



REVIEW OF *BANG FOR THE BUCK*

Reviewed By

Casey D. Cobb
University of Connecticut

July 2016

Summary of Review

The authors of this report created “efficiency scores” for Milwaukee schools and, on the basis of those scores, draw conclusions about the relative efficiency of Milwaukee’s charter and traditional public schools. A first set of scores was generated by dividing test scores in math and science by per-pupil funding. In a second analysis, “efficiency scores” were estimated for school types using a regression model that also accounted for selected student demographic characteristics. Several major problems arise from the authors’ approach: First, test scores do not comprehensively represent the purposes of schools. Second, threats to the validity of the author’s assumption that there is uniform financial accounting across schools and types are not addressed. Third, the analytic description of the study was incomplete, making interpretation difficult. Fourth, the report makes strong but unmeasured claims about the superior “efficiency” of charter schools based on their having greater autonomy. Finally, the authors did not appear to adjust for selection effects – effects that would prove fatal to their conclusions. As a result of its manifold flaws, and because its conclusions are not supported by the evidence presented, the report is of little if any use to policymakers.

Kevin Welner

Project Director

William Mathis

Managing Director

Alex Molnar

Publishing Director

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>

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 nepe@colorado.edu.

This review is one of a series made possible in part by funding from the Great Lakes Center for Education Research and Practice.



<http://www.greatlakescenter.org>
GreatLakesCenter@greatlakescenter.org

REVIEW OF *BANG FOR THE BUCK*

Casey D. Cobb, University of Connecticut

I. Introduction

The report under review, “Bang for the Buck: Which public schools in Milwaukee produce the best outcomes per dollar spent?” was released by the Wisconsin Institute for Law & Liberty and co-authored by the Institute’s Education Research Director, Will Flanders, and its Vice President for Policy, CJ Szafir.¹ The report generates “efficiency scores” by dividing test scores by per-pupil spending scores for each public school in Milwaukee and draws conclusions about the relative efficiency among charter and traditional public schools operating in that city.

Wisconsin law permits three different types of charters. These include instrumentality charter schools, which are authorized by a home school district and staffed with district employees; non-instrumentality charter schools, which are authorized by a home district but allow staff to be hired from outside the district; and independent charter schools, which are authorized by state-approved higher education institutions or the Milwaukee Common Council and whose staff is hired by one of these governing boards. All three types of charters operate in Milwaukee alongside traditional Milwaukee Public Schools (MPS). According to the report, non-instrumentality and independent charter schools in Milwaukee receive roughly \$2,200 less per-pupil than MPS and instrumentality charter schools (\$8,075 vs. \$10,261 in 2014-15). The researchers assert, without sufficient documentation or rationale, that non-instrumentality and independent charter schools have higher degrees of autonomy relative to MPS and instrumentality charters, and set out to determine if such autonomy results in more efficient performance among these charter school types.

Many charter school advocates claim charter schools are underfunded compared to traditional public schools. They also contend that the autonomy and freedom granted to charter schools allows them to run more efficiently—that is, to perform better given the same funding--than traditional public schools. These two themes serve as a broader context for the present report that examines “which public schools produce the best outcomes per taxpayer dollars spent.” (p.1)

II. Findings and Conclusions of the Report

The report concludes that the most efficient public schools in Milwaukee are charter schools that experience greater autonomy from the district. It finds that non-instrumentality and independent charter schools “do more with less” (p.4). In particular, the report highlights independent charters, which it claims are “far more efficient in their expenditures than public schools” (p.16). According to the report, independent charter schools enjoy the most freedom from district oversight, followed by non-instrumentality charters. The report argues that instrumentality charter schools are not truly charters because they are staffed by MPS teachers, are afforded limited administrative flexibility, and follow traditional district curriculums.

The report offers four recommendations for policymakers based on its findings. First, it calls for closer examination of the “comparative inefficiencies” (p.16) of traditional public schools and instrumentality charter schools. Second, it recommends that instrumental charters be reclassified as something other than charter schools, due to their lack of autonomy from district authorizers. Third, the report recommends expanding the number of charter school authorizers in Wisconsin, as well as their authoritative purview. Lastly, the report calls for increased student access to non-instrumentality and independent charter schools in Wisconsin, given their relative effectiveness and efficiency.

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

The report’s conclusions are based on comparisons of “efficiency scores” among charter and traditional public schools in Milwaukee. The authors calculated efficiency scores in two different sets of analyses and reported scores at the individual school level and by school type (i.e., traditional MPS, instrumental charter, non-instrumental charter, independent charter).

