The Foundation for Educational Choice

WWW.EDCHOICE.ORG

NATIONAL RESEARCH

SECOND EDITION

A Win-Win Solution

The Empirical Evidence on School Vouchers

March 2011

Greg Forster, Ph.D.

About the Foundation for Educational Choice

The Foundation for Educational Choice is a 501(c)(3) nonprofit and nonpartisan organization, solely dedicated to advancing Milton and Rose Friedman's vision of school choice for all children. First established as the Milton and Rose D. Friedman Foundation in 1996, the foundation continues to promote school choice as the most effective and equitable way to improve the quality of K-12 education in America. The foundation is dedicated to research, education, and outreach on the vital issues and implications related to choice in K-12 education.

Commitment to Methods & Transparency

The Foundation for Educational Choice is committed to research that adheres to high scientific standards, and matters of methodology and transparency are taken seriously at all levels of our organization. We are dedicated to providing high-quality information in a transparent and efficient manner.

All individuals have opinions, and many organizations (like our own) have specific missions or philosophical orientations. Scientific methods, if used correctly and followed closely in well-designed studies, should neutralize these opinions and orientations. Research rules and methods minimize bias. We believe rigorous procedural rules of science prevent a researcher's motives, and an organization's particular orientation, from pre-determining results.

If research adheres to proper scientific and methodological standards, its findings can be relied upon no matter who has conducted it. If rules and methods are neither specified nor followed, then the biases of the researcher or an organization may become relevant, because a lack of rigor opens the door for those biases to affect the results.

The author welcomes any and all questions related to methods and findings.

THE FOUNDATION FOR Educational Choice

Table of Contents

List of Tables

Executive Summary

This report collects the results of all available empirical studies using the best available scientific methods to measure how school vouchers affect academic outcomes for participants, and all available studies on how vouchers affect outcomes in public schools. Contrary to the widespread claim that vouchers do not benefit participants and hurt public schools, the empirical evidence consistently shows that vouchers improve outcomes for both participants and public schools. In addition to helping the participants by giving them more options, there are a variety of explanations for why vouchers might improve public schools as well. The most important is that competition from vouchers introduces healthy incentives for public schools to improve.

Key findings include:

- Ten empirical studies have used random assignment, the gold standard of social science, to examine how vouchers affect participants. Nine studies find that vouchers improve student outcomes, six that all students benefit and three that some benefit and some are not affected. One study finds no visible impact. None of these studies finds a negative impact.

- Nineteen empirical studies have examined how vouchers affect outcomes in public schools. Of these studies, 18 find that vouchers improved public schools and one finds no visible impact. No empirical studies find that vouchers harm public schools.

- Every empirical study ever conducted in Milwaukee, Florida, Ohio, Texas, Maine and Vermont finds that voucher programs in those places improved public schools.

- Only one study, conducted in Washington D.C., found no visible impact from vouchers. This is not surprising, since the D.C. voucher program is the only one designed to shield public schools from the impact of competition. Thus, the D.C. study does not detract from the research consensus in favor of a positive effect from voucher competition.

- The benefits provided by existing voucher programs are sometimes large, but are usually more modest in size. This is not surprising since the programs themselves are modest – curtailed by strict limits on the students they can serve, the resources they provide, and the freedom to innovate. Only a universal voucher program could deliver the kind of dramatic improvement our public schools so desperately need.

Effects of Vouchers on Participants
(Number of Studies)

Table 1

	Positive Effect		No Visible Effect	Negative Effect
	All Students	Some Students		
Milwaukee	2	0	0	0
Charlotte	2	0	0	0
D.C.	2	0	0	0
New York	0	2	1	0
Dayton	0	1	0	0

Note: All studies examining the impact of vouchers on participants using random-assignment methods are included. A total of eight studies are included; one study is represented more than once because it produced findings in three different cities.

Effects of Vouchers on Public Schools
(Number of Studies)

Table 2

	Positive Effect	No Visible Effect	Negative Effect
Milwaukee	6	0	0
Florida	11	0	0
Others	4	0	0
D.C.	0	1	0

Note: Because random-assignment methods cannot be used to study the impact of vouchers on public schools, all studies are included. A total of 19 studies are included; three studies are represented more than once because they each produced findings in two different locations.

Introduction

School choice programs, which allow parents to send their children to the school of their choice, public or private, are among the most prominent and successful reforms in the education field. There are now 26 school choice programs in 16 states and Washington D.C. More than 190,000 students use these programs to attend private schools using public funds.[1]

The most well-known form of school choice is school vouchers, which give parents the ability to redirect their children's education funding to support their schooling in private schools. An alternative approach, tax-credit scholarships, gives donors a tax credit if they donate money to support private school scholarships.

Perhaps the most important question about vouchers is how they impact academic outcomes, both for the students who use them and in the public schools at large. Defenders of the government school monopoly claim that vouchers do no good for the students who use them, and harm public schools by "draining money" or by "creaming" – that is, skimming off the best students who rise to the top and would be sought by private schools. Voucher proponents, on the other hand, argue that vouchers improve academic outcomes both for the participating students and for public schools at large. They say vouchers save money for public school budgets rather than "draining" money, and send all types of students to private schools rather than "creaming." They also point to the benefits of allowing each student to find the right school, and the healthy incentives created by competition.

A large body of empirical evidence examines these issues. The effects of these programs have been studied using scientific methods and are no longer the subject of mere speculation and anecdotal observation.

This report reviews all available empirical studies on participant effects that use the "gold standard" method of random assignment and all available empirical studies (using any scientific method) of how voucher programs affect academic achievement in public schools. It also discusses the most important methodological issues confronted by research on this subject and some of the larger implications of what the research finds.

This is an updated version of a January 2009 report, published under the title "A Win-Win Solution: The Empirical Evidence on How Vouchers Affect Public Schools." That report included only the research on how vouchers impact public schools; in this updated version we are including the research on participant effects as well. The research reviewed in this report is complete and up to date, including all available empirical studies of which the author is aware as of February 2011.

Choice and Competition in Education

Unfortunately, Americans are not accustomed to thinking of K-12 education in terms of choice.

They expect and demand the right to select their own goods and services in everything from food, housing, clothing, transportation and medical care to magazines, haircuts, dry cleaning and video games. If government attempted to assign people to live in certain neighborhoods or shop at certain grocery stores, they would howl in protest. Americans even expect and demand choice when it comes to education outside of K-12 schools – everywhere from colleges to trade schools to night classes. But when it comes to K-12 education, the idea that they would choose for themselves rather than having government dictate what they receive is new and sometimes uncomfortable.

This helps explain why many Americans readily accept claims about school vouchers that are empirically false or poorly reasoned. When teachers' unions claim that vouchers "drain money" from public schools, many Americans nod in agreement. But how would those same people respond if they were told that from now on they would have to receive all their medical care from a doctor assigned to them by the government, rather than from their current family doctor, on grounds that their choice to seek care from their current doctor "drains money" from the budget of the doctor chosen by the government?

In fact, voucher programs have historically improved public schools financially. From 1990 to 2006, the nation's school choice programs saved $422 million for local school districts and $22 million for state budgets.[2] When students leave public schools using vouchers, not all the funding associated with those students goes with them. This means public schools are left with more money to serve the students who remain. Educating students in private schools rather than public schools not only accomplishes better results, it also costs less.

Similarly, the claim that vouchers "cream" or attract the best students from public schools has no empirical evidence to support it. The best available analyses of this question have found voucher applicants to be very similar to the population of students eligible for vouchers in terms of demographics and educational background.[3] In the Washington, D.C. voucher program, applicants were very similar to a representative sample of the eligible population citywide not only in terms of demographics, but also in their baseline test scores.[4]

Meanwhile, for similar reasons, the idea that vouchers might improve public schools seems counterintuitive to many Americans. In fact, it is not hard to explain. One reason is because vouchers allow parents to find the right particular school for each child. Every child is unique and has unique educational needs.

But probably the most important reason vouchers would improve public schools is because they give parents a meaningful way to hold schools accountable for performance. Under the current system, if a school isn't doing a good job, the only ways to get a better school – purchase private schooling or move to a new neighborhood – are expensive and cumbersome.

These options are especially difficult for low-income and disadvantaged students.

Thus, in the absence of parental choice, schools lack the healthy, natural incentives for better performance that most other types of service institutions take for granted. Hospitals know they must do a good job or lose patients. Professionals like doctors and lawyers must provide good services or lose clients. Stores must provide good value or lose customers.

Vouchers apply to schools these healthy incentives that we take for granted everywhere else. If a public school is providing adequate services, parents can leave their children where they are and be no worse off. But if not, parents can choose a private school that will serve their children better. Either way, schools know that parents have the power to hold them accountable.

