NEPC Review: Promising Start: An Empirical Analysis of How EdChoice Vouchers Affect Ohio Public Schools

Christopher Lubienski
Indiana University, clubiens@iu.edu

Follow this and additional works at: https://scholar.colorado.edu/nepc
Part of the Education Commons

Recommended Citation

This NEPC Review is brought to you for free and open access by Centers and Research Institutes at CU Scholar. It has been accepted for inclusion in National Education Policy Center by an authorized administrator of CU Scholar. For more information, please contact cuscholaradmin@colorado.edu.
Summary of Review

A Friedman Foundation report attempts to find empirical support for the contention that competition from private schools, through voucher programs, improves the effectiveness of public schools. In the first year of Ohio’s new EdChoice voucher program, the report claims to have found substantial academic gains at public schools exposed to the possibility of losing students to vouchers. Despite being presented as scientifically rigorous, the report suffers from serious methodological shortcomings. The analysis uses weak variables and an incorrect approach to measuring academic gains and tries to make claims based on cherry-picking uneven results. Moreover, even accepting the study’s analysis, it produces a finding very much at odds with the author’s intent: that vouchers are not likely to close the achievement gap between high- and low-performing schools.

http://epicpolicy.org/thinktank/review-promising-start
I.  INTRODUCTION

Much of the research on programs that provide public funding to assist students in moving from public schools to private schools has focused on the impact on the student using the voucher, and arguments have raged about whether these programs actually boost achievement for those students transferring to private schools. In comparison, the question of the effects of vouchers on public schools has less often been the subject of rigorous empirical analysis. These effects, however, should be a primary concern because most American students remain in public schools, even in areas with established voucher programs. If voucher policies are to have wider impact beyond the small set of students who use them (and for whom the benefits are not clear), then their effects must extend to public schools.

Moreover, the positive effects that policy makers would seek must overcome any negative effects of spiraling declines in public school enrollment and funding potentially caused by vouchers. The overall effect should be the creation of demonstrably better outcomes for the students remaining in public schools. This is the claim made by the most recent report of the pro-voucher Friedman Foundation for Educational Choice, a report that serves as a timely reminder that these programs are also advanced with the stated purpose of generating the competitive incentives necessary to spur public schools to become more effective. The report is timely because it comes in the wake of a controversy ignited when other prominent voucher advocates publicly repudiated the movement on exactly the grounds that the report tries to reclaim here: the apparent failure of vouchers to generate improvement in public schools. In that respect, the report is particularly interesting because it attempts to find empirical support for a key plank of the voucher platform just as other advocates are citing a lack of such evidence.

In theory, vouchers for private schools not only provide benefits for the more active, discerning, or knowledgeable consumers that use them, but also—as “a rising tide lifts all boats”—create more competitive conditions that lead to higher quality educational options for all students when a program is in place. As Milton Friedman noted, competitive market incentives often lead to quality improvements for an elite few, but these improvements tend to filter down to the “basic product” as well if they are adopted by mass producers, thereby benefiting the many. Faced with the loss, or the threat of the loss, of students and the funding attached to those students, public schools may embrace more effective practices in order to attract and keep students who could otherwise use a voucher to transfer to a private school.

While the immediate effect on voucher students has been much studied, there has also been some research done on this secondary effect on public schools. Much of this research has been conducted by pro-voucher organizations such as the Friedman Foundation. (That research is discussed in greater detail later in this review.)
Given the generally positive findings of pro-voucher researchers, it was notable when, over the past year, major champions of vouchers highlighted the failure of vouchers to spark improvement in public schools. Writing in a publication of the pro-voucher Manhattan Institute, Sol Stern decided that school choice is not improving schools, lamenting that “the evidence is pretty meager that competition from vouchers is making public schools better.”5 Earlier, the pro-voucher Wisconsin Policy Research Institute released a report finding little academic improvement in Milwaukee’s public schools since that city’s voucher program began almost two decades ago.6 One of the prime movers of the voucher program in Milwaukee, Howard Fuller, indicated, “we may have oversold that point… I think that any honest assessment would have to say that there hasn’t been the deep, wholesale improvement in MPS that we would have thought.”7 The Friedman Foundation’s report appears at a time when there is a growing hole in the dike.

