NEPC Review: Charter Management Organizations 2017 (CREDO, June 2017)

Reviewers:

Gary Miron and Christopher Shank
Western Michigan University

September 2017

National Education Policy Center

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

Kevin Welner
Project Director

William Mathis
Managing Director

Alex Molnar
Publishing Director

Publication Type: NEPC Reviews are expert third-party reviews of selected non-peer-reviewed publications. Using academic peer review standards, reviewers consider the quality and defensibility of a report's assumptions, methods, findings, and recommendations. Written in non-academic language, reviews are intended to help policy makers, reporters, and others assess the social science merit of a reviewed report and to judge its value in guiding policy.

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

Suggested Citation: Miron, G. & Shank, C. (2017). *NEPC Review: "Charter Management Organizations 2017."* Boulder, CO: National Education Policy Center. Retrieved [date] from http://nepc.colorado.edu/thinktank/review-CMOs.

NEPC
NATIONAL EDUCATION POLICY CENTER

NEPC Review: Charter Management Organizations 2017 (CREDO, June 2017)

Reviewers:

Gary Miron and Christopher Shank
Western Michigan University

Summary of Review

The Center for Research on Education Outcomes (CREDO) assessed the impact of different types of charter school-operating organizations on student outcomes in 24 states, plus New York City and Washington, D.C. The study finds that students in charter schools display slightly greater gains in performance than their peers in traditional public schools, especially students in charter schools operated by certain types of organizations. CREDO's distinctions between organization types are, however, arbitrary and unsupported by other research in the field. This raises concerns about the practical utility of the CREDO findings. In addition, CREDO researchers made several dubious methodological decisions that threaten the validity of the study. A number of the problems flagged by the reviewers of this study have been raised in reviews of prior CREDO studies. Specifically, CREDO studies tend to over-interpret small effect sizes; to fail to justify the statistical assumptions underlying the group comparisons made; and to not take into account or acknowledge either the large body of charter school research beyond their own work or the limitations inherent in the research approach they have taken. Because these problems have gone unaddressed in the present study, and because the CREDO researchers have compounded them by creating a confusing and illogical charter organization classification system, the report is of limited if any value in policymaking. Readers should review the report with care.

NEPC
NATIONAL EDUCATION POLICY CENTER

NEPC Review: Charter Management Organizations 2017 (CREDO, June 2017)

Reviewers:

Gary Miron and Christopher Shank
Western Michigan University

I. Introduction

This review examines the report produced by the Center for Research on Education Outcomes (CREDO) entitled "Charter Management Organizations, 2017".[1] The CREDO report examines how differences in school management structures impact academic growth of charter school students. This report was sponsored by the Walton Family Foundation and the Fisher Fund; both of these sponsors advocate on behalf of charter schools and provide funding to charter schools.

Research on charter schools is a vital tool for informing both policymakers and the public. Starting in the 1990s, for-profit education management organizations, most prominently Edison Schools Inc., fueled the belief that private entities could operate charter schools and produce better results while still managing to secure a profit. When myths surrounding Edison's performance and finances were dispelled,[2] the philanthropic sector shifted their focus to nonprofit education management organizations (EMOs) as the vehicle for charter school expansion. While most nonprofit EMOs have similar management agreements and fee structures, a small subset of nonprofit EMOs have received extensive funding from philanthropic sources as they attempt to bring what are deemed as successful school models to scale. The New Schools Venture Fund—one of the largest financial supporters of this subset of nonprofit EMOs—refers to them as charter management organizations. While both for-profit and nonprofit EMOs have promoted themselves and their accomplishments, only a few studies have explored the relative performance of these private education management organizations.[3] There is a need for closer investigation and questioning of the unverified claims of success made by these private entities. Broadly speaking, policymakers and the general public need to know more about these publicly funded but private entities and whether their involvement with charter schools is wise and effective public policy.

The *Charter Management Organizations 2017* report by CREDO contributes to an expan-

sive body of existing research on charter schools, and expands upon CREDO's earlier work by reexamining their data sets to explore differences in charter school performance based on type of management organization.

II. Findings and Conclusions of the Report

The report comprises 111 pages including appendices. The analyses includes 24 states as well as New York City and Washington D.C. Key findings note that hybrid management organizations had more positive results relative to other organization types (definitions of organization types are included in Section V below), nonprofit charter organizations had better outcomes than for-profit organizations, and charter school organizations have stronger effects with Black and Hispanic students.