The same Americans who have difficulty with the idea that competition improves schools have no difficulty applying the same concept everywhere else. They know that monopolies provide poor quality because they have little incentive to serve their clients well. And when they get bad service, customers say, "I'll take my business elsewhere" because they know that this provides an incentive for better service.

They do this even in fields like medical care where the service providers have other motives besides profit-seeking for being in their professions.

"Vouchers apply to schools the healthy incentives that we take for granted everywhere else."

Let's assume for a moment that your local hospital is motivated purely by a desire to provide medical treatment to people in need and not at all by revenue. Still, if that hospital were losing patients because it provided poor care, that loss of patients would motivate the hospital to improve care – and the patients know it.

This lack of connection between what Americans think about choice and competition in K-12 education and what it thinks about choice and competition in virtually every other aspect of life is a great hindrance to accurate public discussion of school vouchers. One good hope for rectifying that problem is to make the public aware of the large body of empirical research that examines how vouchers impact student outcomes for participants and public schools.

Why Science Matters – the "Gold Standard" and Other Methods

There is no such thing as a "scientifically right" education policy. Science cannot tell us what education policy is most fitting to the intrinsic nature of the human person, or most in alignment with America's traditional ideals of freedom and democratic self-rule. To answer those questions, we need other kinds of knowledge – knowledge about the nature of human life, the meaning of

freedom, and the historic self-understanding of the American people.[5]

However, the public is not content to set education policy based solely on abstract ideas and history. It hears competing claims about the real-world effects of education policy in the concrete world of the here and now. It wants – and rightly so – to know which claims are true and which are false. That is an empirical question, and it is entirely appropriate for science to address it.

When evaluating the effectiveness of an education policy, it is especially important to rely on empirical research of high scientific quality. Student outcomes are affected by many different influences – including demographic factors (income, race, family structure, etc.), school factors (type of school, teacher quality, etc.) and intangibles such as the level of enthusiasm parents and teachers invest in a child's education. The job of empirical science is to disentangle the influence exercised by each of these factors as well as can be done with the available evidence.

A study that uses good methods can overcome these problems and provide reliable information about what is influencing student outcomes. But if scientific procedures are not rigorously followed, or if we don't bother looking at the science and try to make judgments without it, we can come to the wrong conclusions about what factors cause what outcomes.

The gold standard for empirical science is the method known as "random assignment," in which subjects are randomly divided into a treatment group that will receive the treatment being studied (such as vouchers) and a control group that will not receive it. Because the two groups are separated only by a random lottery, they are likely to be very similar in every respect other than the treatment.

There is a substantial body of random-assignment research on the academic achievement of students who are offered vouchers, which is reviewed in the next section of this report. See that section for more about the importance of this method and its results in studies of vouchers.

Where a significant body of random-assignment studies exists, we should give priority to those studies when considering the evidence. Their methodological superiority should be taken into account; it would make no sense to act as though they were no more reliable than any other kind of study.

But while it may be the best kind of research, the gold standard of random assignment is not the only kind of research worth considering. Where it is not possible to conduct a random-assignment study, other kinds of research methods can produce useful information that sheds light on important policy questions.

The next best research method is to track year-to-year changes in outcomes for individual students. Tracking individual students over time removes from the analysis most, though not all, of the influence of unmeasured factors. If a student is

Parents Say School Choice Works

In addition to direct measurements of academic achievement, there is a large body of research on how parents evaluate the schools that serve their families through school choice programs. The knowledge provided by test scores and other direct measurements of academic achievement is important but is not the whole story. Education includes training in knowledge but is also more broadly a part of the process of forming a complete person, which includes many critical variables that test scores don't capture. Also, using test scores and similar measures as our only definition of educational success implicitly puts the school system under the control of test designers rather than parents, who may seek a variety of educational goals for their children that tests don't capture or that the test designers may not fully share.

For these reasons, parental satisfaction has always been an important part of school choice research. Numerous studies have gathered data on parental satisfaction in school choice programs, consistently finding that parents are more satisfied with the education delivered through school choice.[6] There is no room to review all that research here. However, three of the random assignment studies reviewed in this report include information on parental satisfaction; since these studies are the gold standard of research methods, their findings on parental satisfaction are worth noting here.

Greene's study in Charlotte finds that parents in the voucher group were far more likely to rate their school an "A" than parents in the public school group (53% v. 26%). They were also far more likely to be "very satisfied" with their schools across a variety of variables; for example, 54% of voucher parents were "very satisfied" with their schools' teaching, compared with 27% in the public school group (Greene 2001, pp. 57-58). Howell and Peterson's studies in Dayton, New York and Washington measured parental satisfaction across 16 variables and found that voucher parents reported much higher satisfaction on every variable. Examples include "academic program" (54% v. 15%), "what is taught" (55.5% v. 15%), "teacher skills" (58% v. 18%), "school discipline" (53% v. 15%), "safety" (54% v. 18%), "moral values" (52% v. 17%), and "teacher-parent relations" (53% v. 22.5%; Howell and Peterson 2002, p. 173). The Wolf, et. al. study in Washington also found that voucher parents were more likely to be satisfied with their schools. Parents who used a voucher were 25 percentage points more likely to rate their schools an "A" or "B" (Wolf et al 2007, p. xxi).

advantaged in a way that is not measurable, that advantage will typically be present in the student's outcomes for both year one and year two of the study; thus the change in outcomes between year one and year two will mostly be due to other factors – though unmeasured factors will still exert some influence on the level of year-to-year change. Removing the influence of unmeasured factors allows the analysis to isolate the impact of the factors that are being measured, such as exposure to vouchers.

If it is not possible to track individual students, good research still can be done by tracking year-to-year changes in individual schools. It is reasonable to expect that the unmeasured advantages of the students in a given school will be similar from year to year. If a school had highly advantaged students last year, it probably will still have highly advantaged students this year. Mobility among the student population will create some change in student characteristics from year to year, but not so much that we cannot learn from school-level studies.

Although research on other types of outcomes is briefly noted, this report focuses on studies that examine test scores and, in one case, graduation rates. This is not because these outcomes are the only thing that counts in education – far from it. But these outcomes offer the best opportunity to carefully distinguish the effects of schools and policy choices from other factors, such as wealth and family background. Focusing on these outcomes is the best way to test the claims made on both sides of the debate.

Participant Effects: Academic Outcomes of Students Using Vouchers

There have been eight studies using random-assignment methods to examine how vouchers affect outcomes for participants. However, most researchers characterize this body of research as "ten studies" because one of the studies is a compilation of three separate assessments of data from voucher programs in three different cities. That gives us a total of ten separate examinations of programs.

These ten assessments consistently find that vouchers benefit students. Six of them find that vouchers had a positive impact across all students participating. Another three find that vouchers had a positive impact on some student groups and no visible impact on other students. One study – which has grave methodological flaws that undermine confidence in the results – found no visible impact from vouchers. None find that vouchers had a negative effect.

The Importance of Random Assignment – the Gold Standard

This report focuses on studies using random-assignment methods, which separates subjects into "treatment" and "control" groups randomly. Random assignment gives us high confidence that factors other than the one we're really studying – the treatment – are not influencing our results.

The special value of random-assignment research in education policy is that it removes not only the influence of observable factors such as demographics, but also the vast array of unobservable factors that we know influence education but cannot control for in any other way. For example, it is widely agreed that one of the key drivers of student outcomes is parental motivation; parents who highly value the education of their children are an important positive influence on outcomes. Random assignment gives us high confidence that differences in factors such as this are not influencing the results of our research.

Unfortunately, it is usually not possible to conduct random-assignment research on education policy. However, school vouchers have been one of the rare exceptions.

When there are more applicants for a voucher program than there are slots available, a random lottery is often used to determine who may participate. This creates a naturally occurring random-assignment research design. Students who win the lottery and are offered vouchers can be compared to students who lost the lottery and were not offered vouchers. If we see any systematic (i.e. non-random) differences between the outcomes of the two groups, those differences can be attributed to the offer of vouchers, because nothing other than the offer of vouchers and randomness separates the groups.

Because random assignment is so preferable, a significant body of random assignment studies should be given priority wherever such a body of research exists. It would make no sense to consider gold-standard studies side by side with studies that are more methodologically limited, as though we could have the same level of confidence in both.

What the Gold Standard Studies Show

The first random-assignment study of vouchers was conducted by Jay Greene, then of the University of Texas-Austin, and Paul Peterson and Jiangtao Du of Harvard University in 1998. In the Milwaukee program, students do not apply for vouchers directly, but rather apply for seats in specific schools. Each school that is oversubscribed conducts a random lottery among all applicants to determine who will be offered the seats.