In the first analysis, as noted, state test scores in math and science were divided by per-pupil funding to create an efficiency score for each school. Schools were rank-ordered by their efficiency scores and identified by school type (see Appendix A of the report). Higher scores indicate greater efficiency than lower scores. In an attempt to control for student poverty, which is known to be highly correlated with student achievement, schools serving 80% or more students eligible for free and reduce priced lunch were pulled out of the larger list and rank-ordered separately (see Appendix B).

The authors use the rank-ordered lists to draw inferences about the relative efficiencies among public school types. For instance, the bar charts on page 3 of the report show the “top ten most efficient schools” and “ten least efficient schools” in math and science. Here are two additional claims and their warrants:

“4 of the top 5 schools for efficiency are independent public charter schools (Ind)...Nine of the top 10 schools for efficiency are independent (Ind) or non-instrumentality (Non-Instr) schools...Put another way, even with the most difficult students, dollars spent on independent and non-instrumentality charter schools are much more effective.” (p.2)

“The most inefficient schools with taxpayer money are dominated by traditional MPS. The bottom ten schools who took the Badger [math] Exam are all traditional MPS.” (p.3)

In the second analysis, efficiency scores were estimated for school types using a production function regression model. In the first step, 2014-15 state test scores in math and science were regressed against charter school type (MPS served as the reference group), school level controls (i.e., percentage of free and reduced price lunch students, percentage of English language learners, percentage of non-white students), and testing grade (i.e., grade 4, grade 8). The regression coefficients for each charter school type were then divided by their respective per-pupil funding allotments to create efficiency scores for each school type in both math and science. The MPS efficiency score means were estimated using the regression equation. In the second step of the process, charter school type means were contrasted with MPS means to set up “difference-in-mean efficiency estimates” (p.13). Mean differences in efficiency between each charter type and MPS were divided by what appears to be a standard error of the means (see report’s equation 3) to test the statistical significance of each charter-MPS mean difference. The results are presented in Tables 5 and 6 of the report.

The report’s conclusions rest on certain key assumptions. Among them:

- The performance of a school (and a school type) can be represented by aggregate student scores on statewide exams in science (in grades 4, 8, and 10) and math (in grades 4 and 8).
- Per-pupil funding as reported by the Wisconsin Department of Public Instruction represents a complete account of how each school (and each school type) expends its public resources. And further, that per-pupil funding accounts for all of the funds received and expended by MPS and charter schools in Milwaukee.
- The report assumes school efficiency can be validly measured as a function of a school’s performance on state science and math exams given that school’s per-pupil funding, and the demographic characteristics of students it serves.

While there is uniform agreement that student selection effects are present in charter schools as compared to traditional schools, they are not adequately accounted for in this case. In fact, “selection bias” is not even mentioned by the report.

IV. The Report's Use of Research Literature

The report does not rely heavily on existing research literature. The report references two studies that examine the efficiency of schools on an international scale. These include a study published by GEMS Education Solutions that ranked the US 21st out of 30 industrialized countries and a study by the publisher of the present report that compared Wisconsin to OECD countries.²

The authors of the report indicate that they “are aware of only three existing studies that examine charter school efficiency” (p.5) and they provide summaries of each. Two of these reports were published in peer-reviewed journals and analyzed charter schools in Texas. The third was published by the School Choice Demonstration Project at the University of Arkansas.³ However there are several other peer-reviewed studies published on this topic beyond the three cited in this report. These include a 2010 article published in the *Journal of Education Finance* titled, “Measuring Charter School Efficiency,” that found non-charter schools were more efficient than charter schools.⁴ In addition, Palardy and colleagues conducted two studies, one in 2007 that found charter schools in Arizona were less efficient than traditional public schools and another in 2015 that found Ohio charter schools increased technical efficiency at faster rates than traditional public schools.⁵ Another peer-reviewed study found Texas charter schools to be less efficient than traditional public schools.⁶ Most of these publications offered brief critiques on methods used to examine school efficiency and identified stochastic frontier analysis as the most advanced method to date.

