NEPC Review: The Effects of Means-Tested Private School Choice Programs on College Enrollment and Graduation (Urban Institute, July 2019)

T. Jameson Brewer

University of North Georgia

Follow this and additional works at: https://scholar.colorado.edu/nepc

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.
NEPC Review: The Effects of Means-Tested Private School Choice Programs on College Enrollment and Graduation (Urban Institute, July 2019)

Reviewed by:
T. Jameson Brewer
University of North Georgia
October 2019

National Education Policy Center
School of Education, University of Colorado Boulder
Boulder, CO 80309-0249
(802) 383-0058
nepc.colorado.edu
Acknowledgements

NEPC Staff

Kevin Welner
Project Director

William Mathis
Managing Director

Patricia Hinchey
Academic Editor

Alex Molnar
Publications Director


Funding: This review was made possible in part by funding from the Great Lakes Center for Educational Research and Practice.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

This publication is provided free of cost to NEPC’s readers, who may make non-commercial use of it as long as NEPC and its author(s) are credited as the source. For inquiries about commercial use, please contact NEPC at nepc@colorado.edu.
A new report by the Urban Institute, *The Effects of Means-Tested Private School Choice Programs on College Enrollment and Graduation* (Urban Institute, July 2019), compares certain outcomes of three school voucher programs to traditional public school programs. It finds that students using vouchers to attend private schools sometimes have higher rates of college enrollment and completion than their public school counterparts. These findings, however, arise from comparisons of apples to oranges, because the two case studies showing some voucher benefits do not sufficiently account for pivotal differences between choosers and non-choosers. Only in the third case study, which uses random assignment and thus avoids these selection effects, do we see no voucher benefits. Two other concerns are important to note. First, the literature review places an unbalanced reliance on non-peer-reviewed sources. Second, the report attempts to “move the goalposts” away from the test-score outcomes that have been the center of voucher advocacy and debate for decades—coinciding with recent voucher studies finding null or negative effects on test scores. These shortcomings render the report of limited value for evaluating voucher policies.
I. Introduction

As debates rage about the usefulness and appropriateness of market-oriented reforms like school vouchers, questions remain about how different approaches should be measured and compared. Against this backdrop, a new report by the Urban Institute, *The Effects of Means-Tested Private School Choice Programs on College Enrollment and Graduation*, authored by Matthew Chingos, Daniel Kuehn, Tomas Monarrez (each from the Urban Institute), Patrick Wolf (University of Arkansas), John Witte (University of Wisconsin-Madison), and Brian Kisida (University of Missouri), purports to measure the success of three voucher programs by considering their impact on college enrollment and graduation—two educational outcomes other than the test-score goal long set by advocates.²

As the report notes, “research on private school choice has often focused on measuring the impact of attending a private school on students’ test scores, relative to attending public school” (p. 1). Yet this test-score focus has increasingly shown mixed and often negative results. Perhaps not coincidentally, the report’s authors contend that the use of test scores to evaluate the effectiveness of voucher programs is inferior to evaluating outcomes such as high school graduation, college enrollment, and degree attainment, since educational attainment “may be the most consequential outcome for individual students and their surrounding communities over the long term” (p. 2).

The report concludes that in two of their three case studies, voucher programs produce at least some better educational attainment results than traditional public school systems. The third case study did not tease out any attainment benefits associated with being offered a
voucher. The report also notes that the positive results in two of the case studies still do not close the gaps between low-income students and their more affluent peers.

II. Findings and Conclusions of the Report

The report analyzes the results of recent studies of three voucher programs: (1) the Florida Tax Credit Scholarship; (2) the D.C. Opportunity Scholarship Program; and (3) the Milwaukee Parental Choice Program.

For the Florida Tax Credit Scholarship (FTC) analysis, the report highlights the positive impact that voucher use has had on college enrollment and degree attainment. The study matched 16,111 students who used a voucher each with “up to five nonparticipating students” (p. 4), concluding that voucher use in Florida produced higher rates of college enrollment within two years of a student’s expected high school graduation and higher rates of college degree attainment, with results being best for students who remained in private schools for a longer period of time. The report points out, however, that “these results should be interpreted with caution, as we expect students who persisted in the program might differ in unmeasured ways from those who left” (p. 6).

