A Meta-Analysis of the Effect of Charter Schools on Student Achievement

Julian R. Betts, Department of Economics, University of California, San Diego, and National Bureau of Economic Research

Y. Emily Tang, Department of Economics, University of California, San Diego

San Diego Education Research Alliance at UC San Diego (SanDERA)

sandera.ucsd.edu

Discussion Paper 2018-1

Acknowledgments

An abbreviated version of this paper is forthcoming in **School Choice at the Crossroads: Research Perspectives,** Berends, Mark, R. Joseph Waddington and John Schoenig (Eds.), Routledge.

We would like to thank the many researchers who provided supplementary information we needed to incorporate their papers into this or our earlier (2014, 2011 and 2008) literature reviews. Dale Ballou; Richard Buddin, Leesa M. Foreman, Kaitlin Anderson, Gary Ritter, and Patrick Wolf; Caroline Hoxby and Jonah Rockoff; Anna Nicotera and Mark Berends; Macke Raymond; Christopher Reicher and Larry McClure; Tim Gronberg; and Scott Imberman all provided key data we needed for one or more of our analyses.

Abstract

A meta-analysis is performed of the literature on charter schools and achievement, with a focus on lottery-based studies and rigorous value-added studies. On average, for the limited set of charter schools, locations, and years that have been studied to date, charter schools are producing higher achievement gains in math relative to traditional public schools in elementary and middle but not high schools. For reading achievement charter schools on average are producing higher gains in middle schools but not in elementary or high schools. For both math and reading, middle school studies tend to produce the highest effect sizes of all of the grade groupings. The literature shows a large variation in estimated charter school effects across locations, and some studies also show large variations within a given city or state. We examine some of the factors related to these differences.

1. Introduction

Charter schools are public schools which receive public funds but which receive exemptions from parts of states' education codes, ostensibly to allow them to hire teachers differently and to experiment with alternative curricula and pedagogical approaches. In return for this relative independence, charter schools face greater risk of being shut down than traditional public schools. Theoretically, every few years when a charter school must apply to the chartering authority for a renewal, that authority can turn down the renewal.

Charter schools represent one of the most important innovations in school management in the United States. The National Alliance for Public Charter Schools (2017) estimates that in the 2016-17 school year a total of 6900 charter schools, enrolling 3.1 million students, were operating in 42 states as well as the District of Columbia. This represents almost a tripling in enrollment levels over ten years.

Given that charter schools are intended to improve student learning, in this review article, we study the relation between attending a charter school and scores on state-mandated tests in mathematics and reading. The paper updates and extends prior papers we have written (Betts and Tang 2008, 2011, 2014). Our earlier studies found that results varied by subject tested and grade span, but on average achievement in math and reading of students at charter schools was typically the same or higher than that of comparison groups of students at traditional public schools. We use a method known as meta-analysis to estimate the overall relation between attending charter schools and achievement, and also, crucially, to examine the variation across studies and locations.

Following our earlier papers, we focus on charter school studies that adopt one of two methods. The first approach involves comparing students who win and lose lotteries to attend charter schools. The second approach, known as value-added modeling, is not experimental, but takes into account a student's past academic achievement, unlike some of the weaker nonexperimental approaches.[1]

Only nine papers have used the lottery approach to date, covering a total of 142 charter schools. Non-lottery-based studies that take the value-added approach while also constructing a comparison group against which to benchmark the academic progress of charter school students are far more abundant. In total, the present report includes in its analysis 38 value-added papers that use lottery-based or rigorous value-added approaches, although many of these include estimates for multiple locations or multiple types of charter schools within a given location. This is an increase from 29 papers in our 2014 review. In addition, we updated estimates for five of

1. Betts, Tang, and Zau (2010) use data from San Diego and show that models that do not measure individual students' achievement growth produce quite different results from the more sophisticated value-added models, and that the changes in estimated effects of charters are consistent with the idea that the weaker approaches fail to take into account the relatively disadvantaged backgrounds of students who attend charters.

the studies we cited in 2014, using slightly changed estimates from newly published versions. In our main table that examines estimates for all students in charter schools, across six different grade configurations, we have 182 estimates for reading and 190 for math, up from 120 and 125 in our 2014 analysis. The number of estimates exceeds the number of papers because most studies include estimates for sub-populations of their study area. Some studies, for example, make three contributions by presenting separate estimates for elementary, middle, and high schools. Others provide estimates for various types of charter schools in a given location, and one paper studying Los Angeles distinguishes between effects for students who remain at newly converted charter schools and those who switch into charter schools.

