998, and while the evaluation followed students for only two years, PACE continued to award new scholarships through 2008. In 2008 the average scholarship is worth \$1,800, and students are guaranteed a scholarship for at least four years.</p>
Washington Scholarship Fund (Washington, DC)	1998–2001	1,000	<p>Students entering kindergarten through eighth grade whose family income was less than 270 percent of the federal poverty line were eligible. (Students currently enrolled in private school were eligible for the program but not included in the evaluation. The evaluation focuses on students in the first through seventh grades at baseline.)</p> <p>At most, the scholarship was worth \$1,700 or 60 percent of tuition, whichever was less. However, voucher awards were smaller for higher-income families. The voucher program began in 1993 and continued to offer scholarships in 2008. In 2008–09, the scholarships are worth up to \$3,000 per child each year.</p>

Sources: Howell, Wolf, et al. (2002); Parents Advancing Choice in Education, www.pacedayton.org; and Washington Scholarship Fund, www.washingtonscholarshipfund.org.

performance of the affected education system. Unfortunately, to develop a study that would generate unbiased estimates of any such systemic impacts is extremely difficult. One cannot simply compare the outcomes of students who use a voucher (or who were offered a voucher) to the outcomes of students who remained in the public schools (either by choice or from “bad luck” in a lottery) as this would likely underestimate the general equilibrium impact. The problem is that, in theory, the public schools should improve in response to the increased competition and this improvement should be reflected in the achievement of the public school students. As a result, the control (or comparison) students would not adequately represent what would have happened to the voucher students in the absence of the voucher program.

Rather, one would ideally gather a large group of education “markets” (assuming that any general equilibrium impacts remain within a market and there are

no spillovers to others) and randomly assign some markets to a treatment group—in which the students would be eligible for school vouchers—and randomly assign the remaining markets to a control group—in which there would be no vouchers. After some period of time, the researcher would then compare the average outcomes of students in the voucher markets with those of students in the control markets. A simple comparison of the outcomes would yield an unbiased estimate of the general equilibrium impact of vouchers because, on average, the markets would have been similar ex ante. While such an experiment is possible in theory, in practice it would be extremely difficult to implement mostly because it would require the coordination and cooperation of so many different stakeholders. As a result, researchers have turned to other research designs to try to get an estimate of the potential impact of a large-scale voucher program.

Evidence from the expansion of the Milwaukee Parental Choice Program

After the experimental phase of the Milwaukee Parental Choice Program ended in 1995, the program was expanded to allow for a maximum of 15 percent of the public school enrollment; further, in 1998 the Wisconsin Supreme Court ruled that the vouchers could be used in religious schools. These two events led to a dramatic increase in participation in the program by both students and schools. In fact, the program was so popular that in 2006, participation was expanded to 22,500 voucher students. Researchers have attempted to analyze these last two expansions to estimate the potential impact of a large-scale voucher program on student achievement in the public sector (see Hoxby, 2003; Carnoy et al., 2007; and Chakrabarti, 2008). While some of the details differ, the basic strategy of all three studies is to attempt to identify those schools within the Milwaukee Public School District that face differing competitive pressure because of the mix of income levels among their students. (Those schools with a high proportion of low-income students who are eligible for the voucher program presumably face more competitive pressure than those with a low proportion of low-income students who are eligible.) The basic strategy of all three studies also identifies observably comparable districts elsewhere in Wisconsin in which there are no publicly funded vouchers. The following would be evidence of a positive impact of competition on school efficiency, as reflected in student test scores: Disproportionate gains among students attending schools facing competitive pressure compared with their peers at schools within Milwaukee facing relatively little pressure and at schools outside of Milwaukee (facing no pressure from vouchers).

All three studies find evidence that with the expansion of the voucher program in 1998, student performance improved in the first few years, especially in schools that were most likely to be affected by the increased competition. For example, Hoxby (2003) estimates that the fourth grade test scores of students attending schools likely facing the most competitive pressure improved by 0.12σ per year in math and by 0.07σ per year in reading relative to students attending comparison schools outside of Milwaukee.