In reporting their findings related to the Milwaukee Parental Choice Program (MPCP), the report highlights the impact of using a voucher for private school enrollment on college enrollment and degree attainment analyzing two different cohorts: (1) ninth grade and (2) third through eighth grades, with baselines set in 2006 (with student controls being the following: race or ethnicity, gender, and baseline reading and math scores). The cohort of students who were in the voucher program in grades three through eight was larger than the ninth grade cohort, but the researchers did not know whether they remained in the program in their high school years. Further, the researchers had more control variables available for the ninth grade cohort.

Results on educational attainment varied depending on the cohort and on whether or not statistical controls were employed for student backgrounds and/or parental backgrounds. But the majority of analyses showed no statistically significant differences between the voucher and non-voucher students, whether or not controls were used. For the ninth graders, the report presents five sets of findings: “on the probability of ever enrolling in any type of college, the probability of ever enrolling in a two-year or four-year college, the probability of graduating from college, and the total amount of time spent in college” (p. 11). The results are offered for (a) no controls, (b) just student-level controls, and (c) parent-level controls added as well. The only statistically significant (at p < 0.05) benefit teased out regarding the three enrollment analyses was for enrollment in four-year colleges. (The report also includes [on p. 12] the following: “the statistical model with both student and parent controls was marginally statistically significant” of any college enrollment—pointing to the result for p < 0.1.)

The analyses of total time spent in college also showed little difference (again, looking at the ninth grade cohort). No analyses showed any difference for two-year colleges. For four-year
colleges, students who used a voucher in K-12 remained enrolled in college for longer, and this result held up when adding student-level controls: with the voucher students enrolled for 0.2 years more (1.4 years of enrollment for students who used a voucher in K-12 versus 1.2 years for students who did not use a voucher in K-12). However, when also controlling for parent differences, the analyses showed no statistically significant differences. (Parental controls were as follows: parent education, parent income, and parent religiosity.) Interestingly, the report presents these results in an unusual way, highlighting effect size and downplaying statistical significance. (“Once parent controls were added to the model, however, the MPCP advantage in time completed at a four-year college dropped to an extra 15 percent of a year and became statistically insignificant.”) Looking at college graduation rates among the ninth grade cohort, none of the analyses showed a statistically significant difference.

The other voucher cohort was also similar to the matched non-voucher students in quite a few of the analyses, although more analyses showed benefits with this cohort. The report focuses on analyses using student-level (but not parent-level, which were available only for the ninth grade cohort) controls. The following is how the report presents those results:

Students enrolled in the MPCP in grades three through eight in 2006 were 5 percentage points more likely than their matched MPS peers to enroll in any type of college by 2018 (50 percent versus 45 percent, an 11 percent difference). The higher college enrollment rate for the MPCP students was highly statistically significant (figure 6). The MPCP enrollment advantage for the younger sample of students is only clear regarding four-year colleges. MPCP students enrolled in two-year colleges at a rate that was 2 percentage points higher than their MPS peers (30 percent versus 28 percent), but that difference was not statistically significant. The younger sample of MPCP students enrolled in four-year colleges at a rate that was 4 percentage points higher than their MPS peers (30 percent versus 26 percent) and that difference was highly statistically significant. ** A similar pattern is apparent regarding the years of college completed by MPCP and MPS students in our 2006 third-through-eighth-grade sample (appendix tables A.4 and A.5). By 2018, both groups had averaged about slightly more than half a year at a two-year college. The MPCP students, however, averaged 15 percent of a year more time at a four-year college than the matched MPS students (0.94 of a year versus 0.79 of a year), another attainment advantage for the voucher students that was highly statistically significant.

Finally, in contrast to our ninth-grade sample, we see that students who were in grades three through eight in 2006 are graduating from college at higher rates if they were in the MPCP (figure 7). Only 3 percent of both the MPCP and MPS groups had graduated from a two-year college by 2018. The MPCP students graduated from four-year colleges at a rate that was 3 percentage points higher than their MPS peers (11 percent versus 8 percent, a difference of 38 percent), and that difference was statistically significant. (p. 14-16)

Overall, the report concludes that using a voucher through MPCP had no meaningful impact on college graduation rates for the older cohort of students studied, as compared to students who attended public K-12 schools in Milwaukee. But the study also found that the use of a
The analysis of the D.C. Opportunity Scholarship Program (OSP) differed from the other two case studies in three important ways: (1) the study used a random lottery, instead of matching students based on available control variables; (2) the study used “intent to treat” rather than actual receipt of the voucher; and (3) none of the results showed any voucher benefits. Analyzing the impact of the OSP on academic attainment, the report highlights college enrollment comparisons of students who did and who did not use a voucher to attend private schools through an intent-to-treat analysis. In addition to the discussion and table assessing college enrollment, the report’s appendix includes descriptive statistics, regression analyses of college enrollment within two years and five years of expected high school graduation, and an analysis of enrollment within two years, disaggregated by student demographics.