2. Data and Methods

Our study includes all studies of the relation between charter schools and achievement included in Betts and Tang (2014) plus those found in a Google Scholar search for more recent papers. To be eligible for inclusion, a study needed to observe achievement at the student level, and to control for past student achievement. Further, the estimated effect, its standard error, and the standard deviation of test scores must be presented to be eligible for inclusion in a meta-analysis.[2]

Meta-analysis is a widely used method, especially in the medical literature, of combining estimates from a variety of studies to draw overall conclusions. We use this approach to test whether the average impact of charter schools on achievement is zero, to portray visually the variation across studies, and to estimate the degree to which this is real variation in the effectiveness of charter schools as opposed to statistical noise. Details appear in the appendix.[3]

Table 7 shows the set of papers that we included, along with information on the geographic location, the grade- and time-span of the study, and the label we use to identify the study in graphs.

The charter schools studies group schools into a variety of different grade configurations. Accordingly, we characterize charter school outcomes for elementary, middle, and high school levels, and various combinations. The most common of these is studies that combine elementary

[2] In past reviews, we have noted that not all authors follow the norm of including both an estimated effect and either a standard error or t-statistic. This is problematic because without the latter we lack a measure of how precise the estimated effect is, which is key to calculating the study's weight in the meta-analysis. Among new papers since our 2014 work, the main example of papers that lacked standard errors were ten studies by CREDO of New York, New York City, Texas, Ohio, California, Los Angeles and an urban study. We requested the standard errors for these studies, but were not able to obtain them.

[3] We do not replicate section 3 of Betts and Tang (2011), which tested whether one could maintain the hypothesis of no negative effects of charter schools in the literature and, conversely, the hypothesis of no positive effects. That past analysis showed that for all but two combinations of grades studied and subjects tested, there is very strong evidence of both positive and negative effects in the literature.

and middle charter schools into a single analysis. We refer to these studies as “combined elementary/middle” studies. Another popular approach has been to combine elementary, middle, and high schools into a single analysis, which we refer to as an “all” grade span.

There are a number of locations in which multiple charter school effectiveness estimates exist because different authors have studied the same place. In the appendix, table 7 lists these studies.

We note that the literature on charter schools and achievement continues to grow quickly. The nature of the research is evolving as well. More studies have started to compare outcomes for different groups of students and different types of charter schools. This trend is laudable as it will help us to pinpoint what types of programs, and for which students, charters contribute to academic success. The only downside to this trend is that some papers no longer produce an overall estimate of the effects of charter schools in an area, which makes it harder for readers to get an overall sense of success of the charter schools in a given area.

3. Mean Effects and Variations across Studies

We began by obtaining estimates of charter school effects for the main grade spans found often in the literature.

As in our previous studies (Betts and Tang 2011 and 2014), our main results in this section, in table 1, exclude the results for KIPP charter schools from both the middle school results and the results that combine elementary studies, elementary/middle studies, and middle school studies. (KIPP refers to the Knowledge is Power Program, a charter school operator. The KIPP estimates are often much larger than the estimates in studies that include all charter schools in a given region, and they would assume a disproportionate weight if we included them in the main analysis.) We later perform a meta-analysis of the KIPP studies themselves.

Tables 1 and 2 show the main results. Table 1 shows the results in terms of “effect sizes,” that is, the predicted change in a student’s achievement measured in terms of the number of standard deviations of achievement. Although this is the normal way of presenting results in education research, many readers may find it more understandable to read the results in terms of predicted changes in percentile rank for a student attending a charter school. Table 2 shows the results transformed into percentile rankings.[4]

In table 1, results for each grade span for reading and math appear in the first and second columns respectively. For each grade span, the first number shows the estimated overall effect

4. The percentile ranking of a student indicates the number of students out of 100 that the student would score as highly as or higher than. For example, a 99th percentile student scores as highly as or higher than 99 out of 100 students on average, while a 50th percentile student is in the middle of the achievement distribution.

size. Effect sizes that are statistically significant (at the 5 percent level) are indicated with an asterisk. For elementary schools, the overall estimated effect sizes for reading and math achievement are 0.018 and 0.033, although only the latter estimate is significant at the 5 percent level. The corresponding estimates from Betts and Tang (2014) were 0.020 and 0.045 respectively, and both were significant.