The report is selective in its use of research literature on charter schools. For instance, it states “[m]ost existing research has found that public charter schools earn better outcomes than traditional public schools” (p.1) but does not offer any citations for the claim. Such a claim is misinformed given the contested research terrain comparing charter and traditional public school effectiveness. A heavy volume of studies reveals mixed findings or point to the relative ineffectiveness of charters when compared to traditional public schools.⁷

V. Review of the Report's Methods

As described above, the report generated “efficiency scores” by school and school type. Scores were calculated in the first analysis by dividing school test performance in math and science by per-pupil funding in a single school year, 2014-15. School test performance was measured by aggregating student scores on statewide exams in science for grades 4, 8, and 10 and in math for grades 4 and 8.⁸ The student test scores that were aggregated were not raw or scale scores but rather performance level scores (e.g., 1=below basic, 2=basic, 3=proficient, 4=advanced). Using discrete performance levels instead of scale scores reduces variability in the student achievement measure and is thus less discriminating

across students—and ultimately limits the capacity to detect statistical relationships, if they exist.

The efficiency scores in the first analysis were rank ordered by school. In a good faith attempt to control for student poverty, a separate listing of schools that served 80% or more free and reduced price lunch students was also included. The efficiency scores from the first analysis are not readily interpretable in and of themselves; rather, they allow for relative comparisons across schools. Even when interpreted on a relative scale, it is difficult to discern the meaning of differences between schools. For instance, what is the substantive difference between a school that scored a .40 vs. a school that scored .30? Further, comparisons among schools serving different grade levels (e.g., Carmen High School vs. Lowell Elementary) are problematic as the test scores cannot be compared over a broad grade span and cost structures for the two types of schools may vary.

The report is selective in its use of research literature on charter schools.

The second set of analyses was designed to generate efficiency scores by school type. To do so, average test scores were regressed against school type and additional control variables, such as the percentage of students eligible for free and reduced price lunch, the percentage of nonwhite students, the percentage of English language learners in the school, and grade level tested. The control variables were included in the model to help account for other factors that might influence school performance scores. The regression seemed to be executed appropriately, although it was not clear whether any weighting was used to account for varying school sizes. Weighted regression would be the preferred method considering the variability in school enrollments in the model.

The regression yielded coefficients for each charter school type (with MPS serving as the reference group) that were then divided by the average per-pupil funding within that school type to produce a school-type efficiency score. These efficiency scores were then compared across school types.

Here again, the efficiency scores generated by school type are somewhat challenging to interpret. It is not entirely clear what the scores represent in Tables 5 and 6 of the report. As the statistical analyses were incompletely described, whether proper analytic techniques were used was not possible to fully evaluate.

The second set of analyses relies on a production function model, where a set of school-level inputs (i.e., percentage of students eligible for free and reduced price lunch, percentage of English language learners, percentage of non-white students, grade levels tested) is specified to relate to a certain output (i.e., test scores in math and science). The use of production function models that employ multiple regression has been critiqued elsewhere;⁹ however, these limitations were not cited by this report. Nor does the report indicate why multiple regression was chosen over more sophisticated analytic techniques seen in the school efficiency literature (such as stochastic frontier analysis).¹⁰

Although primary findings of the study were presented as autonomy increases efficiencies, no measure of autonomy was included. The report categorizes the autonomy of the four school types in a table on page 1 with instrumentality="limited", non-instrumentality="high", independent="high" and MPS="none." However the autonomy construct and its score categories are not defined in the report, nor are they grounded in the theoretical or empirical literature.

Finally, the report has a relatively high incidence of typos and missing information. Mistakes are to be expected in any human endeavor, but to the extent that they occur here makes it difficult to interpret the report's analyses and raises questions about the report's overall quality. For instance, in Table 5 of the report (p.14), the mean score for "Public" should be .306, not .036, based on the other mean differences in the table. Similarly, in Table 6, the "Difference from Public" score under the "Independent" column should be .101, not .111, if indeed Independent – Public is .365 - .264. On page 17, endnote 17 refers to appendix B and C but there is no appendix C in the report.

VI. Review of the Validity of the Findings and Conclusions

The report's conclusions about the efficiency of individual public schools in Milwaukee (the first analysis) rests almost entirely on the dubious assumption that school efficiency can be captured by dividing scores on math and science exams in a single year by per-pupil funding. There are many problems with this logic. I speak to two fundamental ones here.

First, per-pupil funding does not take into account how such funds were used or parsed for what purpose. In addition to funds that directly support instruction, a portion of funds can be also used for administration, counseling, transportation, or special services to students with disabilities. The per-pupil funding for non-instrumentality and independent charter schools may be lower than that of MPS and instrumentality schools, but the figure does not necessarily account for other costs absorbed by sponsoring districts or expended through other contractual agreements. The 2014-15 Wisconsin Charter Schools Yearbook, published by the Wisconsin Department of Public Instruction, indicates that in those instances where a charter school functions with less money [than its sponsoring district], it "can happen if a charter school shares an existing district facility, and shares management costs with the school district, participates in district services such as co-curricular activities, special education, psychological services, and food service" (p.7).¹¹ Given the likelihood that non-instrumentality charters appreciate some of these district economies of scale, it would be inappropriate to assume per-pupil funding averages as used in the report represent the complete account of public funds expended in those schools.

