NEPC

NATIONAL EDUCATION POLICY CENTER

Review of *SCDP Milwaukee Evaluation Report #29*

Reviewed By

Casey D. Cobb

University of Connecticut

April 2012

Summary of Review

The School Choice Demonstration Project has published a series of reports written in the fifth and final year of its evaluation of the Milwaukee Parental Choice Program (MPCP). This review is of *Report #29*, a five-year longitudinal growth study, which found that a sample of elementary and middle school MPCP students outperformed a matched sample of Milwaukee Public Schools (MPS) students in reading in the fifth year of the program. The MPCP sample also showed trends of outscoring the MPS sample in math, but these were not statistically significant. The report acknowledges that the findings come as something of a surprise given that no differences were found between MPCP and MPS samples in the prior three years. The authors cite the introduction of a high-stakes accountability policy for MPCP schools just prior to that fifth and final year as a likely cause of the sudden spike in test scores among MPCP students. Overall, the report's methods are sound and its findings are appropriately qualified. Some of report's conclusions about the MPCP's effects on achievement growth appear overly general and potentially misrepresentative given the statistical results, however. The isolated jump in reading scores, but not math scores, in that fifth and final year of the study also raises uncertainty about what actually may have caused the sudden increase and what ultimately readers learn from the study.

Kevin Welner
Project Director

William Mathis
Managing Director

Erik Gunn
Managing Editor

National Education Policy Center

School of Education, University of Colorado
Boulder, CO 80309-0249
Telephone: (802) 383-0058

Email: NEPC@colorado.edu
http://nepc.colorado.edu

Publishing Director: Alex Molnar



This is one of a series of Think Twice think tank reviews made possible in part by funding from the Great Lakes Center for Education Research and Practice. It is also available at http://greatlakescenter.org.

This material is provided free of cost to NEPC's readers, who may make non-commercial use of the material 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.

Review of *SCDP Milwaukee Evaluation Report #29*

Casey D. Cobb, University of Connecticut

I. Introduction

The Milwaukee Parental Choice Program (MPCP) is currently the largest urban school voucher program in the nation. At present, nearly 21,000 students use a voucher of up to $6,442 to attend secular or religious private schools in Milwaukee. Voucher programs have long been targets of intense debates over the efficient and appropriate use of public funds for education.

In February 2012, the School Choice Demonstration Project[1] released a series of final reports from its five-year evaluation of the MPCP. The Wisconsin Legislature in 2005 required MPCP schools to administer nationally normed tests in grades 4, 8, and 10 to MPCP students, and to also submit the test scores to the School Choice Demonstration Project for purposes of evaluation. The reports address a range of issues related to the MPCP, including special education services, school climate and contexts, and test score performance. Of particular interest to policymakers has been the performance of MPCP students relative to students in the Milwaukee Public Schools (MPS). The report under review here is *Report #29: MPCP Longitudinal Educational Growth Study Fifth Year Report*. It is authored by John F. Witte, Deven Carlson, Joshua M. Cowen, David J. Fleming and Patrick J. Wolf [2]

II. Findings and Conclusions of the Report

This report represents the final installment of a five-year longitudinal evaluation of the MPCP. The longitudinal evaluation tracked the test score performance of panels of MPCP and MPS students initially enrolled in 2006-07. Evaluation reports have been issued since 2008, starting with a descriptive baseline report and followed each year since then by analyses estimating differences in achievement growth between MPCP and MPS samples. The findings from these reports have indicated "no meaningful differences in average test-score achievement between the two samples of students" (p. ii). The present fifth-year report, however, provides data indicating higher four-year growth (i.e., 2006 to 2010) among the MPCP sample on the Wisconsin Knowledge and Concepts Examination in reading relative to their MPS counterparts (641 matched pairs). The report cites additional evidence of greater growth in math scores among the MPCP sample, but these results were

not statistically significant. The report provides strong evidence that a high-stakes, test-based accountability policy applied to MPCP schools the prior year could have caused the sudden spike in reading performance. In its summary, the report concludes that "[o]ur supplemental analyses provide substantial evidence that the accountability policy could be responsible, in large part, for the higher achievement gains of the voucher students" (p. 25).

Additional analyses that examined student growth across along the entire distribution of scores suggested that lower-scoring MPCP students gained more than low-scoring MPS students in reading. Other analyses considered the relationship between time spent in MPCP and growth in achievement (a substantial number of MPCP students switched back to MPS over the course of five years, which permitted such exposure or "dosage" analyses). Findings indicated that spending "four or five years [in MPCP] results in greater achievement growth than spending only one or two" (p. iii). The report noted that "there is also evidence that, relative to spending zero years in MPCP, students who spend one, two, or three years exhibit lower achievement growth" (p. iii).