The second number for each grade span shows, in parentheses, the number of estimates contributing to the overall estimate, followed by the number of regions examined in the given studies. For example, in the meta-analysis of reading effects for elementary schools, "(20-16)" indicates that we found and used 20 separate estimates from 16 geographic areas in calculating the overall effect.

The third number presented for each grade span shows an estimate of the percentage of the variation across estimates that reflects true variation in the impact of charter schools, as opposed to variation due to random noise. (This is the I^2 statistic introduced by Higgins et al. (2003).) For both reading and math studies at the elementary level, we estimate that 99.0 percent of the variation reflects true variations in impact across studies. Clearly, there appear to be important variations in charter school effects across studies and, implicitly, across geographic areas.

For elementary schools, the average effect of charter schools is not statistically significant for reading, but is positive and significant for math. For middle schools, we find positive and significant effects of charter schools for both reading and math achievement. For high school studies, neither effect is statistically significant.

A number of studies combine elementary and middle schools and, as shown in the fourth row of table 1, on average there is no significant effect of attending a charter school on reading or math achievement in these studies. It is somewhat unusual to combine elementary and middle schools in this way but perhaps a practical convenience due to the fact that the exact grade in which elementary school students transition to middle school differs by place. In a bid to find a representative portrait of the overall evidence on the impact of charter schools from studies of schools at the elementary, middle, and combined elementary/middle levels, the fifth row of table 1 combines all three of these study approaches. When pooling studies in this way, we find a statistically significant positive average estimated effect size for attending a charter school in these studies for both reading and math.

Finally, some studies include test scores from elementary, middle, and high school grades together in one model. We refer to these as "All" grade span models. The sixth row of table 1 shows that neither the mean effect size in reading nor in math is significant in these studies.

Table 1. Effect Sizes and Significance From Meta-Analysis, by Grade Span and Subject Area

Grade Span	Reading Tests (# estimates-# locations), % true variation	Math Tests (# estimates-# locations), % true variation
Elementary	0.018 (20-16), 99.0%	0.033* (21-17), 99.0%
Middle	0.054* (23-17), 99.1%	0.097* (24-17), 99.4%
High	0.038 (21-15), 98.4%	0.042 (20-15), 99.2%
Combined Elementary/Middle	0.000 (23-19), 98.6%	-0.018 (28-19), 99.7%
Elementary, Middle, and Combined Elementary/Middle	0.020* (70-28) 98.7%	0.033* (72-29),-99.3%
All	0.012 (25-18),-98.7%	0.023 (25-19),-99.9%

Note: Asterisks indicate effect size significantly different from zero at the 5 percent level or less. The numbers in parentheses indicate the number of estimates included in the associated estimate of effect size and the number of locations. The percentage refers to the I^2 estimate of the percentage of the variation across estimates that reflect true variation in the effect of charter schools, rather than just statistical noise. Thus, for example, in the reading test result for elementary schools "(20-16), 99.0%" indicates 20 estimates covering 16 locations (with two studies each of Los Angeles and San Diego schools, and three studies of New York City area schools), and that 99.0 percent of the variation across estimates in the literature may reflect true variation in the effect of charter schools. As mentioned in the text, we exclude a large number of studies of KIPP schools from the middle school tabulations as the number of studies greatly outweighs the share of these schools in the charter school population, while the effect sizes are also much larger than the average seen in other studies.

In sum, no significant mean effect size emerges in studies of combined elementary/middle school levels, high school levels, or all grade levels combined. In studies of middle schools only, and combining studies of elementary, middle, and elementary/middle combined levels, a positive and significant effect for both reading and math emerges. Studies at the elementary level have mean positive effects for both math and reading, but only those for math are significant.

How do the estimated effects compare to the estimated impact of other common educational interventions? Focusing on the statistically significant results, the effect sizes for math range between 0.033 and 0.097 among the elementary, middle, and high school grade spans, signifying that after one year of attending charter school a student's test score would increase, relative to those of other students, by 3 to 10 percent of one standard deviation. For the two cases in which the charter effect is significant for reading, the predicted gain is 2 to 5 percent of one standard deviation. Clotfelter, Ladd, and Vigdor (2007) estimate that in North Carolina reducing class size by five students is associated with gains in achievement of 1.0 to 1.5 percent of a standard deviation.