II. THE REPORT’S FINDINGS AND CONCLUSIONS

This new report from the Friedman Foundation8 examines Ohio’s new EdChoice program, a voucher plan that provides $4,375 ($5,150 for high schoolers) for up to 14,000 students in “chronically under-performing” public schools to attend a private school participating in the program.9 The statewide program began in 2005, three years after the US Supreme Court ruled 5-4 that Ohio’s voucher program in Cleveland was constitutional.10

Greg Forster, a fellow at the Friedman Foundation and author of the report, examined student growth in “voucher-eligible public schools”11 from 2005-2006 (as students were applying for the program) to 2006-2007 (the first year the program was in operation) for grades 3-8, finding “substantial academic improvements” in those schools threatened by vouchers, which he ascribes to the competitive effects generated by the voucher program (p. 5). The report, which was released jointly by the Friedman Foundation and nine other pro-voucher organizations, notes beneficial effects in some grades and “no negative effects” in the others (p. 5). It posits that, because the program is likely to grow and since gains may accumulate over time, EdChoice may go a long way in closing the achievement gap between the lowest and highest performing schools. Thus, the report concludes that it provides further evidence that vouchers do not harm public schools, but in fact “vouchers improve academic outcomes at public schools. Vouchers allow families to choose the right schools to meet their children’s needs and introduce competitive incentives for improvement that are lacking in the traditional government-run education system” (p. 5).

III. RATIONALES SUPPORTING FINDINGS AND CONCLUSIONS OF THE REPORT

The report uses data from the state of Ohio on achievement and basic demographics to compare academic gains in “voucher-eligible” schools to all public schools, controlling for several factors. The analysis wisely seeks to factor in the presence of charter (or, in Ohio, “community”) schools, which otherwise could confound efforts to determine if
any competitive effect was actually due to charter schools and not vouchers. The report also appropriately focuses on grades 3-8, since achievement data for secondary grade-levels were not available, and it excludes from the analysis schools with smaller grade cohorts so that individual students could not be identified. The report then uses linear regression to assess the impact of voucher eligibility on academic gains since the program began, comparing voucher-eligible schools to all Ohio public schools at grades 3-8. To consider the possibility of regression to the mean—schools with unusually low scores at Year 1 might be more likely than mid-scale schools to show growth—the report includes a second analysis, focusing only on voucher-eligible schools in state-determined “very poor urban districts” compared to other public schools in such districts.

While I review the data and methods below, it is important to note that the report advances from several assumptions, not all of which are supported or even considered in the analysis. The report assumes that any gains in achievement result from schools’ organizational effects, ignoring alternative explanations—other influences on student achievement are well-documented in the research literature, including peer effects and student and school socioeconomic status (SES). The report’s analysis implements basic demographic controls that are inadequate for addressing SES, as described below, and since it relies on school-level, rather than individual student-level, data it is impossible to determine, as the report claims to do, if its analytic approach in fact “removes most of the impact of confounding variables such as demographic factors and unobserved characteristics” (p. 12).

The report contends that vouchers allow parents “to hold schools accountable for teaching their students,” with the only other viable option for parents and students being “to move” to another district or school (p. 10). Leaving aside the fact that this claim ignores local democratic mechanisms for holding schools accountable, the assertion is premised on a number of questionable assumptions for voucher-based accountability to be workable: parents must have the information and knowledge about academic effectiveness in different schools; they must have the time and motivation as well as resources (such as transportation) and ability to absorb search costs in order to act on their preferences; competitive incentives must be clear to schools, which must then be able to respond in ways that have a discernable impact on school effectiveness. Yet the evidence for many of these premises is less than clear. Still, the report ignores all of these pre-conditions, falling into the problematic “black box” approach of examining only outputs and assuming that vouchers caused any improvements in school effectiveness.