Today, with the Milwaukee program serving more than 20,000 students in 111 schools, there is no way to conduct a high-quality random-assignment study across all the students and schools, unless the state decides to change the program's governance in order to make such a study possible. However, when the program was first created, it was limited by law to a much smaller number of students.

> *Because random assignment is so preferable, a significant body of random assignment studies should be given priority wherever such a body of research exists.*

In its first year (1990-91), only 337 students participated; in subsequent years that grew to 504, 591, 718 and 786 respectively.[7] The program was then expanded significantly in 1995-96, and even more significantly in 1998-99.

In the program's early years, the state collected data on those who applied for seats in private schools through the program and were either accepted or turned away in a random lottery. Greene, Peterson and Du were able to take advantage of data collected on applicants in the years 1990-93.

Comparing those who used vouchers with those who were not offered them, the researchers found that vouchers improved outcomes for participants. Voucher students who stayed in the program all four years had reading scores 6 points higher and math scores 11 points higher than the control group.[8]

Also in 1998, Cecelia Rouse of Princeton University conducted a random-assignment study using the data collected on the Milwaukee program in its initial years. She found that over four years, vouchers produced an improvement of 8 points in math; she found no visible difference in reading scores.[9]

The next opportunity for a random-assignment study of vouchers was a privately-funded voucher program in Charlotte, N.C. In this program, students applied directly to the funders for vouchers. Because there were more applicants than vouchers available, and the funders chose among the applicants by lottery, the program created another natural random-assignment experiment.

In 2001, Greene (then with the Manhattan Institute) conducted a random-assignment study of the program. He found that in one year, voucher students scored 6 points higher on combined math and reading tests.[10]

Joshua Cowen reanalyzed Greene's data in 2008. Cowen applied a different method of accounting for students who won a voucher lottery but chose not to use the voucher. He found that voucher students had reading scores 8 points higher than the control group, and math scores 7 points higher, after one year.[11]

In 2002, William Howell of the University of Wisconsin-Madison and Peterson conducted a study that examined three privately funded voucher programs in three cities – Dayton, New York City, and Washington, D.C.

Studying the privately funded Washington, D.C. voucher program, they found that voucher students scored 7.5 points higher in combined math and reading tests after two years. However, the difference was no longer visible in the third year of the data.[12]

This curious result calls for an explanation. Howell and Peterson point out that among students who won the lottery in that year, a shocking percentage (71%) decided not to use them.[13] Such high attrition from the participating population in the

third year tends to reinforce the legitimacy of the second-year positive finding, over dismissing it in light of the third-year finding.

This was just at the time charter schools were exploding as a presence in Washington, D.C. The vouchers in this program did not cover the full cost of private school education, requiring parents to chip in some of the cost. Charter schools, by contrast, are "free" for parents, because they are funded by the government school monopoly's tax revenue. This creates an unfair competition for the business of parents seeking school choice; they can pay to go to private schools, or attend charters for "free." We cannot know whether the benefits of the voucher program could have been sustained if they had been able to offer parents resources comparable to those of government-owned, government-funded charter schools.

Studying the program in Dayton, Howell and Peterson found no visible difference across all students after two years. However, among black students – making up 72% of the participants – they did find gains. Black students scored 6.5 points higher on combined math and reading tests after two years. Among non-black students as a group there was no visible difference.[14]

They found similar results in New York City. Among all students they found no visible difference after three years. Black students – who made up 42% of the participants – scored 9 points higher in combined test scores as a result of vouchers. Non-black students as a group showed no visible effects.[15]

It is not clear why, in these two cases, only black students gained enough benefit from vouchers for the impact to be statistically visible. The most plausible hypothesis is that these students were served more poorly in their public schools and thus stood to gain the most from the opportunity to use vouchers. Since the program was limited in size and scope, it could only deliver limited benefits; thus it is plausible that only the students most in need of help were helped sufficiently. (See below for more discussion on why existing voucher programs don't produce larger benefits.)

Howell and Peterson's data for New York were subsequently reanalyzed by two other teams of researchers. In 2003, John Barnard of deCODE Genetics, Constantine Frangakis of Johns Hopkins University, Jennifer Hill of Columbia University and Donald Rubin of Harvard University reanalyzed the math scores in the first year of the program. They separated students based on the quality of the public schools they left to use vouchers. Among students leaving low-quality public schools they found a one-year gain of 5 points.[16]

Alan Krueger and Pei Zhu of Princeton University conducted a reanalysis in which they changed the way students were classified by race. Where student self-identification is absent, the generally accepted method is to use the race of the father. Krueger and Zhu classified a student as black if either parent was black – and made this change only for black students. Arbitrarily redefining a critical variable in an abnormal way,

Effects of Vouchers on Participants
(All Studies Using Random-Assignment Methods)

Table 3

Location	Years Covered	Published	Author	Method	Results			
					Positive Effect		No Visible Effect	Negative Effect
					All Students	Some Students		
D.C.	6	2010	Wolf et. al.	Gold Standard	•			
Charlotte	1	2008	Cowen	Gold Standard	•			
New York	3	2004	Krueger & Zhu	Unreliable			•	
New York	1	2003	Barnard et. al.	Gold Standard		•		
New York	3	2002	Howell & Peterson	Gold Standard		•		
D.C.	3			Gold Standard	•			
Dayton	2			Gold Standard		•		
Charlotte	1	2001	Greene	Gold Standard	•			
Milwaukee	4	1998	Greene et. al.	Gold Standard	•			
Milwaukee	4	1998	Rouse	Gold Standard	•			

and arbitrarily doing so only with one selected group rather than for all students, abandons the scientific method.

In addition, Krueger and Zhu also added students with significant missing data to their data set. Further, they failed to control for the student's baseline scores, a standard step in scientific education analysis.

Unsurprisingly, through these manipulations they were able to drag the variable for statistical significance down below the standard threshold. This allows them to claim that the voucher program had no visible effect on black student scores.[17]

In a devastating rejoinder, Howell and Peterson published a series of 120 reanalyses of their data set, each using a different set of specifications. These analyses demonstrate that the positive finding for black students is robust across numerous different assumptions about racial identification.

Howell and Peterson show that the positive effect only disappears if the analysis incorporates Krueger and Zhu's exact combination of arbitrary

racial redefinition, students with missing data, and exclusion of baseline scores. Leave out any two of these three and the results are positive.[18] Clearly, the Krueger-Zhu statistical model is completely discredited.

Finally, in 2010 a team of researchers led by Patrick Wolf of the University of Arkansas released the results of its six-year study of the Washington, D.C. voucher program. This study examined the voucher program created by Congress in 2003, not to be confused with the privately funded voucher program studied by Howell and Peterson.

The study offered an important innovation by examining graduation rates. It is the only study ever to examine the impact of vouchers on graduation rates using a causal model. Other studies could compare graduation rates of public and private schools, but without random assignment there was no way to scientifically examine whether vouchers were causing higher graduation rates for participants.

The researchers found that vouchers improved graduation rates by 12 percentage points. The voucher students had an 82% graduation rate, compared to a graduation rate of 70% for students in the control group.[19]

On test scores, the study found that voucher students scored higher on average than the control group. However, under the statistical method selected for the study six years earlier, the statistical analysis fell a little short of reaching the conventional level of certainty. The result for reading was 94% certain, and prevailing conventions require 95% certainty to report a finding. The study's headline conclusion is therefore that vouchers had no visible impact on test scores.[20] However, the researchers conducted "sensitivity tests" and determined that the results might become statistically certain – producing a positive finding – under alternative statistical models.[21]

Off the Gold Standard

A great deal of empirical research has compared public and private school impacts on test scores using methods other than random assignment. This research question actually goes all the way back to the origins of modern education science. The Coleman Report, produced by James Coleman for Congress in 1966, attempted to identify input and quality factors in public schools that impacted student performance. But Coleman was unable (using the cruder statistical methods of the time) to identify any. In his subsequent research, however, Coleman found one school quality factor he could tie to better outcomes: private schools performed better than public schools.[22]

Since then, education researchers using better data and methods have demonstrated that a number of school variables, most notably teacher quality, have a large impact on student outcomes. But the desire to compare public and private school performance has never been far below the surface of the education research field – if only there were a way to overcome the problem that

private school parents are systematically different because they choose private schools!

Identifying all the non-gold-standard research that has been done on this question over the years would be too cumbersome to do here, and there is no need given that we have a substantial body of random-assignment studies. But it is worth noting that, like Coleman's, most of the studies that rise to a reasonable level of scientific quality have found in favor of private schools.