While interesting, these results must be interpreted as being only suggestive. The identifying assumption is that there are no unobserved changes before and after the voucher program was implemented when comparing the schools with many voucher-eligible students to schools with few or no voucher-eligible students. However, within the Milwaukee Public School District,

all schools were potentially affected by the vouchers. Further, outside of the Milwaukee Public School District, the demographic composition of the schools is quite different (specifically, the students are less likely to be minority and more likely to come from wealthier families) such that it is not clear researchers can adequately account for differences between the students. In addition, Carnoy et al. (2007) present some results that are not consistent with a simple interpretation that performance in the Milwaukee public schools improved because of increased competition. For example, they also find that there was little improvement after 2002 despite the fact that interest in the voucher program increased (as proxied by the number of applications). Further, they find no evidence of a general equilibrium impact when they employ other direct measures of competition (such as the number of nearby private schools or the relative number of voucher applications from a school).

Evidence from Florida's A+ Opportunity Scholarship Program

In order for a voucher program to spur improvement within the public schools, there need not be a substantial number (or proportion) of students who use a voucher to attend a private school. Rather, if public school administrators perceive there is the potential that the students will do so, they may have an incentive to improve the education in their schools. Thus, researchers have attempted to gain some insight into the potential response of public schools to increased competitive pressure a second way: by studying the schooling outcomes of students attending schools that were under the "threat" of becoming voucher-eligible—that is, schools with a high probability of their students becoming eligible to use a voucher. Researchers have done so by taking advantage of the design of Florida's school accountability system—its A+ Plan for Education. Specifically, since 1999, schools in Florida are given a grade of A through F, largely depending on the performance of the students. Schools that receive high grades and are improving receive bonuses. In contrast, low-performing schools (graded either D or F) are subject to increased administrative oversight. (These poor performers are also provided with some additional financial assistance.) In addition, if a school received an F in two out of four years and had an F in the current year, students became eligible for vouchers called Opportunity Scholarships.¹⁸ While the other features of Florida's A+ Plan for Education remain in effect, the voucher program was declared unconstitutional by the Florida Supreme Court in January 2006. Thereafter students

could no longer use a voucher to attend a participating private school; they are, however, still able to use a voucher to attend a higher-graded public school.

Under Florida's A+ Plan for Education, school grades are determined by assigning "grade points" based on student test score performance.¹⁹ Grades are then assigned based on whether the school is above or below the predetermined cut points for each of the letter grades. Arguably, schools earning just above the number of grade points needed to receive an overall grade of D are no different than schools receiving just below the number of grade points needed to receive a D grade. As a result, many of the schools that received an F grade are quite similar to many of those that received a D grade. Figlio and Rouse (2006), West and Peterson (2006), Rouse et al. (2007), and Chiang (2008) therefore compare student outcomes from schools earning D and F grades while controlling for the number of grade points earned so that they can recover the causal effect of the policy on educational achievement.

All of the papers find that test scores of students improve following a school's receipt of an F grade. For example, Rouse et al. (2007) and Chiang (2008) report gains ranging from 0.12 σ to 0.14 σ in math and about 0.10 σ in reading. Further, these two studies also find evidence that the improvements persist even once the students leave the voucher-threatened school, particularly in math. In addition, Rouse et al. (2007) report finding evidence that the F-graded schools responded in educationally meaningful ways. For example, following receipt of an F grade, schools were more likely to focus on low-performing students, lengthen the amount of time devoted to instruction, and increase resources available to teachers. As such, these studies may provide some evidence that increased competitive pressure can generate some improvement in public schools.²⁰

One should note, however, that the F-graded schools in Florida were also stigmatized as "failing" (one of the intents of the public announcements of the grades). So another possibility is that the stigma of being identified as a failing school (and perhaps the subsequent parental pressure to make changes) led the schools to improve. As such, one cannot strictly distinguish a "voucher effect" from a "stigma effect." That said, Figlio and Rouse (2006) indirectly assess the impact of stigma by comparing student achievement following the implementation of Florida's A+ Plan for Education—which enlisted both the threat of vouchers and stigma—with student achievement following the placement of schools on a critically low performers list in 1996, 1997, and 1998 that involved public

stigma but no threat of vouchers. They estimate that the student gains in reading were nearly identical under the two regimes and were actually larger in math following placement on the critically low performers list, suggesting that the relative improvements among the low-performing schools may have been due more to stigma than to the threat of vouchers.