The report’s conclusions are also based on the assumption that any voucher effects on academic achievement are cumulative (see below). However, this ignores academic slips such as summer reading loss that detract from a simple accumulation of effects and are more likely to affect poorer children, such as the ones in these “very poor urban districts.” The report also assumes that the claimed impact of the program will increase, rather than dissipate, as the
program is expanded—as Ohio policy-makers seem likely to do. A contrasting possibility would be that the most obvious and feasible responses to voucher competition—to the threat of losing students to EdChoice—were the first ones adopted by public schools, and later responses will have a diminishing rate of return.

Furthermore, the report appears to embrace a very simplistic model of market behavior. The report argues that “Where parents are empowered with school choice, schools that don’t adequately teach their students will lose them” (p. 10). Presumably, then, schools that do a better job of teaching will gain students. Yet the report provides no evidence that this is happening, and research does not necessarily bear out this assumption. For example, very few families have used the exit option in the federal NCLB act to transfer their children from public schools designated as “failing,” and at the same time the fastest-growing schools—conservative Christian schools—are among the lowest performing types of schools.14 This suggests that parental choices are influenced by a much broader set of factors than perceptions of “adequate” teaching.

Instead of being empirically based, the report’s assumptions appear to be more statements of belief based in a rudimentary and simplistic view of economic behavior in markets for education. In fact, rather than applying a generic model of markets to education, it would have been more useful to think about the many different types of consumer markets, and how education does or does not exhibit attributes associated with those various markets, such as asymmetries of information, ease of entry, and repeat purchases.

Finally, it has become fashionable for voucher advocates to claim that vouchers “do no harm” to public schools and do not “skim” better students from public schools.15 This largely unsupported contention (particularly considering the optimistic promises of a “panacea” made by earlier voucher advocates16) is repeated in this report. Since the results in the report are presented as averages for large sets of schools, we cannot tell from these data whether or not voucher dynamics harmed individual schools or students. In fact, there is growing evidence that choice programs such as charter schools—which are much more widespread than voucher programs—do in fact lead to increased sorting of students by race, SES, and academic ability.17 The study cited to refute this research—what the report calls the “best analysis” (p. 11)—was funded and conducted by pro-voucher organizations, and was not published in a peer-reviewed journal.18 It found that voucher students were demographically and academically similar to other students eligible for vouchers. In fact, by definition, there are qualitative and academic differences between students applying for vouchers to leave public schools and the students who do not apply, even if they are eligible. Families applying for vouchers have demonstrated an otherwise unobserved commitment to investing time in their children’s education, and they differ on observable factors as well.19

IV. THE REPORT’S USE OF RESEARCH LITERATURE

The report’s use of research literature is
somewhat misleading, highly selective and, in its use other advocacy literature, unabashedly incestuous. Specifically, the report overstates the empirical findings from earlier research. It uses a very selective and one-sided review of previous studies to support its assertions, ignoring the more respected (and peer-reviewed) studies in favor of reports from Friedman and other voucher advocacy organizations. And on the rare occasion when it does refer to reputable research, it mischaracterizes those findings in order to support its own agenda.

For instance, the report claims that “there is a large body of high-quality empirical evidence showing that vouchers make public schools better, not worse” (p. 10). It is a stretch to say that this literature is “large”—most observers would agree that much more attention has been focused on the immediate question of whether choice improves academic outcomes for the students exercising choice. And the assertion of “high-quality” is itself highly questionable; indeed, “research” is not defined. For example, the report contends as follows:

Numerous fiscal studies have examined whether vouchers and tax-credit scholarships (a similar type of school choice program) “drain money” from public schools. This body of research has shown consistently that these programs save money both for state budgets and for local public school districts, even after the fixed costs of public schools (costs that do not go away when students leave a school) are taken into account (pp. 10-11).

But it is not clear which studies are being cited. In a footnote, the report indicates that “Most of these studies are available in the research database hosted on the Friedman Foundation’s website (www.friedmanfoundation.org/friedman/research/ShowResearch.do)” (p. 17, Note 1).

However, this links to all of the research the Friedman Foundation posts on multiple topics. Presumably, the report is referring to the “fiscal impact” category, which lists about two dozen “studies,” many focused on single states. All of these were produced by the Friedman Foundation, and none were peer-reviewed.