Another way of gauging the size of the charter school effect sizes is to translate them into how a charter school student's academic ranking is predicted to change over time. Table 2 translates the effect sizes in table 1 into a student's predicted percentile after attending a charter school for one year. Table 2a makes the assumption that the student starts at the 50th percentile. Table 2b assumes the student starts at the 25th percentile. For the latter case the predicted movement in percentile points is very slightly smaller.

With the important exception of middle schools, the predicted gains in achievement from attending a charter school for one year are small, typically 0.5 to 1 percentile point. These are not huge effects, but if a student stayed in charter schools for 4, 6 or even 12 years cumulatively the predicted effects would be meaningful.

The predicted effects are much bigger for charter schools operating at the middle school level, for both math and reading achievement, and these predicted changes are statistically significant. For reading, depending on whether a student started at the 25th or 50th percentile, a single year in a charter school is predicted to boost a student by 1.7 to 2.2 percentile points. For middle school math, a student is predicted to gain between 3.2 to 3.9 percentile points in a single year depending on the starting point. Over several years, a student's rank is predicted to move up considerably.

Differences in Effect Sizes Across Studies

It is useful to look at the effect sizes of individual studies and how they contribute to the overall estimates shown in table 1. In all cases, almost 100 percent of the variation across studies appears to be true variation, and not just a lack of precision in the estimates. This is as important

Table 2a. Effect Sizes Expressed as Charter Students' Predicted Percentile After One Year, Starting at 50th Percentile, by Grade Span and Subject Area

Grade Span	Reading Tests	Math Tests
Elementary	50.7	51.3*
Middle	52.2*	53.9*
High	51.5	51.7
Combined Elementary/Middle	50.0	49.3
Elementary, Middle, and Combined Elementary/Middle	50.8*	51.3*
All	50.5	50.9

Note: Asterisks indicate effect size significantly different from zero at the 5 percent level or less. The numbers show the predicted test score percentile of a student who started at the 50th percentile, after one year of charter school attendance.

Table 2b. Effect Sizes Expressed as Charter Students' Predicted Percentile After One Year, Starting at 25th Percentile, by Grade Span and Subject Area

Grade Span	Reading Tests	Math Tests
Elementary	25.6	26.1*
Middle	26.7*	28.2*
High	26.2	26.4
Combined Elementary/Middle	25.0	24.4
Elementary, Middle, and Combined Elementary/Middle	25.6*	26.1*
All	25.4	25.7

Note: Asterisks indicate effect size significantly different from zero at the 5 percent level or less. The numbers show the predicted test score percentile of a student who started at the 25th percentile, after one year of charter school attendance.

a finding as the estimates of the average effects. The variation probably reflects both variation in the effectiveness of different charter schools across studies but also in the effectiveness of the traditional public schools against which the research studies compare the charter schools.

Figures 1 and 2 provide an illustration of the variation in the effect sizes across studies of elementary schools for reading and math respectively. The figures use horizontal lines to indicate the 95 percent confidence interval for each estimate. The rightmost column shows the weight attributed to each study. (The size of each square is proportional to these weights.) The diamond at the bottom of each figure illustrates the overall estimated effect size, with the width of the diamond indicating the 95 percent confidence interval.