Some studies have simply conducted descriptive comparisons, gathering information about outcomes without the availability of statistical methods that would establish a causal link between the public/private variable and the outcomes. The most recent of these, just released, is a study (the latest in a series) comparing graduation rates for students in Milwaukee public schools and students participating in the Milwaukee voucher program. It finds that Milwaukee voucher students had a graduation rate of 82% in 2008-09, compared to 70% in public schools.[23]

> *In his subsequent research, Coleman found one school quality factor he could tie to better [learning] outcomes: private schools performed better than public schools.*

Other studies have attempted to use methodology to overcome the central problem, which is that a direct comparison between public and private schools compares students with dissimilar family backgrounds – "choosers" and "non-choosers." There are various methods for doing this.

One method is instrumental variable analysis, which uses proxy measurements to estimate the probability that a student is in private school. A 1997 study by Derek Neal used an instrumental variable technique to calculate the probability that a student would be in a private school, and found that students more likely to be in private school were less likely to drop out.[24] The limitation of this method is that the instrumental variable is never more than an imperfect proxy.

Another approach is to compare the experiences of the same students in public and private schools. A 2003 study by Jay Greene and Greg Forster found that disabled students using Florida's special-needs voucher program got better services and had better outcomes in private schools than those same students had received in public schools.[25] This method is much better for purposes of measuring the impact of vouchers on those students, but limits the ability to generalize the finding to others.

Still another approach is to use "matched" samples, in which students with similar observable characteristics (such as demographics and initial test scores) are matched to one another, then tracked and compared over time. A team of researchers led by John Witte is conducting an

ongoing study of matched samples in Milwaukee, in which voucher students were matched to similar public school students. In the most recent report, few visible differences were found.[26] The limitation of this method is that it only matches students by observed characteristics, not the unobserved ones (such as parental motivation) that are the core of the methodological problem.

Competitive Effects: How Vouchers Change Outcomes in Public Schools

It is not possible to conduct random-assignment research on how vouchers impact public schools. Random assignment is only possible in studies of participants because of the naturally occurring opportunity to conduct a random lottery when there are more voucher applicants than there are vouchers. There is no naturally occurring equivalent that would permit random-assignment research methods in studying the effects of vouchers on public schools.

We must therefore turn to other evidence. Fortunately, this question has only been studied more recently and the amount of evidence is manageable. It is also of good methodological quality – increasingly so over time.

The absence of random assignment is not as great a problem in studies examining only public schools. Here, we have no act of parental choice that needs to be overcome methodologically. We are not comparing "choosers" and "non-choosers." All the relevant students are non-choosers. We are only comparing schools exposed to vouchers and schools not exposed – which is usually an easier methodological barrier to overcome.

Nineteen empirical studies have been conducted on how voucher programs (and one tax-credit scholarship program) impact academic achievement in public schools. Of these studies, 18 find that vouchers improve public schools. The one remaining study found that vouchers had no visible impact on public schools. No empirical study has ever found that vouchers had a negative impact on student outcomes in public schools.

Significantly, the one study to find no visible impact was also the only study conducted on a voucher program that intentionally protects public schools from the impact of competition. Thus, it does not detract from the research consensus that choice and competition provided by vouchers improve public schools.

Milwaukee Vouchers

Six empirical studies have been conducted on how the Milwaukee voucher program affects academic outcomes at public schools. All

> *No empirical study has ever found that vouchers had a negative impact on student outcomes in public schools.*

six unanimously find that vouchers improve Milwaukee public schools.

Vouchers are available to all Milwaukee students who meet certain criteria, most notably an income restriction. Thus, in Milwaukee there is not a simple division between public schools that are and are not exposed to vouchers, as in some other programs. However, some Milwaukee public schools are much more exposed to vouchers than others, based on the demographic makeup of their student populations. Thus researchers have focused on isolating the academic impact of a school being more exposed to vouchers versus being less exposed.

This means the research in Milwaukee will tend to make the effect of vouchers look smaller than it really is. The studies are not comparing "Milwaukee with vouchers" to "Milwaukee without vouchers." They are instead comparing "Milwaukee with lots of vouchers" to "Milwaukee with fewer vouchers." This is like testing the effectiveness of a medicine by comparing the effects of a large dose to the effects of a small dose rather than to the effects of not taking it at all. But it is the best we can do given the absence of a better control.

The first empirical study on the Milwaukee program was conducted by Caroline Hoxby, then of Harvard University, and released in 2001. She compared schools where at least 66% of the student population was eligible for vouchers to schools where fewer students were eligible for vouchers. She found that in a single year, schools

> *"All six studies [conducted on the Milwaukee voucher program] unanimously find that vouchers improve the city's public schools."*

in the "more exposed to vouchers" group made gains that were greater than those of other Milwaukee public schools by 3 percentile points in math, 3 points in language, 5 points in science and 3 points in social studies.[27]

The next study, released in 2002, was conducted by Jay Greene and Greg Forster, then of the Manhattan Institute. Rather than dividing Milwaukee public schools into two groups, they used regression analysis to determine how changes in the percentage of students in a Milwaukee public school who were eligible for vouchers would impact a school's academic results. They found that greater exposure to vouchers had a positive effect on year-to-year changes in public school outcomes; the size of the effect was such that a school with all its students eligible for vouchers could be expected to outperform a school with only half its students eligible by 15 percentile points over four years.[28]

In two studies that were released in 2006, Rajashri Chakrabarti of the Federal Reserve Bank found that the Milwaukee voucher program improved public schools. Chakrabarti conducted multiple analyses using different methods for measuring

public schools' exposure to vouchers. Some are similar to Hoxby's method (though Chakrabarti divided schools into three groups rather than two) and others to Greene and Forster's method. In both studies, Chakrabarti found that increased exposure to vouchers improves academic gains in Milwaukee public schools. A revised version of one of these studies was released in 2008.[29]

A 2007 study was conducted by a team of researchers led by Martin Carnoy of Stanford University. This study used a modified form of the Hoxby/Chakrabarti method. The authors reported that their analysis "confirms the earlier results showing a large improvement in Milwaukee in the two years following the 1998 expansion of the voucher plan to religious schools." Before 1998, religious schools were excluded from the Milwaukee program, so many fewer students participated. When religious schools were admitted to the program in 1998, participation increased dramatically. The study also found that the improvements in public schools caused by vouchers did not get larger in subsequent years and were not dependent on the proximity of private schools to public schools.[30]

Finally, in 2009 Greene (now at the University of Arkansas) and Ryan Marsh (also at Arkansas) conducted the only study on the effects of vouchers on Milwaukee public schools to use individual student data. This improves the scientific quality of the analysis. Greene and Marsh created a "voucher options" variable to measure the different levels of availability of voucher options for different students. For students not eligible for free and reduced lunch, the variable is zero because the student is not eligible for vouchers. For lunch-eligible students – a proxy for voucher availability, as in previous studies – the voucher options variable measures the number of private schools that participate in the voucher program and serve the student's grade level.

Greene and Marsh found that the Milwaukee voucher program improved performance for public school students, with an effect that is modest in size. Each individual private school existing in the city that accepted vouchers and served a student's grade level increased that student's performance by 0.055 points in language arts, 0.047 points in math, and 0.058 points in reading. They ran numerous additional analyses using different controls, sample separations, and other tests for robustness; these caused the size of the results to vary, but the positive effect was consistently present. Like Carnoy et. al., Greene and Marsh also tested the effect of proximity to private schools; while some of their analytical models showed a larger effect from having voucher options closer to the school, overall this did not appear to be an advantage.

Florida Vouchers and Tax-Credit Scholarships

Eleven empirical studies have been conducted on how two voucher programs and one tax-credit scholarship program in Florida have affected academic outcomes at public schools. All eleven

unanimously find that vouchers have improved Florida public schools.

Nine of these Florida studies examine the effects of the state's A+ program, which gave vouchers to students at chronically failing public schools before the program was ended by court order in 2006. Under the A+ program, each public school received an annual grade from the state based primarily on how many of its students either achieved an adequate score on the state test or made substantial progress toward an adequate score. If a school received two (or more) F grades from the state in any four-year period, students who had attended that school in the year of its second (or subsequent) F grade could apply for vouchers. Students were required to apply for vouchers during the two-week period immediately following the public announcement of the second (or subsequent) F grade; after this brief window closed, vouchers were no longer available. However, those students who did manage to apply during the brief eligibility window could continue using vouchers in subsequent years.

The first study of the A+ program was published in 2001 by Greene. At that point, only two schools had ever been eligible for vouchers under the program – too few to provide a basis for meaningful analysis. Instead, Greene studied the impact of the mere *threat* of vouchers on schools that were in danger of becoming eligible for vouchers if they did not improve.