Regarding the question of “saving money” (on which the report tries to make generalizations), the answer obviously depends on individual program design—programs could be structured to either protect or penalize public schools if they lose students. Indeed, it is puzzling as to why the report would assert that public schools are not harmed, since the report’s central argument depends on public schools feeling real penalties from the loss of, or threat of losing, students—as indicated with the report’s discussion of the DC voucher program. That is, the threat of lost students and funding is the key mechanism by which market competition might drive a response from public schools.

In any case, many of the studies appear to be unrelated or marginally related to the topic, or are not studies at all (for instance, one “study” is a statement to legislators). A handful of them focus on topics such as graduation rates or public-private school achievement, with
only tangential conjectures about the introduction of choice programs. Still others, rather than providing empirical analyses of actual data, simply make predictions about non-existent choice programs that the reports propose. (I have critiqued such fiscal impact “research” from the Friedman Foundation previously.20) According to the Friedman Foundation’s own criteria, these are not “high-quality” studies.21

To further support the claim that voucher competition improves public schools, the report cites five studies on the voucher program in Milwaukee, nine on Florida, and two on other states. The report celebrates the prestigious institutional affiliations of the researchers, although most of the studies were actually produced by pro-voucher think tanks, including the Friedman Foundation, Manhattan Institute, and the Hoover Institute.22 Of the 16 studies (one of which is cited twice23), five are unpublished working papers.24 Only two were conducted by respected researchers not associated with a pre-announced agenda regarding vouchers,25 and only one was published in a reputable peer-reviewed journal.26 Moreover, many of these studies conducted by voucher advocates have been questioned in the scholarly community. For instance, Professor Hoxby’s study of competition in Milwaukee27 was criticized for using a school-level analysis (as this report does) rather than a student-level analysis,28 and was published in a journal of “opinion and research” produced by pro-voucher think tanks.

The concern here is not the citation of work published outside rigorous peer-reviewed outlets. Even advocacy venues can produce high-quality research. Rather, the problem is that such a small fraction of the research has ever eventually made its way through such a rigorous peer-review process.

The Friedman report selectively focuses on studies—no matter what the quality—that appear to support its agenda. In doing so, it leaves out much high-quality research, much of it peer-reviewed, that seriously questions the assertion that the threat of losing students has a positive impact on public schools.29 Much of this work has been done on charter schools, which are much more widespread than the voucher programs that the report highlights. These studies explore dynamics comparable to those created by voucher policies, examining schools facing the threat (or real) loss of students. The overall results are mixed, but some of these studies identify essentially no, or even negative, competition effects on public schools.30 The report’s discussion of competition literature would be considerably stronger if it had included such research.

Finally, the report misrepresents findings in order to support its assertions. For instance, it cites a recent study by Carnoy et al. to lend support to its claim that “These studies unanimously found that public schools improve when a voucher system has been implemented” (p.11). This is a curious claim. Carnoy and his colleagues report on their two-part study, and neither part supports the Friedman conclusions. The first part was basically a replication of earlier studies, finding initial public school gains but also finding that those gains then fell off—suggesting that different types of market conditions need to be considered, as do alternative ex-
planations for gains. In the second part, they adopted a more nuanced approach, and found “essentially no evidence that students in those traditional public schools in Milwaukee facing more competition achieve higher test score gains.” The Friedman report misrepresents those major findings.

V. REVIEW OF THE REPORT’S METHODS

The report uses a regression analysis, controlling for basic demographic factors as well as for the presence of charters. Rather than tracking individual student gains over time, which is a preferable method for measuring school effects, the analysis uses school-level data. (This is unfortunate, because mobility between schools might mean the report is measuring substantially different populations within the same school from year to year—with these data we simply do not know.) Setting aside the use of school-level data, the regression approach is generally appropriate. But the analysis ultimately falters on some serious methodological missteps, fatally damaging the study.