Within the body of the report, the findings presented suggest that “students offered a scholarship [voucher] were somewhat less likely to enroll in college within two years of expected graduation from high school”—but again with no statistical significance (p. 18). However, even given the lack of statistical significance, the report’s authors hypothesize that the possible negative impact is “perhaps because private schools are more likely to hold [students] back for multiple grades or because [students’] entrance into college is more likely to be delayed for other reasons” (p. 19). Overall, the report concludes that there is “no statistically detectable difference between students who won and lost the voucher lottery” (p. 21) and that despite the use of a randomized experiment, it is possible that the results are not as reliable as those in Florida and Milwaukee because students in Washington D.C. have “access to a significantly greater amount of public school choice (in the form of charter schools)” (p. 21).

The report concludes that students who used vouchers in Florida and Milwaukee were “more likely to enroll in and graduate from college than similar students who remained in public schools” (p. 21). The lack of a similar impact in D.C. was speculatively attributed to a small sample size or students attending charters rather than using vouchers to attend private schools.

The report notes that while voucher users showed higher rates of attainment than their public school counterparts in at least some analyses in two of their three case studies, low-income students still complete college at lower rates. Despite this persistence of a college completion gap between the affluent and the poor, the report’s authors suggest that vouchers can help “move the needle” (p. 21).

III. The Report’s Rationale for Its Findings and Conclusions

Using data from the National Student Clearinghouse (NSC), the report calculates the difference in academic outcomes for students who used (or, in one case, were offered) a voucher
to attend a private school and those who remained in local public schools. For the case study of the Florida Tax Credit (FTC) program, the report compares data from two existing, non-peer-reviewed, Urban Institute reports with data from the NSC, data available from the FTC, and from a non-profit, Step Up for Students, that administers vouchers for the FTC program. In the case of the D.C. Opportunity Scholarship Program (OSP), the report relies on combining data from an existing, non-peer-reviewed, Urban Institute report with additional data from the voucher administrator, Serving Our Children, to reconstruct baseline data so that attainment outcomes could be matched with data from the NSC. The report’s findings and conclusions on the Milwaukee Parental Choice Program voucher program (MPCP) are the result of a random sample of voucher students in grades three through eight at baseline and all of the MPCP ninth-grade students at baseline, compared with matched school-district students, combined with data available from the NSC.

The authors suggest that educational attainment is a more “consequential outcome” (compared to test scores) as students with higher levels of attainment live longer, are healthier, avoid being on welfare, and avoid criminality.

IV. The Report’s Use of Research Literature

The report cites 30 sources that are not peer reviewed and only 14 that are peer reviewed; however, some of the peer-reviewed studies are not the result of double-blind reviews (that is, reviewers knew the identity of the authors). One of the primary studies relied on was in fact reviewed after its publication, through the same process as this review: The National Education Policy Center’s NEPC Review project. Earlier this year, Professor Jaekyung Lee reviewed the underlying Florida FLC study, and it is worth quoting the entire Executive Summary from that review:

This Urban Institute research report is aimed at assessing the impact of the Florida Tax Credit (FTC) scholarship program on college enrollment and graduation. Through matched comparisons, the study finds consistently positive effects of the program for both two-year and four-year colleges, with relatively stronger effects for four-year private college enrollment and for students who stayed in the program longer. This review acknowledges the study’s contributions, using expanded data tracking to inform school voucher policy debates, but the review also raises three critical questions about the validity of some methods, findings and conclusions. First is the problem of selection bias. The study’s attempt to match FTC and non-FTC students resulted in the non-FTC comparison group having two to three times more students receiving reduced-price lunch. This is on top of the acknowledged problem that choosers could be more advantaged in unmeasured aspects than nonchoosers, and that this tends to result in an upward bias (positive effects on achievement). Second, without real achievement benefits, the estimated impact of FTC on college enrollment may reflect college matching effects rather than true program effects on students’ college readiness. Third, the FTC program’s reporting of much greater effects on college en-
rollment than on graduation suggests that the conditional FTC effect on college completion (i.e., conditional on college entry) could be null or even negative. These concerns call into question the use and misuse of this study by voucher advocates; proper use of the study begins with an understanding of its limitations and sees the need for confirmatory research and future exploration of potential mechanisms driving any increased college attainment.