> *All eleven studies conducted on Florida's programs unanimously find that vouchers have improved public schools.*

Using a simple descriptive analysis, Greene found that schools that had received an F grade, which would be eligible for vouchers if they received another F grade, made much larger year-to-year gains than schools that received a D (18 points in reading and 26 points in math for F schools versus 10 points in reading and 16 points in math for D schools). Greene then drew two further comparisons intended to isolate the impact of the voucher threat: high-scoring F schools compared with low-scoring D schools, and high-scoring F schools compared with low-scoring F schools. There was a substantial difference between high-scoring F schools and low-scoring D schools (16 points in reading and 24 in math versus 13 points in reading and 18 in math). However, a regression analysis showed that among F schools there was no statistical relationship between their test scores in the prior year and their test scores in the subsequent year – high-scoring F schools and low-scoring F schools had about the same results in the following year. Greene concluded that the difference in outcomes was attributable to receiving an F grade from the state, which included the voucher threat.[31]

This analysis was methodologically simple, as is often the case the first time an empirical question is being studied. Greene's analysis in this first study did not examine some alternative possibilities that might account for a relationship

between receiving an F grade and making bigger improvements the next year. His next study, and later studies conducted by others, included additional analyses designed to test whether the improvements associated with the F grade were due to these alternative explanations or to vouchers, or both. (See below for discussion of these alternatives and the results of the analyses examining them.)

In a subsequent 2007 study, Greene, along with Marcus Winters of the Manhattan Institute, used a more advanced statistical method. Greene and Winters divided schools into four categories:

- *Sometimes D* schools were those that had received a D grade, but no F grades and at least one grade above a D, in any of the previous four years;

- *Always D* schools were those that had received D grades in each of the previous four years;

- *Voucher Threatened* schools were those that had received exactly one F grade in the previous four years;

- *Voucher Eligible* schools were those that had received two or more F grades in the previous four years.

They then used regression analysis to compare the year-to-year gains made in schools in each of these four categories with those of other Florida schools.

For both math and reading scores, on both the state test and the national norm-referenced Stanford-9 test, Greene and Winters found that the positive impact of the A+ program closely tracked the schools' distance from vouchers. *Voucher Eligible* schools made the biggest academic gains, followed by smaller gains in *Voucher Threatened* schools, followed by the two categories of schools that had received Ds but no Fs. For example, in math scores on the state test, *Voucher Eligible* schools made improvements 15 points higher than other Florida public schools, while *Voucher Threatened* schools made improvements 9 points higher, *Always D* schools 4 points higher, and *Sometimes D* schools 2 points higher.[32]

When Greene and Winters' analysis was published in the journal *Education Next* in the summer of 2004, it was accompanied by an analysis conducted by Rajashri Chakrabarti, then of Cornell University. Chakrabarti used a simple descriptive analysis to provide further assurance that the relationship between the F grade and school improvements was due to vouchers. She compared the improvements made by F schools under the A+ program with improvements made by schools in the lowest performance category (out of four) under the state's previous school evaluation system. The previous system had no voucher component. Chakrabarti found that under the previous system, putting a school in the lowest-performing category did not improve its performance relative to schools in the next lowest performance category, while F schools did make bigger gains than those of D schools

under the A+ program. Over three years, the gap between F schools and D schools closed from almost 15 points to about 5 points.[33]

In 2006, Chakrabarti released a more sophisticated analysis that compared the impact of the A+ program on public schools to that of the Milwaukee voucher program. (This is one of the two Milwaukee studies by Chakrabarti cited in the section on the Milwaukee program, above.) A revised version of the study was released in 2008.

In this study, she used regression analysis to compare the trends over time in the academic achievement of schools that received F, D, and C grades in 1999, the first year of the A+ program. She compared trends in outcomes at these schools before and after the implementation of the program in 1999. Chakrabarti found that when F and D schools are compared to each other, the F schools made gains 8 points larger in reading and 5 points larger in math over three years. When F and D schools are separately compared to C schools, the F schools made gains 17 points larger in reading and 11 points larger in math than the C schools over three years, while D schools made gains 9 points larger in reading and 4 points larger in math than the C schools.
Chakrabarti then confirmed the impact of the F grade using a method known as "regression discontinuity," which isolates the impact of the F grade from other factors by comparing high-scoring F schools with low-scoring D schools. Regression discontinuity is a very high-quality method that is widely considered the next-best thing to a random-assignment study. However, it does limit the scope of the analysis, since it excludes many of the F schools from the data set. She found that the high-scoring F schools outscored the low-scoring D schools by 7 points in reading and 6 points in math over three years.[34]

Chakrabarti further confirmed this analysis with a follow-up study in 2007. This study took advantage of the fact that school grades are based primarily on how many students are either above or approaching a given cutoff score on the state test. Chakrabarti found that in schools that had received an F grade, students near the cutoff made larger gains relative to the gains of students at other schools, while other students at F schools were not negatively affected. The study used a regression discontinuity design to compare high-scoring F schools and low-scoring D schools.[35]

David Figlio of the University of Florida and Cecelia Rouse of Princeton University have also studied the A+ program to examine its impact on public schools. Their initial analysis, released in 2004, was the first to use student-level data rather than school-level data, providing improved scientific quality. In this analysis they examined data up through 2000-01 (that is, before vouchers became widely available in 2002-03, as was the case in Greene's initial 2001 study). They found that if a school received an F grade, its students made gains on the state test that were 2 points larger in reading and 5 points larger in math than those of other Florida schools over one year. Scores on the national

Stanford-9 test also improved. They confirmed the existence of a positive effect from the F grade using a regression discontinuity model, examining scores on the Stanford-9. A revised version was published in 2006.[36]

In a subsequent study released in 2007, in which they were joined by Jane Hannaway of the Urban Institute and Dan Goldhaber of the University of Washington, they collected data through 2004-05, using these data to track the continuing effects on schools that had received Fs in 2002-03. The study used school data and a regression discontinuity model to compare high-scoring F schools and low-scoring D schools. It found that receiving an F grade in 2002-03 produced academic improvements in students' test scores in the next year relative to those in non-F schools, and that these improvements were sustained in future years. They presented their results in terms of standard deviations rather than test score points; they found that the gains were equal to about a tenth of a standard deviation.[37]

Martin West and Paul Peterson of Harvard University released an analysis using individual student data in 2005. It found that among schools that had not received the lowest possible rating under the state's previous school evaluation system (and thus were "shocked" by the imposition of the F grade), receiving an F under the new accountability system produced an improvement in student's test scores equal to about four percent of a standard deviation.[38]

In 2008, Greg Forster, then of the Friedman Foundation, conducted a study examining the impact of the A+ program in every year from 2001 through 2006. Previous studies had only examined the impact of getting a particular grade, such as F or D, in a single year (usually occurring either in 1999, the first year grades were given out under the A+ program; or in 2002, the first year when a substantial number of schools were eligible for vouchers). Because vouchers were not widely available until 2002, and the voucher element of the A+ program was struck down by court order in early 2006, this study was able to track the changing impact of the A+ program as the status of the vouchers in the A+ program changed.

Forster used Greene and Winters' four categories to examine the impact of the voucher threat. He found that in 2001, before vouchers were widely available, *Voucher Threatened* schools made gains relative to all Florida schools equal to 13 points on Florida's new "developmental scale," which uses a single scale to track student scores from 3rd grade through high school. The next year, when vouchers were widely available, *Voucher Threatened* schools gained 15 developmental points, but *Voucher Eligible* schools gained 67 developmental points relative to other Florida schools. Over the next three years, as the percentage of eligible families using vouchers decreased due to the artificial obstacles created by the state department of education, the positive voucher effect was not as large but remained substantial (*Voucher Eligible* schools gained from 20 to 27 developmental

points each year). Then, in 2006, the first year after the voucher element of the program was removed, *Voucher Eligible* schools gained only 11 developmental points. Results for *Voucher Threatened* and D schools followed similar patterns.[39]

The only empirical study conducted on Florida's McKay voucher program was also released in 2008. The McKay program allows any disabled student in Florida public schools to use a voucher. Conducted by Greene and Winters, now of the University of Arkansas, the study used student-level data over five years to measure the relationship between the academic performance of disabled students in public schools and the number of private schools accepting McKay vouchers nearby. They used the number of nearby private schools to measure exposure to vouchers because the McKay program is statewide; as in Milwaukee, it is not possible to simply distinguish schools that are exposed to vouchers from schools that are not. As in Milwaukee, this will tend to make the effect of vouchers look smaller.