There is some evidence from the expansion of the Milwaukee Parental Choice Program and from the threat of vouchers created by Florida's A+ Plan for Education suggesting that the achievement of students attending schools facing increased competition improves. However, the research strategies do not allow one to definitively rule out other explanations for the improvements. As such, we conclude that the jury is still out on the potential for vouchers to spur public schools to improve.

Other potential social gains from vouchers

There may be other reasons why providing school vouchers may be appealing from a public policy standpoint. One might argue in favor of vouchers as a way to increase equity by giving poor families more opportunities to choose private schools over their neighborhood public schools. Also, based on parents' reports for the publicly funded DC voucher program (DC OSP), the schools that are chosen (private schools) may be safer. Parents of students offered a voucher reported a significantly lower level of perceived school danger than parents of students not offered a voucher.²¹

In a related fashion, student achievement may not be the only criterion by which to judge the success of voucher programs. If school choice means that parents are more satisfied with the education their children are receiving and if voucher programs are no more expensive than our current system, then a voucher program may be a cost-neutral way to increase social welfare. Importantly, one consistent finding in this literature is that voucher parents report being more satisfied with their current schooling than do nonvoucher parents. For example, in the DC OSP, parents of students offered a voucher gave their children's schools a significantly higher overall grade on a five-point scale (grades A through F) and were significantly more likely to give their children's schools a grade of A or B. Further, they reported significantly greater satisfaction with their children's schools on all aspects asked, including location, class sizes, discipline, academic quality, and the racial mix of the students (Wolf et al., 2007). These results have generally been reported for other voucher programs, such as those in New York City (Mayer et al., 2002) and Milwaukee (Witte, Sterr, and Thorn, 1995).²²

Yet, the potential net improvement in social welfare depends on both the general equilibrium effects of vouchers and the cost advantage over current public schools—two issues that are not well understood. While small-scale voucher programs indicate that parents offered a voucher are more satisfied with their children’s schools than those not offered a voucher, a large-scale voucher program might result in some parents who are more satisfied and some who are less satisfied. In order for social welfare to be increased with a cost-neutral voucher program, the gains to the parents who benefit must be large enough to outweigh the losses to parents who do not benefit.

In addition, there is not much information about whether a well-developed voucher program would, indeed, be cost-neutral. On its face an education voucher system should be no more expensive than the current system as the state (or some other public entity) would simply send a voucher check to participating schools for each participating child rather than to the local public school or district. However, if truly implemented on a large scale, there may be other, less obvious costs that would depend critically on the actual design of the program. Levin and Driver (1997) caution that, depending on a number of factors, the cost of a voucher system could actually exceed those of the current geographically based system. These factors include the transportation of children to and from school, recordkeeping, and the monitoring of student enrollment. Two additional concerns are how a program deals with students currently attending private schools and how disputes are adjudicated (particularly if there are differing voucher amounts). While Levin and Driver’s estimates are rough, based on hypothetical voucher programs and crudely estimated costs, their analysis suggests, at a minimum, that we should not assume a voucher program would be cost-neutral. Further, there may be large costs associated with the transition to a voucher system that should be considered.

Finally, the studies to date necessarily focus on short-run effects of vouchers when in fact there may be longer-run impacts on high school graduation, college enrollment, or even future earnings. For example, Altonji, Elder, and Taber (2005a) study the effect of Catholic education on a variety of outcomes and find little evidence that Catholic schools raise student test scores. At the same time, their results suggest that Catholic schools increase the probability of graduating from high school and potentially the probability of enrolling in college. These longer-run effects have yet to be credibly examined in studies of school vouchers.

Conclusion

The best research to date finds relatively small achievement gains for students offered education vouchers, most of which are not statistically different from zero, meaning that those gains may have arisen by chance. Further, the very little evidence about the potential for public schools to respond to increased competitive pressure generated by vouchers also suggests that one should remain wary that large-scale improvements would result from a more comprehensive voucher system.