Results varied noticeably across states and between charter management organizations. The differences between key charter school types in the study were very small and according to common practice could be considered noninterpretable or not discernable. As with CREDO's earlier reports, the authors chose to interpret minute differences between groups and to suggest that these differences were in fact large and meaningful. Although differences between groups or school types were small, their results for individual management organizations revealed some differences that were sufficiently large and meaningful to interpret. Other key findings highlighted in the report include the following:

- Charter school operators that hold non-profit status post significantly higher student academic gains than those with a for-profit orientation. For-profit operators have results that are at best equal to the comparison traditional public school students in reading or worse in math.
- The effectiveness of charter school organizations varies extensively across states.
- Schools that contract with external vendors for much or all of the school operations post significantly lower results than network operators that maintain direct control over their operations.
- Online charter schools continue to present significantly weaker academic performance in reading and math compared with their counterparts in conventional schools.

A key conclusion highlighted in the report is there are wide variations in student performance in schools operated by private management organizations. Authorizers and charter school boards should be aware of this so that they can make informed decisions on which management organizations should be encouraged to grow.

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

Similar to earlier CREDO studies, findings are based on a quantitative analysis of existing state assessment data that is compiled into longitudinal datasets by CREDO. For this particular report, the dataset includes student-level data from 26 states plus New York City and the District of Columbia. Although full details on the methods are not included, it is apparent that conclusions are based on the analysis of the student-level data set.

The data collection and analytic methods are not sufficiently described in the report. In addition to some details included in the report, there was a 1.5-page technical appendix (Appendix B) that also touched on methods. The absence of details on methods led us to conclude that methodological limitations from earlier reports might also apply to this report. Data sources and the methods will be discussed further in Sections V and VI, below.

IV. The Report's Use of Research Literature

The report does not contain a literature review and little effort was made to link the findings to the larger body of evidence related to charter school impact on student learning.

There is a very broad research base on charter schools and their performance on standardized tests. The research base has clearly improved over time. In the mid-1990s, this research largely relied on group level data. Increasingly, the research includes a range of outcome measures and incorporates student-level data. Over the past decade there have been a large number of studies and evaluations that draw upon student-level data which has resulted in more rigorous designs. The review by Miron and Applegate (2009)[4] provides an overview of major relevant research studies not referenced by CREDO. Since 2009, the level of rigor of new studies has continued to increase as more researchers are gaining access to student level data.

Besides citing their own work, the report does include citations for a few studies that found that charter schools performed better than traditional public schools. Interestingly, the most expensive and rigorous study to date, which was commissioned by the US Department of Education and involved a simulated random-assignment experimental design, is neither cited nor recognized by CREDO. This study by Mathematica[5] showed a very small negative effect for charter schools that have large waiting lists (i.e., popular charter schools that had sufficiently large waiting lists so that a comparison group could be created. Schools with few or no students waiting to get in were excluded). The authors recognized that the test result differences were very small and concluded that students in these highly sought after charter schools are performing at similar levels to matched students in traditional public schools. They also recognized that urban students performed slightly better, while suburban charter schools perform slightly less well. Unfortunately, CREDO did not demonstrate a command of the literature.

V. Review of the Report's Methods

The methods section of the report is short. It lacked the conventional descriptions used in documenting methods, clearly reporting limitations, and fairly presenting findings. In fact, the report tended to overstate findings that are too small in size to be policy-relevant.

In this study, CREDO uses gain scores to measure student growth over time, and compares charter students' gain scores to those of "Virtual Twins" created by averaging the test scores of similar public school students. Unfortunately, the authors do not highlight limitations; an omission that can undermine the credibility of even very sound analyses. More in-depth criticisms of CREDO's methodology are well-reported in reviews of their previous charter school reports (e.g. Miron & Applegate, 2009; Maul, 2013; Maul & McClelland, 2013).[6] The key criticisms of these reviews are as follows:

- CREDO's technical reporting is insufficient, and inadequately describes the analytic process by which the report's conclusions are reached.
- CREDO's outcome variable of student gain is flawed, presuming that a student exhibiting a certain degree of improvement on their state's test is indistinguishable from a student in another state exhibiting the same degree of improvement on an entirely different test.
- CREDO's Virtual Control Record (VCR) approach to quasi-experimental control deviates from widely-accepted propensity matching methods and is not a legitimate substitute for true random assignment to experimental conditions.[7]
- CREDO's choice to model school-level and student-level effects separately is counter-indicated by the hierarchical structure of the data. Absent statistical adjustments that CREDO neglected to describe, their analysis fails to account for within-school variance.
- CREDO's regression model treats school achievement scores as fixed instead of randomly sampled; this approach is justified if describing the observed population of schools in the sample, but the report suggests inference to the theoretical population of all charter schools.
- CREDO's gain scores, which are a longitudinal measure of student growth, from 2012-13 to 2014-15, represent average instead of individual growth and assumes the growth trends of both the experimental and matched comparison groups are stable.

Since CREDO does not appear to have modified its methodological approach, these technical concerns undermine the validity and utility of the findings.

Of further relevance to this study, CREDO broke out charter schools into four categories based on the existence and nature of management organizations:

1. Schools operated by *charter management organizations* (CMOs). Entities that operate at least three charter schools and are the charter holder for all schools

they operate. These can be for-profit or nonprofit in nature.

2. *Vendor-operated schools* (VOS). Entities that operate at least three charter schools but do not hold the charter. This includes organizations or companies that provide limited services to organizations or companies that fully operate the schools. Some of these are nonprofit organizations and some are for-profit.

3. Schools operated by *Hybrid organizations*. This refers to small subset of entities that operate schools that would be in a CMO relationship and other schools in a VOS relationship. This can also include a network that holds contracts to operate multiple management organizations. It was estimated that 1 percent of charter schools are managed by a Hybrid organization.

4. *Independent charter schools*. This refers to charter schools without management organizations that operate them. These are also referred to as "mom & pop" charter schools and they typically fit the original idea of charter schools being small, locally run, autonomous schools.

While CREDO bases much of their analyses on differences between CMOs, VOS, Hybrid, and Independent schools, these classifications are not conventional in either research or practice. The New Schools Venture Fund, for example, describes CMOs as a small subset of nonprofit education management organizations that are attempting to bring successful models to scale with support from the philanthropic sector.

Neither does CREDO's classification system overlap with the definitions and classifications of other research on charter management organizations, including the definitions established by the National Education Policy Center, which has compiled 14 Editions of Education Management Organization Profiles[8], and the Center for Reinventing Public Education and Mathematica which conducted a study on management organizations.[9] In fact, the classification system used in this report also differs from an earlier CREDO report.[10]

The new VOS category needlessly complicates rather than clarifies the taxonomy. The authors group all schools with relationships with private companies, including those providing limited services and those that have contracts to operate the whole school. While management organizations that operate the school are accountable for outcomes, vendors that provide limited services are not accountable for outcomes. Given this, it makes no sense to combine these diverse vendors since the CREDO analysis is based on outcomes, and many vendors are not responsible for outcomes.

Beyond the concerns about the VOS category, there are many other conceptual and methodological problems with the classifications created by CREDO; some of which we highlight below:

- CREDO excluded more than 60 organizations and companies that operate charter schools during the years covered by their study.

- The CREDO team claims that vendors are management organizations that are under contract to an independent charter school board; however, most of the entities they list as education management organizations started the schools and re-

tain executive control of the schools based on the conditions of the management agreement. (These are what the NEPC group refers to as Education Management Organizations, whereas vendors are not profiled by NEPC since they do not have executive control of the school and they are not responsible for outcomes).

- The CREDO team did not consider the literature on nonprofit organizations and they do not address the substantive differences between a nonprofit education board and a nonprofit education management organization. While the independent schools are seen as nonprofit entities, the other 3 types of organizations can be classified as either for-profit or nonprofit in nature. Both past research and CREDO's findings in this report suggest for-profit and nonprofit management organizations may conduct educational activities in different manners and with different motivations.

- CREDO did not contact the schools or organizations to allow them to confirm or refute the classification the CREDO assigned to them (this is a common practice in the EMO Profiles compiled by the National Education Policy Center).

- CREDO identified and grouped schools as affiliated with CMOs in states that require each charter school to have its own independent governing board, contradicting their own operational definition of CMO.