While studies in Milwaukee found no relationship between proximity to private schools and voucher improvements, the Greene and Winters study found a strong relationship between the presence of private schools participating in the McKay program and voucher improvements. This may be because Milwaukee is a small, dense city with lots of public transportation, while most of Florida has low population density and less public transportation.

The strongest effect of the McKay program was among students classified as learning disabled, representing 61 percent of all Florida disabled students. At a public school with an average number of private-school McKay competition within five miles, the positive impact of the McKay program was equal to 16 points in math and 24 points in reading among learning disabled students.[40]

Finally, Cassandra Hart and David Figlio published a study on Florida's tax-credit scholarship program in 2011. They used individual student data to compare public school outcomes in the year before the program was enacted (1999-2000) with outcomes in subsequent years through 2006-07.

Tax-credit scholarships are an alternative way of providing private school choice; donors are provided with a tax credit for donating money to support K-12 scholarships. Other tax-credit programs exist in Arizona, Georgia, Iowa, Indiana, Pennsylvania and Rhode Island. Florida's program serves only low-income students, enrolling nearly 29,000 in 2009-10.

The tax-credit scholarship program is statewide, so as in Milwaukee and the McKay program (see above) it is not possible to simply compare schools that are exposed to school choice with schools that are not. Hart and Figlio used four types of measurements to differentiate public schools that were more or less exposed to school choice: distance to the nearest private school, the number of private schools within a five-mile

radius, the number of different types of private schools (ten types were differentiated) within a five-mile radius, and an index that combined the second and third measurements. Again, as in the previous cases, the inability to simply compare "vouchers" and "no vouchers" cases will tend to make the effect of vouchers look smaller.

On all four measurements, Hart and Figlio found that increased exposure to private school competition caused public school performance to improve. Moving the nearest private school 1.1 miles closer to a public school would increase its math and reading scores by 0.015 standard deviations. Adding 12 private schools within five miles would increase its performance by 0.03 standard deviations. Adding two additional types of private schools within five miles would increase it by 0.02 standard deviations. And a one-standard-deviation increase in the index of diversity and density of private schools would increase it by 0.01 standard deviations.

On all four measurements, Hart and Figlio found that increased exposure to private school competition caused public school performance to improve.

Other Programs

Four studies have been conducted on the impact of voucher programs in other places. Three of these studies find that vouchers improve public schools; one finds that vouchers make no visible difference to public school outcomes.

The first of these studies was conducted by Christopher Hammons of Houston Baptist University in 2002. Hammons examined century-old voucher programs in Maine and Vermont. When these states first created public schools, they gave small towns the option of "tuitioning" their students – using public funds to pay for their students to attend private schools or nearby public schools – rather than building their own public schools.

Hammons measured the relationship between a public school's academic achievement and its distance from the nearest "tuitioning" town. Using regression analysis, he found a positive relationship. The relationship was strong enough that if a town one mile away from a school began tuitioning its students, the percentage of students at the school passing the state's achievement test could be expected to go up by 3 percentage points.[41]

In the same 2002 study in which they examined the impact of the Milwaukee program, Greene and Forster also examined the impact of a large-scale privately funded voucher program targeted to the Edgewood school district in San Antonio, Texas. Unfortunately, it was not possible to

Effects of Vouchers on Public Schools
(All Empirical Studies)

Table 4

Location	Years Covered	Published	Author	Method	Results		
					Positive Effect	No Visible Effect	Negative Effect
Florida	7	2010	Hart & Figlio	Student data	•		
Milwaukee	7	2009	Greene & Marsh	Student data	•		
Florida	5	2008	Grenee & Winters	Student data	•		
Ohio	2	2008	Forster	School data	•		
Florida	5	2008	Forster	School data	•		
Milwaukee	10	2008	Chakrabarti	School data	•		
Florida	8			School data (regression discontinuity)	•		
Florida	8	2007	Chakrabarti	School data (regression discontinuity)	•		
Milwaukee	15	2007	Chakrabarti	School data	•		
Florida	6	2007	Rouse et. al.	School data (regression discontinuity)	•		
Milwaukee	8	2007	Carnoy et. al.	School data	•		
D.C.	1	2006	Greene & Winters	School data		•	
Florida	6	2006	Figlio & Rouse	Student data (regression discontinuity)	•		
Florida	3	2005	West & Peterson	Student data	•		
Florida	2	2004	Greene & Winters	School data	•		
Florida	4	2004	Chakrabarti	School data (descriptive only)	•		
Milwaukee	4	2002	Greene & Forster	School data	•		
San Antonio	3			District data	•		
Maine	1	2002	Hammons	School data	•		
Vermont	1			School data	•		
Milwaukee	3	2001	Hoxby	School data	•		
Florida	2	2001	Greene	School data (descriptive only)	•		

differentiate between Edgewood schools that were more or less exposed to competition from the voucher because the program offered vouchers to every student in the Edgewood district. Greene and Forster instead examined the performance of the district as a whole. District-wide data are less methodologically desirable than school data. But where no other data can be examined, district data at least provide a rough indication of how Edgewood performed in the presence of vouchers.

Controlling for demographics and local resources, they found that Edgewood's year-to-year test score gain outperformed those of 85% of school districts in Texas. Given that Edgewood is a high-poverty (93% eligible for lunch programs) and high-minority (97% Hispanic) district, the study concludes that such a high statewide academic rank for Edgewood suggests that vouchers produced public school improvements.[42]

In 2006, Greene and Winters released a study of how the federal voucher program in Washington, D.C. impacts public schools. Because eligibility for the voucher program is restricted to a relatively small number of students, particularly in the program's first year (when the study was conducted), Greene and Winters measured exposure to the voucher program by measuring the distance between each public school and the nearest private school participating in the voucher program. They found no visible relationship.[43]

The D.C. voucher program is the nation's only voucher program with a "hold harmless" provision that allocates additional money to the public school system to "compensate" for the loss of students. This is intended to insulate the public school system from the impact of competition from vouchers. Thus, the absence of a visible effect in this study does not detract from the research consensus in favor of a positive impact from vouchers on public schools.

Finally, in a 2008 study, Forster examined the impact of Ohio's EdChoice voucher program on public schools. The EdChoice program offers vouchers to all students attending chronically failing public schools. In the program's first year – the year covered by Forster's study – schools were eligible if they had been designated in a state of "academic emergency" by the state in each of the last three years. The definition was subsequently expanded to include more schools. Forster used regression analysis to examine year-to-year test score changes in schools where students were eligible for vouchers. Forster found positive effects from the EdChoice program in math scores for 4th and 6th grade students and reading scores for 6th grade students, and no visible effect in other grades. The positive effects ranged from 3 to 5 points in one year.[44]

Alternative Theories

As the first studies on this subject emerged, some speculated that the improvements in public schools might be caused by other factors besides positive incentives from vouchers. Subsequent

research rigorously tested the alternative hypotheses that were offered and found them to be unsupported.

These theories were more extensively discussed in the 2009 edition of this report. A brief summary of the issues appears here.[45]

One theory speculates that the worst students may be more likely to use vouchers, leaving behind the better students in public schools. On this theory, rather than "creaming" the best students, vouchers "dredge" the worst. As has already been noted, the direct evidence on this question (such as we have) supports neither the creaming nor the dredging hypothesis.[46] Also, nine studies show Florida's A+ program has improved public schools that are merely threatened with vouchers, without actually removing students from schools. It is also worth noting that a number of studies have tracked the achievement of individual students rather than whole schools, and still found that vouchers improve public school outcomes.

Another theory speculates that since vouchers improve schools' per-student finances (see above), the positive effects of vouchers could result from this fiscal benefit rather than competition. Of course, this would not be an argument against vouchers even if it were true. However, as we have seen, the evidence in Florida shows a positive voucher effect even where no students have actually changed schools, simply as a result of the threat of vouchers.

Two other alternative theories emerged after Greene's initial study of Florida's A+ voucher program. One is the hypothesis of a "stigma effect" – schools assigned an F grade improve in order to remove the stigma of being labeled as failing, rather than responding to voucher competition. First, stigma cannot explain the positive findings for Milwaukee, Florida's two other programs, or town tuitioning vouchers in Maine and Vermont. Second, later studies (Greene and Winters 2004; Chakrabarti 2004, 2007 and 2008; Figlio and Rouse 2006; Forster 2008) have used various methods to check for the possibility of a stigma effect in Florida. All found that voucher competition had a positive impact independent of any stigma effect.