So why has it been so difficult for researchers to observe large improvements in student achievement with school vouchers in the U.S.? One explanation may be that schools already compete for students through residential choice such that the public sector does not operate as poorly as perceived by many. Another explanation may be that the education sector does not meet the conditions for perfect competition (Garner and Hannaway, 1982). For example, information on school quality may be costly and difficult for parents to obtain, so having more choice may generate less additional competitive pressure on schools than one would expect in a perfect information environment. Further, education is not a homogenous good. Therefore, while competition for students may make schools more responsive to parents, this may be achieved through changes in other areas of school life, such as religious education or sports, rather than academic achievement.

Despite the heretofore lackluster empirical findings, the theoretical rationale behind school vouchers remains compelling: If parents choose schools based on academic performance and if we allow them more choice, then the schools will need to improve academically in order to attract students. In addition, others have endorsed vouchers to promote greater equity: If rich families have the means to opt out of the public school system, should not poor families have a similar opportunity? It is perhaps for these reasons—combined with frustration that other approaches to improve the U.S. education system have proven weak or futile—that school vouchers remain high on the agenda for many policymakers.²³ However, expectations about the ability of vouchers to drastically improve student achievement, at least as measured by test scores, should be tempered by the results of the studies to date, and arguments for vouchers as a cost-neutral alternative should be subject to more careful analysis of the full costs.

NOTES

¹The National Assessment of Educational Progress is the only nationally representative and continuing assessment of what students in the U.S. know and can do in various subject areas, such as mathematics and reading. The commissioner of education statistics, who heads the National Center for Education Statistics in the U.S. Department of Education, is responsible by law for carrying out the NAEP project; for further details, see <http://nces.ed.gov/nationsreportcard/>. See also Hoxby (2003).

²The Florida A+ Opportunity Scholarship Program—a publicly funded voucher program initially created for students to attend private schools—was declared unconstitutional by the Florida Supreme Court in January 2006. After this ruling, students could no longer use the voucher to attend a participating private school; they are, however, still able to use the voucher to attend a higher-rated public school. For further details, see the discussion on the Florida A+ Opportunity Scholarship Program later in the text and in table 1.

³A less efficient public sector and a less competitive (public schooling) environment may explain the larger impacts of school vouchers that have been estimated in other countries, such as Columbia (see, for example, Angrist et al., 2002). In the U.S., elementary and secondary public schooling has largely depended on local financing, meaning that choice between local school districts may already generate strong competitive pressure. As a result, there may be less potential for vouchers to generate large efficiency gains (see, for example, Barrow and Rouse, 2004).

⁴See, for example, Coleman, Hoffer, and Kilgore (1982a, 1982b); Evans and Schwab (1995); Neal (1997); and Altonji, Elder, and Taber (2005b).

⁵See, for example, Goldberger and Cain (1982), Cain and Goldberger (1983), and Altonji, Elder, and Taber (2005a).

⁶Of course, even if vouchers improved outcomes in the long run, there might be a transition period in which the full benefits were not realized. A more complex version of this hypothetical experiment would be needed to identify both the transitional costs and long-run effects of a voucher program.

⁷Ironically, this also means that this literature bears striking similarity to that of the differential effectiveness of private and public schools.

⁸The range reflects estimates from different model specifications. Other studies using these early data from Milwaukee include Witte (1997) and Witte, Sterr, and Thorn (1995), as well as Greene, Peterson, and Du (1999). Using only the sample of low-income students from the Milwaukee Public Schools as a comparison group, Witte (1997) and Witte, Sterr, and Thorn (1995) estimate no impact of the program on student achievement. Greene, Peterson, and Du (1999) only use the unsuccessful applicants as a comparison group and estimate a positive impact in both math and reading. See Rouse (1998) for further discussion of the differences between the studies.

⁹The voucher is also progressive in that it pays 90 percent of tuition up to \$3,450 for those with family income below 200 percent of the poverty line and only 75 percent of tuition up to \$3,450 for those from families earning above 200 percent of the poverty line. The original program paid tuition up to a maximum of \$2,250 (Metcalf et al., 1998). The Cleveland Metropolitan School District changed its name from the Cleveland Municipal School District in 2007.