- The actual number of organizations that operate charter schools identified by CREDO is far lower than the number of organizations identified by the National Education Policy Center, which is responsible for generating regular reports that include inventories of nonprofit and for-profit EMOs.[11]

It is troubling that the CREDO researchers have not updated their taxonomy from relevant literature and have created categories that do not reflect policy-relevant questions. Further, they have effectively eliminated the comparison of their findings to others, thus further undermining the validity and utility of their findings.

VI. Review of the Validity of the Findings and Conclusions

As addressed above, validity-relevant criticisms of CREDO's methodology are well-reported in reviews of their previous charter school reports.[12] However, it is worth emphasizing the inadequacy with which CREDO reports the effects of charter school attendance on student achievement outcomes.

Three issues related to CREDO's handling of effect size warrant further attention. First, CREDO fails to interpret effect sizes in any meaningful way. Second, and related to the first point, CREDO refers to observed student growth in terms of instructional days, a misleading and erratic error-prone comparison. And third, CREDO implicitly conflates statistical significance with policy or practical significance. The experienced reader of the educational research literature will likely notice these issues and adjust their interpretation of the re-

sults. Unfortunately, to journalists and policymakers who may act upon this report, and to parents and students impacted by those actions, CREDO's casual approach to reporting the impact of charter schools presents an inaccurate and misleading picture of charter school performance and the role of management organizations.

Effect Size

Traditional measures of effect size are based generally on rule-of-thumb and approximation. The purpose is to determine whether the difference is big enough to warrant a policy change. The most common benchmarks of effect size refer to an effect greater than 0.2 of a standard deviation as "small", an effect over 0.5 as "medium", and an effect over 0.8 as "large" (Cohen, 1988).[13] By these standards, almost none of CREDO's findings would qualify as even a small effect.[14] Cohen, it seems, never intended for these benchmarks to become the industry standard for effect size interpretation regardless of context; depending on the discipline of study, the practical significance of a given effect size can vary greatly, so the traditional context-independent interpretation, though occasionally a reasonable approximation of practical significance, can also be seen as flawed.[15]

Numerous researchers have suggested, instead, interpreting effect size in terms of the existing research base on the topic of interest, or in terms of the effect sizes that would make an intervention relevant for policy purposes).[16] A report for the Institute of Education Sciences[17] addressed effect size reporting and suggested "four complementary perspectives whereby intervention effects are assessed (a) relative to normal student academic growth, (b) relative to policy-relevant gaps in student performance, (c) relative to the size of the effects found in prior educational interventions, and (d) relative to the costs and benefits of the intervention" (p. 26). A well-reported effect size for CREDO's repeated comparisons between independent charter schools and externally managed charter schools, then, might cast the comparison in terms of the known difference between schools in the 50th and 75th percentiles of achievement, or in terms of the regularly observed gaps between high- and low-SES students. CREDO has incorporated none of these four perspectives in any useful manner.

Instructional Days

The report does offer the differences between types of schools in terms of instructional days gained or lost. According to Hanushek, Peterson, and Woessmann (2012),[18] whom CREDO cites as the source for equating effect size to days of instruction: "student growth is typically about one full standard deviation on standardized tests between 4th and 8th grade, or about 25 percent of a std. dev. from one grade to the next" (p. 4). While this interpretation of effect size is relevant to the paper it originates from, it is misapplied to CREDO's report for several reasons. In the introduction to the report CREDO acknowledges that Hanushek, Peterson, and Woessmann's calculations are based on "examining average growth from fourth grade to eight grade on the National Assessment of Educational Progress (NAEP)" (p. 12). The NAEP is administered to representative samples of 4th and 8th grade students, and is wholly unrelated to the state tests that CREDO used to assess student achievement. Setting aside

issues with CREDO's approach to equating student performance across different state tests (addressed in detail in Miron & Applegate, 2009),[19] it is unreasonable to assume that the growth students display on the NAEP is equivalent to that of any single state test, much less all of them. Furthermore, research suggests wide variation in growth trajectories by year, by level of achievement, and by demographic group; between 4th and 5th grade, a student's academic achievement in math might be expected to increase by 0.56 (± 0.11) of a standard deviation, while between 7th and 8th grade it might be expected to increase by 0.32 (± 0.05) of a standard deviation.[20] While CREDO's Virtual Twin matching approach may account for some differences in growth trajectory at the student level, it does not justify days-of-instruction as an interpretation for effect size (for a critique of the Virtual Twin matching approach, see Maul & McClelland, 2013).[21] Even if the NAEP were psychometrically equivalent to the tests that CREDO aggregates, a cross-sectional measure of growth averaged over four years and scaled by instructional days in a year is likely to wildly overestimate instructional days in some years, making this interpretation of effect size at best imprecise and at worst extremely misleading.