The report is based on two comparisons: first, achievement growth for “voucher-eligible” schools compared to other Ohio public schools, and then, to check for the possibility of regression to the mean, it compares “voucher-eligible” schools in “very poor urban districts” to other public schools in such districts. This second comparison using a district-wide scope is problematic, since school quality and resources can vary substantially within districts. That is, the report’s approach implicitly assumes that the comparison district schools should also be very low scoring due to demographics—about the same as the voucher-eligible schools. But there is, by definition, something about these non-“voucher-eligible,” comparison schools that are also in poor districts that distinguishes them from “voucher-eligible” schools in those districts. This something might be leadership, teaching staff, curriculum, or the demographic backgrounds of students. Forster desperately wants to assume that vouchers are making up the difference, but it is likely that he is comparing the already more-effective schools to schools that had more room for improvement. Indeed, those differences between the two sets of schools would have pre-dated the introduction of the voucher program. A more useful comparison would be to examine student academic growth and instructional programs of individual schools where students are eligible for vouchers relative to demographically similar schools that are more academically successful in order to also consider other factors that could be responsible for greater school effectiveness.

Another problem is that the demographic measures used in this report are extremely weak. The analysis controls for the percentage of white students at a school, as well as the percentage of students eligible for free or reduced-price lunch. This white versus other approach prevents the analysis from distinguishing among African-American, Asian, or Hispanic students, and yet these are important considerations. (For example, Hispanic students might be more likely to apply for a voucher to attend a Catholic school, and their academic gains often differ from those of African-American students; and pro-voucher research, even in the best light,
has usually found an effect only for African-American students.\textsuperscript{34}) Furthermore, the free or reduced-price lunch classification is often regarded—especially by voucher proponents such as Forster\textsuperscript{35}—as an inappropriate variable in itself, which it can be when used as a sole measure of SES, as is the case in this report.

The report does make the reasonable decision to control for the presence of charter schools, presumably under the assumption that the availability of this additional option by which students can leave a school could confound the report’s attempt to measure the effect of the possibility of exit through vouchers. Yet the report only looks at the number of charter schools as a percentage of all schools in various cities. A more useful approach would have looked at market share: the percentage of students or spaces in charter schools. This is important because charter schools tend to be smaller schools, so school counts used in isolation could over-estimate charter competition. Moreover, it is unclear why the report would account for charter competition but fail to consider the number of private schools, especially those accepting vouchers, and the number of available spaces in those schools. There would be no threat of exit to a failing public school if there are no private schools nearby (or none that accept vouchers), or if the private and charter schools are full. Thus, between these factors, not to mention the availability of open-enrollment options (that is, the choice to enrollment to a public school other than one’s neighborhood school) in Ohio, it is nearly impossible to determine if EdChoice and the exit option it provides are the cause of any changes in achievement in public schools.

Finally, the report is concerned about disproving the possibility that “regression to the mean” is influencing any apparent gains in achievement—that a school’s low score in Year 1 would increase the next year as a result of a statistical artifact, rather than as a result of true gains. Although this is a valid concern, the possibility is present only because the analysis used differences in achievement between Year 1 and Year 2 as the dependent variable. Instead, the analysis should have used the Year 1 test as a predictor for Year 2 outcomes, as the author of this report has done elsewhere, but strangely not here.\textsuperscript{36} This preferred approach would control for the fact that some schools are starting Year 1 very high or very low, and thus have more “room” to make gains (or more limits, in the case of a ceiling effect for initially high-scoring schools). This more sophisticated approach would account for any tendency of regression toward the mean, much more so than the simplistic approach used in the Friedman Foundation report.

VI. REVIEW OF THE VALIDITY OF THE FINDINGS AND CONCLUSIONS

Perhaps the greatest weakness of the report is its interpretation of the results. The report concludes that its analysis demonstrates “substantial beneficial effects on academic outcomes in public schools from EdChoice vouchers, and no harmful effects” (p. 13). But these benefits are hardly clear. Consider the results of the two comparisons in Tables 1 and 2 of the report. Regressions were run on growth at five grade-to-grade stages (i.e., 3\textsuperscript{rd} to 4\textsuperscript{th}) separately in reading and mathematics, for two
different analyses, giving us a total of 20 coefficients for which there could be a voucher effect. Of those 20, 6 show a statistically significant positive effect (grades 4-5 and 6-7 in math, and grades 6-7 in reading in both analyses), with gains in math (+5 scale points) outpacing those in reading (+2). Yet there was no statistically significant effect for the vast majority of grades. Based on these highly uneven and inconsistent results, the report argues for a cumulative effect over time that can close the achievement gap by one standard deviation in four years.