The final words of the report offer this summary:

> Our study established a common educational starting point for MPCP and MPS students. The gun sounded and a five-year race was run. The test-based accountability policy was applied to our MPCP runners during the final leg of the race, like a surge of adrenaline, and they clearly crossed the finish line ahead of their MPS peers in reading. That is what we learned from our Longitudinal Educational Growth Study and we think the lesson is an important one. (p. 24)

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

The rationale for the report's findings and conclusions are, for the most part, reasonable. Thorough care was taken to reduce or account for missing data. Multivariate and supplemental analyses were employed to confirm initial findings. The finding of greater average growth in reading in the final year among MPCP students was qualified by substantial evidence that the test-based accountability policy implemented the previous year was likely responsible "in large part" for the higher gains (p. 24). However, the insinuations of greater math achievement among voucher students were unsupported statistically. Further, some of the conclusions offered in the executive summary and closing sections are overly general and potentially misrepresentative given the report's analyses. For example, the report concludes, "the results…demonstrate that, conditional on spending any years in MPCP, spending four or five years results in greater achievement growth than spending only one or two" (p. iii). The report's results do not consistently support this conclusion. The results reported in Table 8 (p. 20), derived from the report's more robust analyses, show only one statistically significant finding (in reading) out of four possible findings in year 4—and this was at the very liberal and unconventional statistical probability level of $p<.10$. Similarly, in year 5 only one statistically significant finding was evident out of four possible, again in reading. No evidence in Table 8 indicated

MPCP students were better in math. Results presented in Tables 1 and 2 from the report's initial matched-pair analysis showed a trend toward "greater achievement growth" among MPCP students, but here again only two out of six results were statistically significant in math (Table 1), and two of six in reading (Table 2), after four and five years in the MPCP.

IV. The Report's Use of Research Literature

Several references were included that cast light on the short term leveraging effect of test accountability policy on school performance. The report cited very few evaluations and peer-reviewed studies of voucher programs and their effect on academic achievement although many prominent studies are available. For instance, it did not cite examinations of Cleveland's Scholarship and Tutoring Program,[3] recent studies of voucher programs in Chile,[4] and other peer-reviewed articles on vouchers.[5]

V. Review of the Report's Methods

Inferences about the effects of the MPCP on student academic growth compared with those of the MPS hinge on the comparability of the two group's initial samples. Considerable attention was given to the process by which MPS students were matched in the study to MPCP counterparts. The report uses sophisticated matching techniques, including matching students by race, gender, academic ability, and neighborhood location. Follow-up sensitivity analyses provided evidence that selection bias was not a particular concern. The growth analysis relied on a sample of students in grades 3, 4, and 6 at the baseline year 2006 who were also in grades 7, 8, and 10 four years later (for a combined MPCP and MPS sample of 1,282 students).[6]

Several analyses were conducted, including fairly straightforward mean comparisons between sectors (see the report's Tables 1 and 2) and more sophisticated linear regressions that modeled achievement as conditional upon student baseline characteristics, such as gender, race, and test scores (the report's Tables 4 and 5). Tables 4 and 5 present findings using two different sampling approaches: one that ignores sector switching and another that considers only those students who remained in their initial sector throughout the duration of the study. In the first analysis, students who switch from MPCP to MPS is treated as if they remained in MPCP all along (and vice versa). The report offers justification for counting MPCP leavers as permanent MPCP students by noting that this approach is standard protocol in clinical medical trials. In the second analysis, switchers were dropped. That is, the 849 students who stayed in the same sector across all four years were analyzed separately.

Student mobility is a concern for all longitudinal studies of this sort. Mobility refers to students switching schools or districts, transferring across public-private sectors, leaving the state, or dropping out of school entirely. For this study, the level of attrition from the original panels of 4,007 students was 32% by year five. The report finds this attrition rate to be better than expected and superior to those in similar studies. The report examines whether the attrition unduly biased the findings by comparing student characteristics of

those missing from the study to those who remained. Comparisons were also made between missing students in the two sectors; no differences worth noting were evident, save for the lower baseline math scores found among missing MPCP students relative to missing MPS students. All differences were statistically adjusted in subsequent multivariate analyses. Missing data were recovered or addressed in a variety of other ways, including through telephone inquiries, database searches, and additional statistical adjustments. Overall, the methods employed were sound.

VI. Review of the Validity of the Findings and Conclusions

The findings presented in the report's Tables 4 and 5 are particularly noteworthy due to the advanced level of analysis. Each table shows statistically significant "effects" of the MPCP on reading scores, even after controlling for student race, gender, and baseline test score. The effects of the MPCP are presented in standard deviation units, or effect sizes, which range from 0.07 to 0.15. Effects on reading scores were statistically significant across all three models but low in magnitude with effect sizes of 0.12 and 0.15. Effects on math scores were not statistically significant. Table 4 displays results for a sample of around 1,300 students across both sectors, and does not make distinctions for students who switch out of their initial sector. Table 5 shows results for a much smaller combined group of "stayers" or "non-sector switchers." Recall that these participants stayed in their original groups.

The report describes its analysis of 849 students who remained in their initial sectors all four years subsequent to baseline as "the purest and sharpest contrast regarding the extent to which school sector influenced student achievement" (p. 13). It does, however, offer caution due to the inequality of the two sector groups:

> Thus, although this represents perhaps the cleanest contrast between the MPCP and MPS students in our sample, it does not necessarily provide the best estimate of the average effect of the MPCP on achievement because stayers—in both MCP and MPS—are a distinctive subgroup of students (p. 13).