Practical Significance vs. Statistical Significance

To compound the spurious sense of accuracy and precision CREDO promotes through the use of inadequately interpreted effect sizes, the report does nothing to dispel any illusions the reader might have that statistically significant results are in fact practically significant. CREDO writes that "it is important to consider if an effect size is significant or not" (p. 12), but offers no caveat to couch this statement in terms of the context of this study. Considering that statistical significance is in large part a function of sample size, a sample of 3.7 million student observations results in innumerable effects "detected" as statistically different from zero – but practically inconsequential. What is the reader to make of independent charter schools' 0.01 standard deviation difference from traditional public schools on math scores, significant at $p < .01$? CREDO would have done better to express confidence in their results in terms of margins of error around their effect sizes. Upper and lower bounds on the error of estimation are a necessary inclusion in order to inform the reader of the accuracy and reliability of effect size estimates, and provide an unambiguous alternative to the inscrutability of p-values in large samples (Kelly & Preacher, 2012; Sun, Pan, & Wang, 2010).[22] By relying on statistical significance testing without giving the implications of this decision the fair treatment that it deserves, CREDO has added a veneer of academic professionalism while contributing nothing to the reader's interpretation of their results.

The findings in CREDO's report do not offer sufficient education-relevant benchmarks by which to evaluate the magnitude of their observed results

To summarize, the findings in CREDO's report do not offer sufficient education-relevant benchmarks by which to evaluate the magnitude of their observed results. The interpretive framework that CREDO does offer, instructional days, is deeply flawed. And CREDO's reliance on statistical significance testing, absent discussion of the practical significance of their results, creates a false sense of precision and conclusiveness that a single study on a constantly evolving body of research does not warrant. These issues are further compounded by the methodological issues outlined above. Methodologi-

cally speaking, one of the key flaws in the study that can invalidate findings outright is the peculiar and unfounded classification of management organization types. Creating arbitrary classification variables for analysis and then over-interpreting the results of this analysis makes any attempt at inference to the greater population of charter school management organizations highly suspect and perhaps dangerously misleading.

Stanford University has a long tradition of teaching and training in the fields of social science research and statistics. Given the proximity to excellent training opportunities and input from highly qualified scholars, it is surprising that the CREDO team continues to generate similarly flawed reports. The reports could be relevant in the body of scholarship, but as long as the authors of the reports overstate miniscule or non-discernable differences, conflate statistical significance with size of impact, and do not fairly present the limitations and caveats to readers, these reports should be seen as advocacy works, and not contributions to the already large body of scholarship on the topic of charter school performance.

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

In summary, this report dramatically overstates its findings, ignores relevant literature, fails to address known methodological issues, and does not bring the limitations of the study to the fore. Consequently, the general findings from this study do not have relevance to policymakers.

In academia, scholars recognize the relevant literature and build upon it, using past research to contextualize and support their findings. The strength of the charter school effectiveness research base depends not on a single piece of well-executed research but the comprehensive research output of all scholars working in the field. In the case of CREDO, their continued disinterest in the relevant research neutralizes and marginalizes their work. Whatever the reason, it is unfortunate that the researchers at CREDO approached this work as if in an academic vacuum.

One of the great obstacles to research in the charter school area is the excessive politicization of research issues. Thus, when a study overstates its conclusions to the level demonstrated in this effort, it suggests that it is more about serving the needs and agenda of the funding agencies (i.e., the Walton Family Foundation and Fischer Fund) rather than providing sound evidence for policy. This is a disservice to advancing our knowledge as well as to the common good.

The tables that display effect size estimates for individual management organizations in Appendix A are, however, relevant since these are reported without grouping the organizations by the typology that CREDO developed for this report. This is far from being a definitive study, as the authors might wish for readers to believe, although it can be seen as a useful contribution to the already extensive field of charter school research provided the reader approaches CREDO's work with a critical eye. We remain hopeful that the concerns presented here and in other reviews will be incorporated into CREDO's future work on charter schools;

this would help ensure that the findings they present are more accurate and fairly presented, and more useful for driving sound policy.