The other alternative theory is the "regression to the mean" or "mean reversion" hypothesis; failing schools are more likely to improve than to get worse simply because they can't get much worse than they already are. First, as with the stigma hypothesis, regression to the mean cannot explain the positive findings for Milwaukee, Florida's other programs, and town tuitioning vouchers in Maine and Vermont. Second, again, subsequent studies (Greene and Winters 2004; West and Peterson 2005; Figlio and Rouse 2006; Chakrabarti 2007 and 2008; Forster 2008) have examined this theory using various methods. All found no effect from regression to the mean. Perhaps most important, the studies using regression discontinuity (Figlio and Rouse 2006; Rouse, Hannaway, Goldhaber and Figlio 2007; and Chakrabarti 2007 and 2008) all confirmed the positive effect from the voucher

program. A regression discontinuity design excludes regression to the mean because the schools in the treatment group (high-scoring F schools) and the control group (low-scoring D schools) begin with very similar test scores.

Universal Vouchers Would Deliver Much Bigger Results

If vouchers are so great, why are the public school systems in cities and states with vouchers still showing little to no overall improvement? Milwaukee public schools were widely dysfunctional in 1990 when the voucher program was enacted, and they remain widely dysfunctional today. There has been no Milwaukee Miracle.

But the absence of a dramatic miracle is not a valid reason to conclude that vouchers aren't helping. The U.S. public school system is so tenaciously resistant to change that expecting miraculous results from any education reform is unreasonable.

The empirical evidence consistently shows that vouchers have succeeded in improving public schools. The size of the effect is often moderate, but a moderate positive effect is still a positive effect. Claims that vouchers "don't work" directly contradict a clear consensus in the scientific evidence.

> *... the absence of a dramatic miracle is not a valid reason to conclude that vouchers aren't helping.*

And yet, while it might be unreasonable to expect miracles, there is still an urgent need for larger improvements than vouchers are now delivering. Are the results of today's voucher programs the best that vouchers can do? Or can we reasonably expect to do better?

Yes, Vouchers Work

The overall performance of a school system is affected by countless factors. Some of these factors, such as political policymaking, can change quickly and dramatically. Others, such as demographic factors, are highly stable.

As a result, the overall performance of a school system can never by itself provide a reliable guide to whether any one factor (such as vouchers) is having a positive effect. If a man with asthma starts taking a new medication, and at the same time takes up smoking, his overall health and ability to breathe may not improve but this has no bearing on the question of whether the medicine is helping.

The only way to know whether vouchers are having a positive impact is to conduct empirical research using high-quality scientific methods. The whole purpose of using scientific methods is to isolate the impact of vouchers from all other factors that influence academic outcomes, so that we can measure it accurately.

Given the remarkably unanimous research on the impact of vouchers everywhere they are allowed to affect public schools, it is clear they are having a positive effect. It is wrong to say vouchers must be doing no good simply because a lot of public schools are still failing.

Vouchers Could Work Much Better… If We Let Them

The positive impact of voucher programs identified in the empirical research is sometimes large, but it is more often modest in size. That is hardly surprising, given that existing voucher programs are also modest in size. If modest programs produce modest benefits, not dramatic benefits, is the logical conclusion to deny that voucher programs have any benefits and give up on them? Or to expand them until they are large enough to have a dramatic impact?

Existing voucher programs are hindered in a number of ways, such as:

- limits on the number of students they may serve;
- limits on the types of students they may serve;
- limits on the purchasing power they are allowed to provide;
- limits on families' ability to supplement that purchasing power;
- limits on how students may be admitted to participating schools, and so forth.

The Foundation for Educational Choice's 2008 report *Grading School Choice*, by Robert Enlow, covers these limitations in more detail.[47]

Some of the most restrictive limits are imposed in the Milwaukee voucher program. That program was the early pioneer that founded the modern school choice movement. Because it started at a time when vouchers did not have a national movement behind them, the Milwaukee program had to accept more political compromises than recent school choice programs have had to accept. For the same reason – because it was the original pioneer – the Milwaukee program is taken to be the flagship voucher program and is the nation's most prominent school choice program. So, ironically, the program that labors under the worst restraints, which therefore ought to be expected to produce especially modest results, is actually expected to produce a Milwaukee Miracle and vindicate all voucher programs everywhere. This is unreasonable.

> *"The size of the [voucher] effect is often moderate, but a moderate positive effect is still a positive effect. Claims that vouchers 'don't work' directly contradict a clear consensus in the scientific evidence."*

Likewise, the privately funded voucher programs studied in some of the random-assignment research were restricted to small numbers and low-income students. As in Milwaukee, we are examining the effects of very limited programs. Even the more recent programs, which have been able to offer more choice, are horribly restricted. There is no program in the U.S. even close to offering school choice for all.

Only Universal School Choice Can Sustain Dramatic Change

Ultimately, the only way to make school reform work on a large scale is to break the government monopoly on schooling. The monopoly ensures that no meaningful accountability for performance can occur, except in rare cases as a result of herculean efforts. The monopoly maintains power because a dense cluster of rapacious special interests fights back against all efforts to improve schools. Worst of all, by making it impossible for an education market to emerge, the monopoly removes all the necessary preconditions for sustainable innovation. When it comes to improving schools, the monopoly isn't just one powerful obstacle among many; it's what makes all the many obstacles as powerful as they are.

No social system can hope to advance unless it has three things. First, there must be opportunities for productive innovation outside the dominant service providers – new providers must be able to set up shop and do things differently. But in education, except for a few niche markets, new entrants can't emerge. They must charge tuition; the government crowds them out by raising money through taxes and then providing the service for "free."

Second, everyone must understand that the standard of success is whether the people who actually use the service find that it meets their needs. The only reason schools exist is to serve parents by helping them educate their children, so parents' decisions to stay in a school or seek schooling elsewhere is the only standard that truly measures whether schools are doing what is actually their job. Yet the monopoly ensures that standard can never be applied.

Third, service providers who do their jobs well must succeed, and be seen to succeed, while those who do not must not. This applies at the individual level (those who do their job well should be promoted) and the organizational level (those that serve people well should prosper). The monopoly ensures no such thing can take place.

Existing school choice programs don't provide enough students, dollars, and freedom to sustain new schools and allow a robust education market to emerge. Only universal vouchers can break the education monopoly and produce the dramatic improvements we need.

Conclusion

Even if vouchers did not improve test scores for participants and in public schools, there

would still be other reasons to implement them. Vouchers put students into schools that graduate more students, earn significantly greater satisfaction from parents, provide better services for disabled students, improve racial integration and students' civic values, save the public money, and so forth.[48]

There are also other reasons one might support vouchers independent of their impact on test scores. Perhaps the most important argument is that they return control of education to parents, where it had rested for much of our nation's history. The seizure of power over education by a government monopoly and attendant interest groups (especially unions) has had far-reaching implications for our nation. The American founders would have viewed it as incompatible with a free and democratic society, as well as a realistic understanding of the natural formation of the human person in the family.

However, when all these issues have been considered, the empirical question of how vouchers impact student test scores remains – and it remains important. Vouchers do, in fact, improve test scores for both participants and public schools. The benefits of competition in education are clearly established by the evidence. The only remaining question is whether the evidence will be permitted to shape public debate on the question of vouchers.