¹⁰The nonrecipient group potentially contains both students who did not win the voucher lottery and students not entered into the lottery due to the preference given to students from low-income families (Metcalf, 2001).

¹¹The public school sample was generated by using the first grade classmates of voucher recipients who did not use their voucher, as well as the first grade classmates of program applicants who were not awarded a voucher (Metcalf, 2001).

¹²Although Belfield (2007) only reports results for the third and fifth years of the program, he notes that the results are similar for the fourth year when the cohort was in third grade.

¹³Belfield (2007) finds a statistically significant -0.06σ difference in math between voucher winners and the public school sample.

¹⁴In addition, Belfield (2007) includes some measures in his empirical specifications that are arguably outcomes of the voucher program, namely, class size and teacher's years of experience. That said, his results are largely similar when these controls are excluded.

¹⁵See Wolf et al. (2007) for more details. Students attending low-performing public schools were given a better chance of winning the lottery. Although private school students were eligible for the vouchers, they were excluded from the study.

¹⁶The U.S. General Accounting Office's legal name became the U.S. Government Accountability Office on July 7, 2004. For further details, see www.gao.gov/about/namechange.html.

¹⁷In contrast, Howell, Peterson, et al. (2002) estimate a negative impact for African American students after three years in the privately funded voucher program in Washington, DC, although the impact is not statistically different from zero. Results for the third year of the privately funded programs apply only to those in Washington, DC, and New York City because the Dayton, Ohio, program was evaluated for only two years.

¹⁸Currently Florida has two other voucher programs as well: an income tax credit for corporations to fund vouchers for low-income students and the McKay Scholarship for Students with Disabilities Program. Greene and Winters (2008) study the impact of the McKay Scholarships on the achievement gains of students with disabilities who remain in the public schools. Because their estimation strategy identifies the general effect of vouchers by using students whose disability status changes, the extent to which these results generalize to overall improvements in the public schools is unclear.

¹⁹Literally speaking, school grades were not assigned using "grade points" before 2002 when Figlio and Rouse (2006) studied the system. Nevertheless, their strategy is quite similar in spirit.

²⁰A statistical issue with which all of the researchers wrestle is whether the disproportionate gains by students in the F-graded schools resulted from mean-reverting measurement error or reflected actual changes in response to Florida's A+ Plan for Education. Mean-reverting measurement error occurs when gains the year after a school scores unusually low—and is thereby labeled as F—reflect the measurement error in test scores. That is, the test scores of students might have increased in many of the F-graded schools even in the absence of Florida's education plan simply because they were transiently low in the prior year. The reliance on a regression discontinuity design (one that compares the D-graded and

F-graded schools while also controlling for the grade points) helps to mitigate against the presence of mean-reverting measurement error, although the researchers employ other strategies as well.

²¹See Wolf et al. (2007), table H-3. While student perceptions also suggest that the chosen schools are safer on average, the difference was not statistically significant (see table H-4 of the same study).

²²At the same time, not all parents are satisfied with the voucher schools. Focus groups among parents of DC OSP participants found that they believed a few schools misrepresented aspects of their programs and that there was a need for an evaluation of participating schools (Stewart et al., 2007). Similarly, in the early years of the Milwaukee Parental Choice Program, 43 percent of the parents who took their children out of the voucher schools cited the poor

quality of the voucher school as one of the primary reasons they withdrew their children from the program. More specifically, they cited being unhappy with the staff, the education their children were receiving, and the lack of programs for special needs; they also noted that the teachers were too disciplinarian. Thirty percent cited the poor quality of the overall Milwaukee program—including hidden school fees, difficulties with transportation, and the limitation on religious instruction—as the primary reason for withdrawing their children (Witte, Sterr, and Thorn, 1995).

²³Most recently the George W. Bush Administration proposed the strengthening of the choice provisions in the reauthorization of the federal No Child Left Behind Act, and there were (unsuccessful) ballot initiatives in California and Utah to create statewide voucher programs open to all students.

REFERENCES

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber, 2005a, “An evaluation of instrumental variable strategies for estimating the effects of Catholic schooling,” *Journal of Human Resources*, Vol. 40, No. 4, Fall, pp. 791–821.