Notes and References

1 Woodworth, J., Raymond, M., Han, C., Negassi, Y., Richardson, P., & Snow, W. (2017). *Charter Management Organizations 2017*. Palo Alto, CA: Center for Research on Education Outcomes. Retrieved July 5, 2017, from https://credo.stanford.edu/pdfs/CMO%20FINAL.pdf

2 Miron, G. & Applegate, B. (2000). *An evaluation of student achievement in Edison schools opened in 1995 and 1996*. Kalamazoo, MI: Western Michigan University. Retrieved August 5, 2017, from http://homepages.wmich.edu/~miron/publics/wmu_edison_rpt.pdf

Saltman, K. (2005). *The Edison Schools: Corporate Schooling and the Assault on Public Education*. Abingdon, UK: Taylor & Francis Ltm.

3 Loveless T. (2003). *The 2003 Brown Center report on American Education. How well Are American Students Learning?* Washington, DC: Brookings Institution.

Miron, G. & Nelson, C. (2002). *What's public about charter schools: Lessons learned about choice and accountability*. Thousand Oaks, CA: Corwin Press.

Mathematica & CRPE. (2012). *The National Study of Charter Management Organization (CMO) Effectiveness; Charter Management Organizations: Diverse Strategies and Diverse Student Impacts*. Mathematica Policy Research Group, Washington, D.C. & Center on Reinventing Public Education, University of Washington. Retrieved August 4, 2017, from http://www.crpe.org/sites/default/files/pub_cmofinal_Jan12_0.pdf

4 Miron, G. & Applegate, B. (2009). *Review of Multiple Choice: Charter school performance in 16 states*. Boulder and Tempe, CO: Education and the Public Interest Center & Education Policy Research Unit.

5 Gleason, P., Clark, M., Tuttle, C.C., & Dwoyer, E. (2010). *The Evaluation of Charter School Impacts: Final Report (NCEE 2010-4029)*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Gill, B., Walsh, L., Smither Wulsin, C., Matulewicz, H. Severn, V., Grau, E., Lee, A., & Kewin, T. (2015). *Inside Online Charter Schools*. Cambridge, MA: Mathematica Policy Research. Retrieved November 12, 2015, from http://www.mathematica-mpr.com/~/media/publications/pdfs/education/inside_online_ charter_schools.pdf

6 Miron, G. & Applegate, B. (2009). *Review of "Multiple choice: Charter school performance in 16 states."* Boulder and Tempe, CO: Education and the Public Interest Center & Education Policy Research Unit.

Maul, A. & McClelland, A. (2013). *Review of "National charter school study 2013."* Boulder, CO: National Education Policy Center.

Maul, A. (2013). *Review of "Charter school performance in Michigan."* Boulder, CO: National Education Policy Center.

7 Maul, A. & McClelland, A. (2013). *Review of "National charter school study 2013."* Boulder, CO: National Education Policy Center.

8 These reports can all be downloaded from the following website: http://nepc.colorado.edu/ceru/annual-report-education-management-organizations

9 Mathematica & CRPE. (2012). *The National Study of Charter Management Organization (CMO) Effectiveness; Charter Management Organizations: Diverse Strategies and Diverse Student Impacts*. Mathematica Policy Research Group, Washington, D.C. & Center on Reinventing Public Education, University of Washington. Retrieved August 4, 2017, from

http://www.crpe.org/sites/default/files/pub_cmofinal_Jan12_0.pdf

10 Woodworth, J., & Raymond, M. (2012). *Charter School Growth and Replication: Volume II*. Palo Alto, CA: Center for Research on Education Outcomes. Retrieved July 5, 2017, from https://credo.stanford.edu/pdfs/CGAR%20Growth%20Volume%20II.pdf

11 The National Education Policy Center, which has already put out 14 Editions of a comprehensive inventory of all EMOs and the schools they operate. The group is currently working on the 15th Edition of EMO Profiles, which is expected to be published in December 2017. See http://nepc.colorado.edu/ceru/annual-report-education-management-organizations

12 Miron, G. & Applegate, B. (2009). *Review of "Multiple choice: Charter school performance in 16 states."* Boulder and Tempe, CO: Education and the Public Interest Center & Education Policy Research Unit.

Maul, A. & McClelland, A. (2013). *Review of "National charter school study 2013."* Boulder, CO: National Education Policy Center.