The report's Figure 3 presents point estimates for achievement differences between MPCP and MPS. The point estimates are presented with confidence intervals, which help the reader discern the range of possible point estimates. However the report uses 90% confidence intervals. There is nothing inherently wrong with using 90% confidence intervals but one would expect to see 95% confidence intervals, which would be consistent with conventional levels of statistical probability (i.e., $p<.05$). There does not appear to be much gained here by using 90% intervals, other than the fact that the intervals themselves are shorter than what would be produced by 95% intervals. The difference in 2010 reading scores is still statistically significant at the .05 level.

As a whole, the statistical findings in the report are valid. As noted above in section III, some of the summative conclusions about the MPCP's effects on achievement growth are too general given the statistical results presented in the report.

VII. Usefulness of the Report for Guidance of Policy and Practice

It is unclear how this report contributes to policy and practice. Three prior reports using similar methodology found no differences in achievement between MPCP and MPS students. This would suggest that the MPCP does not provide a value above and beyond the education provided by MPS. This final report, however, offers evidence of significantly higher reading performance among MPCP students who started in the program four years earlier. The increase was sudden and large, and it immediately followed the introduction of a high-stakes-test accountability policy. Perhaps the most useful aspect of the report, then, is this apparent influence of a high-stakes test accountability policy on voucher students' performance. Unclear are the potential effects of the accountability policy on MPCP student mobility in that last year.

The concluding text of the report, which makes use of a racing metaphor, is somewhat confusing. As noted above, the final sentences read, "our MPCP runners . . . clearly crossed the finish line ahead of their MPS peers in reading. That is what we learned from our Longitudinal Educational Growth Study and we think the lesson is an important one" (p. 24). It is unclear, however, what exactly the lesson learned might be, and why it is important. Is it that MPCP students do better in reading than MPS students in their fifth year in the program, even after demonstrating no superior achievement in the first four years? Is it that a high-stakes accountability policy influences actions and behaviors in some way that leverages achievement gains, at least in reading? Elaboration on this final point would have been helpful to policymakers trying to make sense of the report's findings.

Notes and References

1 The School Choice Demonstration Project (SCDP) is housed in the Department of Education Reform at the University of Arkansas (http://www.uark.edu/ua/der/SCDP/Research.html).

2 Witte, J. F., Carlson, D., Cowen, J. M., Fleming, D. J., & Wolf, P. J. (2012). *MPCP Longitudinal Educational Growth Study Fifth Year Report*. Fayetteville: University of Arkansas School Choice Demonstration Project. Retrieved February 27, 2012, from http://www.uaedreform.org/SCDP/Milwaukee_Research.html.

3 Plucker, J., Muller, P, Hansen, J., Ravert, R., & Makel, M. (2006). *Evaluation of the Cleveland Scholarship and Tutoring Program: Summary Report, 1998- 2004*. Center for Evaluation & Education Policy. Indiana University;

Belfield, C. (2006). *The evidence on education vouchers: An application to the Cleveland Scholarship and Tutoring Program*. National Center for the Study of Privatization in Education. Queens College, City University of New York.

4 Hsieh C.-T. & Urquiola M. (2006). The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program. *Journal of Public Economics, 90* (8-9), 1477-1503;

Lara, B., Mizala, A., & Repetto, A. (2011). The effectiveness of private voucher education: Evidence from structural school switchers. *Educational Evaluation and Policy Analysis, 33*(2), 119-137.

5 Lubienski, C., Weitzel, P., & Lubienski, S. (2009). Is there a consensus on school choice and achievement?: Advocacy research and the political economy of knowledge production. *Educational Policy, 23*(1), 161-193;

McEwan, P. J. (2004). The potential impact of vouchers. *Peabody Journal of Education, 79*(3), 57-80;

Rouse, C. E., & Barrow, L. (2008). School vouchers and student achievement: Recent evidence, remaining questions. *Annual Review of Economics*, *1,* 17-42.

6 It is not entirely clear why 1,282 students are reported as qualifying for the growth analysis while Table 4 shows *N*s of just over 1,300.

DOCUMENT REVIEWED:	**SCDP Milwaukee Evaluation Report #29: MPCP Longitudinal Educational Growth Study Fifth Year Report**
AUTHORS:	John F. Witte, Deven Carlson, Joshua M. Cowen, David J. Fleming, & Patrick J. Wolf
PUBLISHER/THINK TANK:	The School Choice Demonstration Project, Department of Education Reform, University of Arkansas
DOCUMENT RELEASE DATE:	February 2012
REVIEW DATE:	April 19, 2012
REVIEWER:	Casey D. Cobb, University of Connecticut
E-MAIL ADDRESS:	casey.cobb@uconn.edu
PHONE NUMBER:	(860) 486-6278

SUGGESTED CITATION:

Cobb, C. D. (2012). *Review of "Report 29 of the SCDP Milwaukee Evaluation."* Boulder, CO: National Education Policy Center. Retrieved [date] from http://nepc.colorado.edu/thinktank/review-Milwaukee-Choice-Year-5.