A second problem lies on the output side of the efficiency equation, which relies on test scores in math and science in a single year, across only tested grades, to capture an entire school's performance. Doing so ignores many other valued outcome measures such as

graduation rates, post-secondary rates, student extracurricular success, and student achievement in the arts, reading, writing, and social studies. Further, it is grossly oversimplistic to assume that a test score in a single year represents a school's unique contribution to student achievement in those subjects. Schools obviously affect student learning, but there are myriad other influences on student learning that cannot be accounted for by statistical controls. Students are not randomly distributed across charter and public schools, and most statistical controls are unable to attenuate such selection bias. As parents of children choosing a charter school are different from non-choosers, it is unclear as to whether the differences in test scores are due to selection effects, the type of school or some other "third factor." The correlational research design employed here is simply insufficient to lend support to any causal statements about effectiveness or efficiency in performance (of which this report offers many).

To its credit, the report acknowledges at least two serious limitations of the study, including cautions with inferring causal relationships from statistical analysis and examining only one year of data. Despite these acknowledgements, there are several instances where the report uses language that strongly implies causality. As one example, the report stretches the utility of its efficiency estimates for school type, claiming "[t]he efficiency estimate for Independent charter schools is .408, indicating that, on average, a \$1000 change in expenditure in this type of school will yield .408 more points on the 4-point scale of the WKCE" (p.13). Even if one buys the authors' efficiency estimate, all they are doing is estimating a test score difference per dollar spent. Nothing in the data or this correlational research design would indicate that spending an additional \$1,000 would result in a proportional increase in test score.

The report further propositions,

If a conclusion must be drawn, it would seem that the more autonomous a school is, the better it performs. This is consistent with much of existing research. This model should be replicated and the state should encourage policies that incentivize such schools. (p.4)

Given this statement it is more than a little odd that no data were collected to validate a school's level of autonomy, nor was the study design sufficient to test its potential influence on school performance.

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

In the report, strong claims are made about the relative efficiencies of charter schools in Milwaukee, particularly those with more autonomy from traditional school districts. For instance the report concludes that the "rankings [of schools in Table 3] strongly suggest that independent and non-instrumentality charters are better stewards of taxpayer money" (p.12). Since a huge variety of factors are in play, relative strength of "stewardship" is

merely speculative. And at the heart of the matter, such conclusions are based on the construction of school “efficiency scores” that are of questionable value. Therefore the report is potentially misleading to policymakers making decisions in challenging fiscal climates. Taken on the whole, the research warrants are simply too weak to support the report’s very strong claims.

Notes and References

- ¹ Flanders, W., & Szafir, CJ (2016). *Bang for the buck: Which public schools in Milwaukee produce the best outcomes per dollar spent?* Wisconsin Institute for Law & Liberty. Retrieved June 9, 2016, from <http://www.will-law.org/wp-content/uploads/2016/05/Bang-for-the-Buck-FINAL.pdf>
- ² The report cites “Dalton [sp], Marcenaro-Gutierrez and Still” in its narrative (p.5) but does not list it in the endnotes or any reference list. Based on my own search, I had to assume that the study the report is referring to is:

Dolton, P., Marcenaro-Gutierrez, O., & Still, A. (2014). *The efficiency index: Which education systems deliver the best value for money?* GEMS Education Solutions. Retrieved on June 9, 2016, from <http://www.edefficiencyindex.com/book/>. Incidentally, this report was reviewed by NEPC (<http://nepc.colorado.edu/newsletter/2014/09/review-efficiency-index>).

The second study cited was “Leuken, Esenberg and Szafir (2015)” (p.5) and published by the Wisconsin Institute for Law & Liberty, the same publisher of the report under review here. This study also was not referenced in the report. I had to assume this was the following report:

Leuken, M.F., Esenberg, R., & Szafir, CJ (2015). *Diminishing returns in K12 education: Has Wisconsin hit a wall where an additional dollar in education spending will not bring improvements in student outcomes?* Wisconsin Institute for Law & Liberty. Retrieved on June 9, 2016, from <http://www.will-law.org/wp-content/uploads/2016/02/2015-04-03-WILL-Report-Diminishing>Returns-in-Education.pdf>
- ³ This report was reviewed by Gene Glass of the NEPC. See <http://nepc.colorado.edu/files/ttr-charter-school-productivity.pdf>
- ⁴ Carpenter, D.M., II, & Noller, S.L. (2010). Measuring charter school efficiency: An early appraisal. *Journal of Education Finance*, 35(4), 397-415.
- ⁵ Palardy, J., & Nesbit, T. (2007). Traditional public schools versus charter schools: a comparison of technical efficiency. *Economics Bulletin*, 9(9), 1-10.