Notes

[1] See http://www.edchoice.org/School-Choice/School-Choice-Programs.aspx.
[2] Susan Aud, "Education by the Numbers: The Fiscal Effect of School Choice Programs, 1990-2006," Friedman Foundation for Educational Choice, April 2007.
[3] William Howell and Paul Peterson, *The Education Gap: Vouchers and Urban Schools*, second edition, Brookings Institution, 2006, p. 61-65.
[4] Patrick Wolf, et. al., "Evaluation of the D.C. Opportunity Scholarship Program: Final Report," U.S. Department of Education, June 22, 2010.
[5] Two illuminating books on the historic American understanding of education as it relates to the principles of human personhood, character, freedom and democracy are James Davison Hunter, *The Death of Character*, Basic Books, 2000; and Charles Glenn, *The Myth of the Common School*, Institute for Contemporary Studies, 2002.
[6] For example, see Greene and Forster, "Vouchers for Special Education Students"; and Greg Forster and Christian D'Andrea, "An Empirical Evaluation of the Florida Tax-Credit Scholarship Program," August 2009.
[7] See http://www.schoolchoiceinfo.org/facts/index.cfm?fl_id=1.
[8] Jay Greene, Paul Peterson and Jiangtao Du, "School Choice in Milwaukee: A Randomized Experiment," in *Learning from School Choice*, eds. Paul Peterson and Bryan Hassel, Brookings Institution, 1998; see also Jay Greene, Paul Peterson and Jiangtao Du, "Effectiveness of School Choice: The Milwaukee Experiment," *Education and Urban Society*, February 1999.
[9] Cecilia Rouse, "Private School Vouchers and Student Achievement," *Quarterly Journal of Economics*, May 1998.
[10] Jay Greene, "Vouchers in Charlotte," *Education Next*, Summer 2001.
[11] Joshua Cowen, "School Choice as a Latent Variable: Estimating the 'Complier Average Causal Effect' of Vouchers in Charlotte," *Policy Studies Journal*, May 2008.
[12] Howell and Peterson, *The Education Gap*, p. 146.
[13] Howell and Peterson, *The Education Gap*, p. 149.
[14] Howell and Peterson, *The Education Gap*, p. 146.
[15] Howell and Peterson, *The Education Gap*, p. 146.
[16] John Barnard, Constantine Frangakis, Jennifer Hill and Donald Rubin, "Principal Stratification Approach to Broken Randomized Experiments: A Case Study of School Choice Vouchers in New York City," *Journal of the American Statistical Association*, June 2003.
[17] Alan Krueger and Pei Zhu, "Another Look at the New York City School Voucher Experiment," *American Behavioral Scientist*, January 2004.
[18] Paul Peterson and William Howell, "Voucher Research Controversy," *Education Next*, Spring 2004.
[19] Wolf, et. al., "Evaluation of the D.C. Opportunity Scholarship Program," p. xv.
[20] Wolf, et. al., "Evaluation of the D.C. Opportunity Scholarship Program," p.35.
[21] Wolf, et. al., "Evaluation of the D.C. Opportunity Scholarship Program," p. 35 and Appendix C.
[22] Paul Peterson, "A Courageous Look at the American High School," *Education Next*, Spring 2010.
[23] John Warren, "Graduation Rates for Choice and Public School Students in Milwaukee, 2003-2009," School Choice Wisconsin, January 2011.
[24] Derek Neal, "The Effects of Catholic Secondary Schooling on Educational Achievement," *Journal of Labor Economics*, January 1997.
[25] Jay Greene and Greg Forster, "Vouchers for Special Education Students: An Evaluation of Florida's McKay Scholarship Program," Manhattan Institute, June 2003.
[26] John Witte, et. al., "The MPCP Longitudinal Educational Growth Study: Third Year Report," School Choice Demonstration Project, April 2010.
[27] Caroline Hoxby, "Rising Tide," *Education Next*, Winter 2001.
[28] Jay Greene and Greg Forster, "Rising to the Challenge: The Effect of School Choice on Public Schools in Milwaukee and San Antonio," Manhattan Institute, October 2002.
[29] Rajashri Chakrabarti, "Impact of Voucher Design on Public School Performance: Evidence from Florida and Milwaukee Voucher Programs," Federal Reserve Bank of New York Staff Report #315, January 2008 (see also 2006 manuscript); Rajashri Chakrabarti, "Can Increasing Private School Participation and Monetary Loss in a Voucher Program Affect Public School Performance? Evidence from Milwaukee," *Journal of Public Economics*, June 2008.
[30] Martin Carnoy, et. al., *Vouchers and Public School Performance: A Case Study of the Milwaukee Parental Choice Program*, Economic Policy Institute, 2007. Curiously, the authors seem anxious to deny the positive results of their own study. Although their data clearly show that the Milwaukee voucher program improved public schools, the authors argue that we should not conclude the voucher program improved public schools. Their arguments amount to little more than sophistry. For specific discussion see p. 17 of the 2009 edition of this report (http://www.edchoice.org/Research/Reports/A-Win-Win-Solution----The-Empirical-Evdence-on-How-Vouchers-affect-Public-Schools.aspx).
[31] Jay Greene, "An Evaluation of the Florida A-Plus Accountability and School Choice Program," Manhattan Institute, February 2001.
[32] Jay Greene and Marcus Winters, "Competition Passes the Test," *Education Next*, Summer 2004.
[33] Rajashri Chakrabarti, "Closing the Gap," *Education Next*, Summer 2004.
[34] Chakrabarti, "Impact of Voucher Design."
[35] Rajashri Chakrabarti, "Vouchers, Public School Response, and the Role of Incentives: Evidence from Florida," Federal Reserve Bank of New York Staff Report #306, October 2007.
[36] David Figlio and Cecilia Rouse, "Do Accountability and Voucher Threats Improve Low-Performing Schools?" *Journal of Public Economics*, January 2006 (see also 2004 manuscript).
[37] Cecilia Rouse, Jane Hannaway, Dan Goldhaber and David Figlio, "Feeling the Florida Heat?" How Low-Performing Schools Respond to Voucher and Accountability Pressure," National Center for Analysis of Longitudinal Data in Education Research, November 2007.
[38] Martin West and Paul Peterson, "The Efficacy of Choice Threats within School Accountability Systems: Results from Legislatively Induced Experiments," *Economic Journal*, March 2006 (see also paper presented to the Royal Economic Society, March 23, 2005).
[39] Greg Forster, "Lost Opportunity: An Empirical Analysis of How Vouchers Affected Florida Public Schools," Friedman Foundation for Educational Choice, March 2008.
[40] Jay Greene and Marcus Winters, "The Effect of Special-Education Vouchers on Public School Achievement: Evidence from Florida's McKay Scholarship Program," Manhattan Institute, April 2008.
[41] Christopher Hammons, "The Effects of Town Tuitioning in Vermont and Maine," Friedman Foundation for Educational Choice, January 2002.
[42] Greene and Forster, "Rising to the Challenge."
[43] Jay Greene and Marcus Winters, "An Evaluation of the Effects of D.C.'s Voucher Program on Public School Achievement and Racial Integration after One Year," Manhattan Institute, January 2006.
[44] Greg Forster, "Promising Start: An Empirical Analysis of How EdChoice Vouchers Affect Ohio Public Schools," Friedman Foundation for Educational Choice, August 2008.
[45] Readers seeking more detail should consult the 2009 edition, which can be accessed at http://www.edchoice.org/Research/Reports/A-Win-Win-Solution----The-Empirical-Evdence-on-How-Vouchers-affect-Public-Schools.aspx.
[46] Howell and Peterson, *The Education Gap*, p. 61-65; and Wolf, et. al., "Evaluation of the D.C. Opportunity Scholarship Program."
[47] Robert Enlow, "Grading School Choice," Foundation for Educational Choice, February 2008.
[48] See Greg Forster, "Monopoly versus Markets: The Empirical Evidence on Private Schools and School Choice," Foundation for Educational Choice, October 2007.

About the Author

Greg Forster, Ph.D., is a senior fellow with the Foundation for Educational Choice. He conducts research and writes on school choice policy. Forster has conducted empirical studies on the impact of school choice programs in Milwaukee, Ohio, Florida and Texas, as well as national empirical studies comparing public and private schools in terms of working conditions for teachers; racial segregation; and teacher and staff misconduct. He also has conducted empirical studies of other education topics, including charter schools, accountability testing, graduation rates, student demographics, and special education.

Forster's research has appeared in the peer-reviewed publications *Teachers College Record and Education Working Paper Archive*, and his articles on education policy have appeared in the *Washington Post*, the *Los Angeles Times*, the *Philadelphia Inquirer*, *Education Next*, the *Chronicle of Higher Education*, and numerous other publications. He is co-author of the book *Education Myths: What Special-Interest Groups Want You to Believe about Our Schools and Why It Isn't So*, from Rowman & Littlefield. Forster is also a contributor to Jay P. Greene's Blog (jpgreene.com).

We Welcome Your Support

Our goal is to promote Milton and Rose Friedman's vision of a society where all parents have the freedom to choose the school that works best for their children, regardless of whether that school is publicly or privately run. One way we achieve this goal is by producing studies and reports on school choice that educate the public and policymakers on the need for and benefit of educational freedom. As a nonprofit organization, our work relies solely on the generous support of our many friends and donors.

Please send your tax-deductible gift today to help advance liberty and choice in our educational system. With your help, America can achieve the Friedmans' vision of universal school choice.

To order copies of this report,
please call (317) 681-0745,
or visit us online at www.edchoice.org

THE FOUNDATION FOR Educational Choice

ONE AMERICAN SQUARE, SUITE 2420
INDIANAPOLIS, IN 46282
317-681-0745 • WWW.EDCHOICE.ORG

The Foundation for Educational Choice is a 501(c)(3) nonprofit and nonpartisan organization, solely dedicated to advancing Milton and Rose Friedman's vision of school choice for all children. First established as the Milton and Rose D. Friedman Foundation in 1996, the foundation continues to promote school choice as the most effective and equitable way to improve the quality of K-12 education in America. The foundation is dedicated to research, education, and outreach on the vital issues and implications related to choice in K-12 education.

www.edchoice.org