_____, 2005b, “Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools,” *Journal of Political Economy*, Vol. 113, No. 1, February, pp. 151–184.

Angrist, Joshua, Eric Bettinger, Erik Bloom, Elizabeth King, and Michael Kremer, 2002, “Vouchers for private schooling in Columbia: Evidence from a randomized natural experiment,” *American Economic Review*, Vol. 92, No. 5, December, pp. 1535–1558.

Barrow, Lisa, and Cecilia Elena Rouse, 2004, “Using market valuation to assess public school spending,” *Journal of Public Economics*, Vol. 88, Nos. 9–10, August, pp. 1747–1769.

Belfield, Clive R., 2007, “Achievement effects of the Cleveland Scholarship and Tutoring Program,” City University of New York, Queens College, unpublished mimeo.

Cain, Glen G., and Arthur S. Goldberger, 1983, “Public and private schools revisited,” *Sociology of Education*, Vol. 56, No. 4, October, pp. 208–218.

Carnoy, Martin, Frank Adamson, Amita Chudgar, Thomas F. Luschei, and John F. Witte, 2007, *Vouchers and Public School Performance: A Case Study of the Milwaukee Parental Choice Program*, Washington, DC: Economic Policy Institute.

Chakrabarti, Rajashri, 2008, “Can increasing private school participation and monetary loss in a voucher program affect public school performance? Evidence from Milwaukee,” *Journal of Public Economics*, Vol. 92, Nos. 5–6, June, pp. 1371–1393.

Chiang, Hanley, 2008, “How accountability pressure on failing schools affects student achievement,” Harvard University, mimeo, January.

Coleman, James, Thomas Hoffer, and Sally Kilgore, 1982a, *High School Achievement: Public, Catholic, and Private Schools Compared*, New York: Basic Books.

_____, 1982b, “Cognitive outcomes in public and private schools,” *Sociology of Education*, Vol. 55, Nos. 2–3, April/July, pp. 65–76.

Evans, William N., and Robert M. Schwab, 1995, “Finishing high school and starting college: Do Catholic schools make a difference?,” *Quarterly Journal of Economics*, Vol. 110, No. 4, November, pp. 941–974.

Figlio, David, and Cecilia Elena Rouse, 2006, “Do accountability and voucher threats improve low-performing schools?” *Journal of Public Economics*, Vol. 92, Nos. 1–2, January, pp. 239–255.

Friedman, Milton, 1962, *Capitalism and Freedom*, Chicago: University of Chicago Press.

Garner, W., and J. Hannaway, 1982, “Private schools: The client connection,” in *Family Choice in Schooling: Issues and Dilemmas*, Michael E. Manley-Casimir (ed.), Toronto: Lexington Books, pp. 119–133.