Palardy, J., Nesbit, T.M., & Adzima, K. A. (2015). Charter versus traditional public schools: A panel study of the technical efficiency in Ohio. *Education Economics*, 23(3), 278-295.
- ⁶ Gronberg, T.J., Jansen, D.W., & Taylor, L.L. (2012). The relative efficiency of charter schools: A cost frontier approach. *Economics of Education Review*, 31(2), 302-317.
- ⁷ For example, see:

Bifulco, R., & Ladd, H.F. (2006). The impacts of charter schools on student achievement: Evidence from North Carolina. *Education Finance and Policy*, 1(1), 50-90;

Carnoy, M., Jacobsen, R., Mishel, L., & Rothstein, R. (2005). *The charter school dust-up: Examining the evidence on enrollment and achievement*. Washington, DC: Economic Policy Institute;

Lubienski, S.T., & Lubienski, C. (2006). School sector and academic achievement: A multilevel analysis of NAEP mathematics data. *American Educational Research Journal*, 43(4), 651-698;

Ni, Y., & Rorrer, A.K. (2012). Twice considered: Charter schools and student achievement in Utah. *Economics of Education Review*, 31(5), 835-849;

Powers, J.M., & Glass, G.V (2014). When statistical significance hides more than it reveals. *Teachers College Record*. Retrieved June 9, 2016, from <http://www.tcrecord.org/Content.asp?ContentID=17591>;

Yeh, S.S. (2013). A re-analysis of the effects of KIPP and the Harlem Promise Academies. *Teachers College Record*. Retrieved June 9, 2016, from <http://www.tcrecord.org/content.asp?contentid=16917>;

Zimmer, R., Gill, B., Booker, K., Lavertu, S., Sass, T.R., & Witte, J. (2009). Charter schools in eight states: Effects on achievement, attainment, integration, and competition. Washington DC: RAND Corporation. Retrieved June 9, 2016, from <http://www.rand.org/pubs/monographs/MG869.html>;

Zimmer, R., Gill, B., Booker, K., Lavertu, S., & Witte, J. (2012). Examining charter student achievement effects across seven states. *Economics of Education Review*, 31(2), 213-224.

- ⁸ Wisconsin Knowledge and Concepts Exam (WKCE) science exam was administered in 4th, 8th, and 10th grades; 4th and 8th grade math scores from the Badger Exam were included in the study.
- ⁹ Carpenter, D.M., II, & Noller, S.L. (2010). Measuring charter school efficiency: An early appraisal. *Journal of Education Finance*, 35(4), 397-415.
- ¹⁰ Several school efficiency studies have used analytic techniques that have advantages over straightforward multiple regression models. See:

Grosskopf, S., Hayes, K.J., & Taylor, L.L. (2014). Efficiency in education: Research and implications. *Applied Economic Perspectives and Policy*, 36(2), 175-210.
- ¹¹ Evers, T., Pahnke, B., Collins, T., Abel, J., & Eagleburger, S. (2014, September). Wisconsin charter schools yearbook, 2014-2015. Madison, WI: Wisconsin Department of Public Instruction. Retrieved June 13, 2016, from <http://dpi.wi.gov/sites/default/files/imce/sms/pdf/2014-15yearbook.pdf>

DOCUMENT REVIEWED: Bang for the Buck: Which Public Schools in Milwaukee Produce the Best Outcomes Per Dollar Spent?

AUTHORS: Dr. Will Flanders and CJ Szafir

PUBLISHER/THINK TANK: Wisconsin Institute for Law & Liberty

DOCUMENT RELEASE DATE: May 19, 2016

REVIEW DATE: July 12, 2016

REVIEWER: Casey D. Cobb, University of Connecticut

E-MAIL ADDRESS: casey.cobb@uconn.edu

PHONE NUMBER: (860) 486-0253

SUGGESTED CITATION:

Cobb, C.D. (2016). Review of "*Bang for the Buck*." Boulder, CO: National Education Policy Center. Retrieved [date] from <http://nepc.colorado.edu/thinktank/review-WILL>