Actually, ten of the twenty outcomes are either zero or negative, and those, along with four other positive effects, are not statistically significant at the 95% (or .05) level. Here it should be noted that the report’s author has insisted elsewhere (in arguing for evidence of academic outcomes) that statistical significance at the 95% level of certainty is not a cut-point but an arbitrary convention; he has argued instead for a continuum of confidence.37 If we accept this logic, the report has to some degree disproved its own claim that vouchers have “no harmful effects” on public schools (p. 13), since the results indicate a negative five-point effect in math from grades 5-6 in one analysis, with 89.2% certainty that it was not just statistical “noise” (p. 15).

The report asserts that vouchers have a “substantial” effect despite the fact that they do not have any measurable effect in the majority of situations analyzed. Further, Table 3 of the report cherry-picks the three findings where the analysis was able to get positive findings that were also statistically significant. This is inappropriate. In view of the inconsistency of the results, the patterns strongly suggest that there are other unmeasured factors at play here and that the three positive results are a fluke. In fact, the report should explain to readers why there was no effect (and possibly even a sizable negative one) in the vast majority of cases.

The report tries to build an even bigger case on even this paltry evidence, claiming: “If the effects accumulate over time, in three to four years the voucher-eligible schools will have improved by one standard deviation (equal to one-sixth of the distance between the top-scoring and bottom-scoring schools in Ohio)” (p. 5). Yet the report offers no support for this contention that any gains are cumulative. And even if this optimistic and unfounded assessment were correct, that means that a K-12 voucher program would not close the achievement gap between the lowest and highest scoring schools in Ohio.

VII. USEFULNESS OF THE REPORT FOR GUIDANCE OF POLICY AND PRACTICE

In addition to this and other reports, the Friedman Foundation produces a “Guide to Evaluating the Scientific Quality of Education Research.”38 The guide lists nine considerations in assessing research—issues such as alternative explanations, generalizability, appropriate comparisons, and the availability of other research. This new report arguably falls short on at least half of the Friedman Foundation’s own criteria.

While this analysis draws on quantitative data and asks interesting questions, the quality of the empirical analysis

http://epicpolicy.org/thinktank/review-promising-start
falls far short of the “highest standards of scientific rigor” promised by the Friedman Foundation. In view of the announced advocacy mission of the Friedman Foundation for Educational Choice regarding vouchers, and the notable flaws on this report, it is better read as a statement of belief than as an empirical analysis.
Notes & References


10 Zelman v. Simmons-Harris (U.S. Supreme Court 2002)

11 The term “voucher-eligible public schools” is somewhat misleading, since the schools are not eligible for vouchers, but different categories of students who were or would have been assigned to those schools can apply for a voucher


http://epicpolicy.org/thinktank/review-promising-start


15 Greene, J.P. (2000). *A survey of results from voucher experiments: Where we are and what we know* (Civic Report No. 11). New York: Center for Civic Innovation, Manhattan Institute

Greene, J.P. (2001). *A Survey of Results from Voucher Experiments: Where We Are and What We Know*. In C. R. Hepburn (Ed.), *Can the Market Save Our Schools?* (pp. 121-149). Vancouver, BC: The Fraser Institute


Chakrabarti, R. (2004). *Closing the Gap*. *Education Next (Summer)*


32 As Forster and others voucher advocates have noted here and elsewhere, this is not the ideal approach to determining school effects on learning. However, in view of the data limitations, that was the only route available for the analysis. See Forster, G. (2007). Monopoly vs. Markets: The Empirical Evidence on Private Schools and School Choice. Indianapolis, IN: Friedman Foundation


The Think Tank Review Project is made possible by funding from the Great Lakes Center for Education Research and Practice.