Maul, A. (2013). *Review of "Charter school performance in Michigan."* Boulder, CO: National Education Policy Center.

13 Cohen, J. (1988). *Statistical power analysis for the behavior sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

14 To illustrate the application of CREDO's equating of effect sizes to days of instruction, consider their 2013 report (Cremata et al., 2013) in which they found and interpreted effect sizes between charter school students and district school students that ranged from 0.007 to 0.01. These are incredibly small, even though they were statistically significant (meaning if they repeated the study with the same measures and methods, 95 out of 100 times, the researchers would find incredibly small differences between the two groups. In the current report, CREDO found that CMO-operated charter schools had a positive effect size of 0.03, which is still incredibly small given Cohen's guidance that a small effect size should be at least 0.2. In their 2015 report (Woodworth et al., 2015) on online charter schools, CREDO was finally able to find an effect size that finally met the threshold of "small." The effect size was -0.25 in math and CREDO equated this to 180 days of instruction (i.e., over the course of one year, the students in virtual charter schools lost one full year of instruction relative to matched students in traditional public schools). This is telling since when CREDO finally captures an effect size that is sufficiently large to be classified as "small" it turns out that this is equated to 180 days of instruction; this illustrates that CREDO's equating of effect size to days of instruction is absurd and nonsensical.

Cremata, E., Davis, D., Dickey, K., Lawyer, K., Negassi, Y., Raymond, M., & Woodworth, J. (2013). *National Charter School Study*. Palo Alto, CA: Center for Research on Education Outcomes. Retrieved August 5, 2017, from http://credo.stanford.edu/documents/NCSS%202013%20Final%20Draft.pdf

Woodworth, J., Raymond, M., Chirbas, K., Gonzalez, M., Negassi, Y., Snow, W., & Van Donge, C. (2015). *Online Charter School Study, 2015*. Palo Alto, CA: Center for Research on Education Outcomes, Retrieved August 5, 2017, from http://credo.stanford.edu/pdfs/Online%20Charter%20Study%20Final.pdf

15 Sun, S., Pan, W., & Wang, L.L. (2010). A comprehensive review of effect size reporting and interpreting practices in academic journal in education and psychology. *Journal of Educational Psychology*, *102*(4), 989-1004.

16 Hill, C. J., Bloom, H.S., Black, A.R., & Lipsey, M.W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives*, *2*(3), 172-177.

17 Lipsey, M.W., Puzio, K., Yun, C., Hebert, M.A., Steinka-Fry, K., Cole, M.W., Roberts, M., Anthony, K.S., & Busick, M.D. (2012). *Translating the Statistical Representation of the Effects of Education Interventions into More Readily Interpretable Forms* (Report No. NCSER 2013-3000). Washington, D.C.: Institute of Education Sciences at the U.S. Department of Education.

18 Hanushek, E.A., Peterson, P.E., & Woessmann, L. (2012). *Achievement Growth: International and U.S. State*

Trends in Student Performance (Report No. 12-03). Cambridge, MA: Harvard Kennedy School.

19 Miron, G. & Applegate, B. (2009). *Review of "Multiple choice: Charter school performance in 16 states."* Boulder and Tempe, CO: Education and the Public Interest Center & Education Policy Research Unit.

20 Hill, C.J., Bloom, H.S., Black, A.R., & Lipsey, M.W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives*, *2*(3), 172-177.

Hanushek, E.A., Peterson, P.E., & Woessmann, L. (2012). *Achievement Growth: International and U.S. State Trends in Student Performance* (Report No. 12-03). Cambridge, MA: Harvard Kennedy School.

Scammacca, N.K., Fall, A-M., & Roberts, G. (2015). Benchmarks for expected annual academic growth for students in the bottom quartile of the normative distribution. *Journal of Research on Educational Effectiveness*, *8*(3), 366-379.

21 Maul, A. & McClelland, A. (2013). *Review of "National charter school study 2013."* Boulder, CO: National Education Policy Center.

22 Sun, S., Pan, W., & Wang, L. L. (2010). A comprehensive review of effect size reporting and interpreting practices in academic journal in education and psychology. *Journal of Educational Psychology*, *102*(4), 989-1004.

Kelly, K., & Preacher, K. J. (2012). On effect size. *Psychological Methods*, *17*(2), 137-152.