- Goldberger, Arthur S., and Glen G. Cain**, 1982, "The causal analysis of cognitive outcomes in the Coleman, Hoffer, and Kilgore report," *Sociology of Education*, Vol. 55, Nos. 2–3, April/July, 103–122.
- Greene, Jay P., Paul E. Peterson, and Jiangtao Du**, 1999, "Effectiveness of school choice: The Milwaukee experiment," *Education and Urban Society*, Vol. 31, No. 2, February, pp. 190–213.
- Greene, Jay P., and Marcus A. Winters**, 2008, "The effect of special education vouchers on public school achievement: Evidence from Florida's McKay Scholarship Program," Manhattan Institute for Policy Research, civic report, No. 52, April, technical version available at www.manhattan-institute.org/pdf/Effect_of_Vouchers_for_SE_Students_on_Public_School_Achievement_2-19-08.pdf.
- Hill, Carolyn J., Howard S. Bloom, Alison Rebeck Black, and Mark W. Lipsey**, 2007, "Empirical benchmarks for interpreting effect sizes in research," MDRC (Manpower Demonstration Research Corporation), working paper, July.
- Howell, William G., and Paul E. Peterson with Patrick J. Wolf and David E. Campbell**, 2002, *The Education Gap: Vouchers and Urban Schools*, Washington, DC: Brookings Institution Press.
- Howell, William G., Patrick J. Wolf, David E. Campbell, and Paul E. Peterson**, 2002, "School vouchers and academic performance: Results from three randomized field trials," *Journal of Policy Analysis and Management*, Vol. 21, No. 2, Spring, pp. 191–217.
- Hoxby, Caroline M.**, 2003, "School choice and school productivity: Could school choice be a tide that lifts all boats?," in *The Economics of School Choice*, Caroline M. Hoxby (ed.), Chicago: University of Chicago Press, pp. 287–341.
- Krueger, Alan B., and Pei Zhu**, 2004, "Another look at the New York City voucher experiment," *American Behavioral Scientist*, Vol. 47, No. 5, January, pp. 658–698.
- Levin, Henry M., and Cyrus E. Driver**, 1997, "Cost of an educational voucher system," *Education Economics*, Vol. 5, No. 3, December, pp. 265–283.
- Mayer, Daniel P., Paul E. Peterson, David E. Myers, Christina Clark Tuttle, and William G. Howell**, 2002, "School choice in New York City after three years: An evaluation of the School Choice Scholarships program," Mathematica Policy Research Inc., report, No. 8404-045, February 19.
- Metcalf, Kim K.**, 2001, "Cleveland Scholarship Program evaluation: 1998–2000 technical report," Indiana University Bloomington, Indiana Center for Evaluation, mimeo, September.
- Metcalf, Kim K., William J. Boone, Frances K. Stage, Todd L. Chilton, Patty Muller, and Polly Tait**, 1998, "A comparative evaluation of the Cleveland Scholarship and Tutoring Grant Program: Year one: 1996–97," Indiana University Bloomington, Indiana Center for Evaluation, report, March 16.
- Neal, Derek**, 1997, "The effects of Catholic secondary schooling on educational achievement," *Journal of Labor Economics*, Vol. 15, No. 1, part 1, pp. 98–123.
- Rouse, Cecilia Elena**, 1998, "Private school vouchers and student achievement: An evaluation of the Milwaukee Parental Choice Program," *Quarterly Journal of Economics*, Vol. 113, No. 2, May, pp. 553–602.
- Rouse, Cecilia Elena, Jane Hannaway, Dan Goldhaber, and David Figlio**, 2007, "Feeling the Florida heat?: How low-performing schools respond to voucher and accountability pressure," National Bureau of Economic Research, working paper, No. 13681, December.
- Stewart, Thomas, Patrick J. Wolf, Stephen Q. Cornman, and Kenann McKenzie-Thompson**, 2007, "Satisfied, optimistic, yet concerned: Parent voices on the third year of the DC Opportunity Scholarship Program," Georgetown University Public Policy Institute, School Choice Demonstration Project, report, No. 0702, December.
- U.S. General Accounting Office**, 2002, "School vouchers: Characteristics of privately funded voucher programs," report, Washington, DC, No. GAO-02-752, September 10.

West, Martin R., and Paul E. Peterson, 2006, “The efficacy of choice threats within school accountability systems: Results from legislatively induced experiments,” *Economic Journal*, Vol. 116, No. 510, March, pp. C46–C62.

Witte, John F., 1997, “Achievement effects of the Milwaukee voucher program,” University of Wisconsin–Madison, Department of Political Science and Robert M. La Follette Institute of Public Affairs, mimeo, January.

Witte, John F., Troy D. Sterr, and Christopher A. Thorn, 1995, “Fifth year report: Milwaukee Parental Choice Program,” University of Wisconsin–Madison, Department of Political Science and Robert M. La Follette Institute of Public Affairs, mimeo, December.

Wolf, Patrick, Babette Gutmann, Michael Puma, Brian Kisida, Lou Rizzo, and Nada Eissa, 2008, *Evaluation of the DC Opportunity Scholarship Program: Impacts after Two Years*, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, report, No. NCEE 2008-4023, June.

Wolf, Patrick, Babette Gutmann, Michael Puma, Lou Rizzo, Nada Eissa, and Marsha Silverberg, 2007, *Evaluation of the DC Opportunity Scholarship Program: Impacts after One Year*, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, report, No. NCEE 2007-4009, June.