The specifics of Professor Lee’s critique here are important, but so is the basic truth that important problems and limitations tend to be surfaced when we put our work through the peer-review process. The extensive reliance on non-peer-reviewed sources in this new report calls into question the strength of the underlying sources and the conclusions.

Also of note here is that the report arguably minimizes and misrepresents prior research on test-score effects. It starts off by noting that test-score research until recently “showed neutral to positive effects of private school choice on student achievement” (p. 1). It then acknowledges that recent studies that have found negative effects but asserts that “these negative effects tend to dissipate over time” (p. 1). Other researchers are less sanguine. As professors Lubienski and Malin recently wrote in *The Hill:*6

When vouchers were first studied, researchers fought vicious battles over relatively minor differences in academic achievement. Voucher advocates like DeVos embraced any evidence of learning gains for students using vouchers to switch to private schools, and a number of think tanks and large philanthropies like the Walton Family Foundation also lined up to support this education reform. Some even saw vouchers as the key for reducing achievement gaps between white and minority students. But while most researchers found that any gains were rather negligible overall, advocates argued that vouchers were at least not harming students’ academic achievement.

Recently though, there has been a sea-change in the results.

As city-based pilot programs in places like Cleveland and Milwaukee were eclipsed by statewide programs in Ohio, Indiana, Louisiana, and elsewhere, researchers are consistently seeing large, significant, negative impacts — outcomes almost unheard of in evaluations of education interventions.

Researchers — including several voucher advocates — have conducted nine rigorous, large-scale studies since 2015 on achievement in voucher programs. In no case did these studies find any statistically positive achievement gains for students using vouchers. But seven of the nine studies found that voucher students saw relative learning losses. Too often, these losses were substantial.

For the references used by those authors, please see their article. But some the most important recent research is set forth in here (in the endnote).7
V. Review of the Report’s Methods

Contrary to the implications of the title, the report is not a comprehensive summary of studies. In referring to the three studies, it does not explain the procedures, noting that full descriptions of the underlying methods (which vary for each of the three case-studies) are “described in detail” in other published works.

The analysis of the Florida Tax Credit Scholarship used “the same propensity score matching methodology described” in other studies\(^8\) to match students using a voucher to those not using a voucher. As reviewed by Lee,\(^9\) while demographics, test scores, family income, etc. were controlled for between the treatment and the control group, it remains impossible to control for the myriad of ways in which parental influence can bias outcomes both prior to use of a voucher or how using (or not using) a voucher may, or may not, alter parental influence. In fact, the “matched” non-voucher students appear to have been poorly matched.

Similar problems limit the usefulness of the analysis of the Milwaukee Parental Choice Program, which also matched a student using a voucher to a “similar student” remaining in a public school. The authors note that “students were matched based on their exact grade level and neighborhood, a narrow band for their initial test scores, and a propensity score that included their initial test scores (again) and student demographic variables, including race, gender, and English language learner status” (p. 11). While the authors employed various statistical controls for students and parents in their analysis of students who were in the ninth grade at baseline, the report notes that all of the analysis relative to the third through eighth grade cohort’s college attendance employed only student controls, because parental controls were not available. The methods used in the Washington D.C. OSP study are stronger, with the researchers able to rely on a random lottery and then backing it up with analyses that use control variables. “Intent to treat” is an appropriate approach for an intervention of this sort, but it cannot tease out any attainment differences between students who actually received the treatment and those who did not. Also, as is the case with the other two case studies, it’s important to note the limitation that the results cannot be generalized to students and families that never opted to apply for a voucher—those without the incentive, motivation or resources to seek out this option.

Because the report brings together three existing studies, the methods, data, and presentation of results are not consistent. For example, the authors write that “the appendix tables contain information on the effects of MPCP student enrollment on all 10 of our measures of attainment, including high school graduation” (p. 11)—yet, it isn’t readily clear why these findings were not highlighted in the same manner for Florida’s FTC or Washington D.C.’s OSP, nor is it clear why descriptive statistics were used for the OSP but not the other two case studies.

VI. Review of the Validity of the Findings and Conclusions

There has been a growing trend among voucher advocates to move the goalposts away from evaluating school choice reforms by comparing student test scores.\(^{10}\) This may be because
the trending results have begun to turn negative\textsuperscript{11}—thus undermining the justification for such reforms. Despite recent claims that pro-market advocates have not pushed for test score comparisons,\textsuperscript{12} it has in fact long been the claim that school choice models, such as vouchers, would improve student test scores, and advocates have long touted what were perceived to be an increase in test scores—albeit a position that has required the use of cherry-picking, questionable methods, and over-generalizations.\textsuperscript{13}

Significantly, the report notes that Florida FTC “schools are not required to accept the scholarship as full tuition”—that is, families must pay the difference between the voucher amount and higher school tuition (p. 3). This can have selection effects, skewing the voucher recipients toward those families who will/can pay the extra amount. Also noted is that the FTC has expanded its qualification parameters over the years to allow more affluent families to participate in the voucher program and that “private schools can continue to use their usual admissions processes to select applicants” (p. 3). These admissions processes often involve requirements concerning factors such as religion and LGBTQ status.

The report’s FTC analysis also notes that students who remained in private schools longer experienced considerable gains in outcomes when compared to their peers who spent less time in private schools. In doing so, it acknowledges that “we expect students who persisted in the program might differ in unmeasured ways from those who left” (p. 6), but it does not provide a meaningful discussion on the realities that might shape such differences. The continual financial burden of paying the difference in tuition as well as pushing out students for behavioral/academic concerns, religious beliefs, or LGBTQ status—all or some of these could account for the observed patterns.

Overall, the main threat to the validity of the relatively upbeat results presented in the report is straightforward: In the two case studies that yield some positive findings, the controls are inadequate to account for the likely impact of contextual factors. These factors can include parent and student motivation, the home, whether or not the family can afford tuition costs not covered by a voucher, and entrance requirements that work to screen out unwanted students. Further, readers may want to consider whether peer effects are part of any voucher “benefit.” Voucher advocates may see vouchers as beneficial in part because parents can choose preferred peers. But is this either fair or scalable, and does it trouble the ability to compare groups?\textsuperscript{14}

\textbf{VII. Usefulness of the Report for Guidance of Policy and Practice}

A critical point acknowledged by the report is that “participants and nonparticipants could differ in unmeasured ways, such as parental engagement, family religiosity, or experiences in public school” and “if these unmeasured characteristics differ, on average, between the treatment and comparison groups and are associated with student outcomes, our results will be biased” (p. 4). With this in mind, the report concedes that there is a possibility that there is a discernable difference between their comparison groups of students who receive

\url{http://nepc.colorado.edu/thinktank/vouchers}
and utilize vouchers and students who do not use vouchers (either as a result of not being offered one or choosing to not use the voucher).

Yet, despite this caveat, the report goes on to conclude that vouchers tend to have attainment benefits. The report’s comparison of students who attend private schools with those who attend public schools is very likely a comparison of apples to oranges because—in the two case studies suggesting some benefit—parents and students self-selected into a treatment group (likely impacting, e.g., a parent’s ability to provide transportation to the private school and/or supplement tuition costs not covered by a voucher). It is notable that the report’s one case study (in D.C.) using random assignment—thus minimizing self-selection and variances among control and treatment groups—found no benefits as a result of using a voucher. Additionally, while the report does acknowledge that “private schools [in Florida] can continue to use their usual admissions process to select applicants” (p. 3), it gives no further treatment of this as a possibility that can tarnish comparisons of treatment and control groups, as private schools are likely more selective in the admissions process.

Now that it is becoming clear that vouchers are not accomplishing their test-score aims, it appears expeditious for advocates to move the goalposts. Yet, as is evidenced in the Urban Institute report itself, even the newly suggested metrics for evaluating vouchers are producing results that are, at best, lackluster, with any seeming benefits likely the result of ignoring contextual selection effects.
Notes and References


2 See, for example:


5 Lee, J. (2019). *NEPC review: The effects of the Florida tax credit scholarship program on college enrollment*


7 See, for example:


10 See, for example:


14 See, for example,