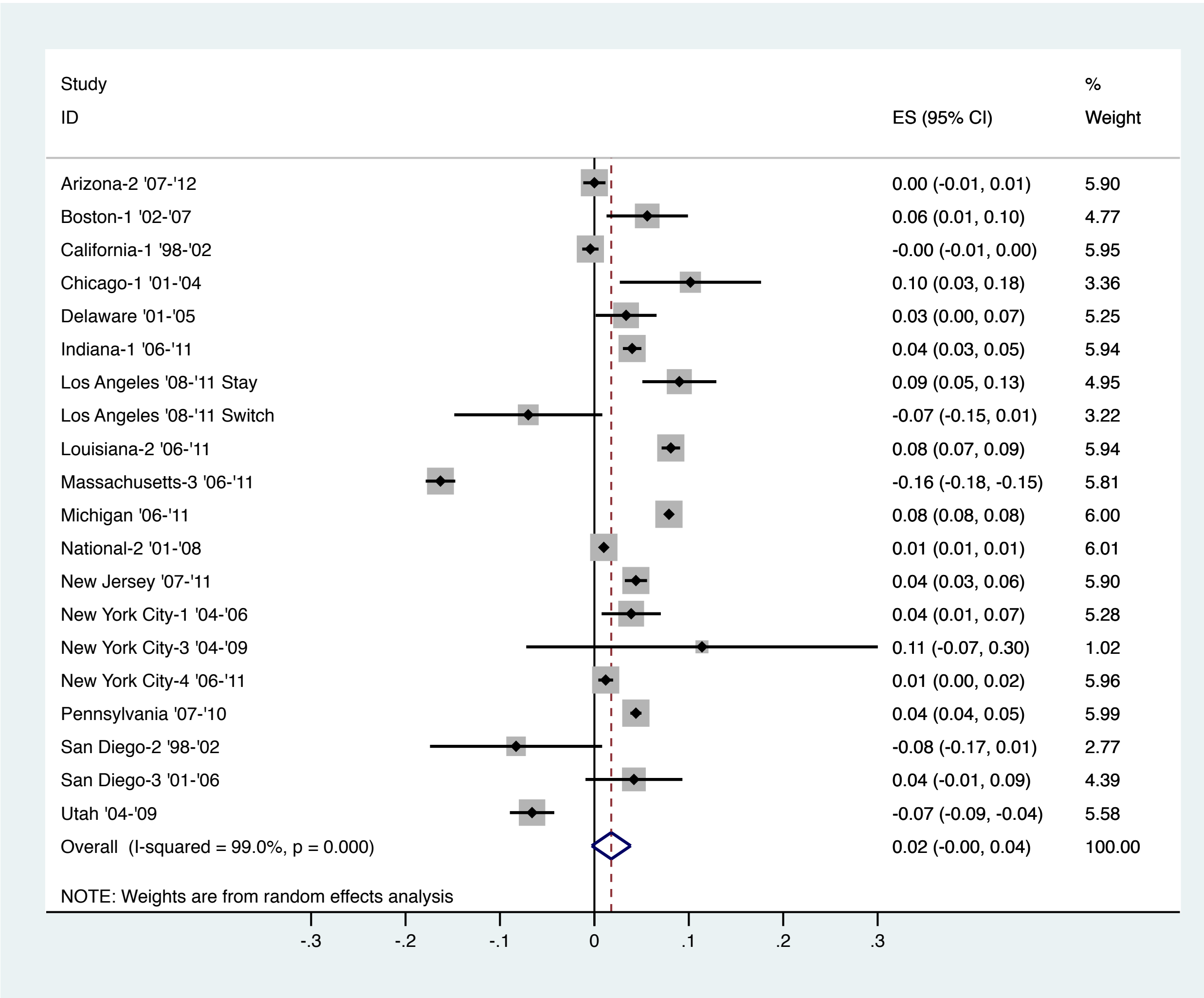
The elementary school studies with the largest estimated reading effect size include studies of New York City, Boston, Los Angeles, Michigan, Louisiana, and Chicago. Two studies show negative and significant results: a study by Ni and Rorrer of Utah (2012), and a CREDO (2013a) study of Massachusetts. A third study with a large negative (but in this case not quite significant) coefficient is a study of San Diego charters (Betts et al. 2005). A study of San Diego by Betts, Tang, and Zau (2010) using the same statistical approach but a later time frame produced a positive and, again, nearly significant coefficient. In math, the studies with the largest positive effect sizes for elementary charter schools were in Idaho, San Diego, New York City, and Chicago. (Again, a study of an earlier period in San Diego produced a negative and this time significant counterpoint. It seems likely that San Diego's charter schools have become more effective with regard to math and reading achievement over time.)

The bottom of the left-hand column in the figures reproduces the I^2 statistic along with the p value of a test for homogeneous effects across studies. The p values are essentially zero, which is what we typically found in our analyses of other samples. Put simply, this result indicates that the idea that charter schools have the same impact in all geographic areas is wrong.

The right-hand column in the figures shows the weights assigned to each study when obtaining our overall estimated effect size. Smaller, less precise estimates get less weight than larger, more precise estimates, but because most of the variation is estimated to be "true," for the most part there is not much difference in the weight assigned to the various studies.

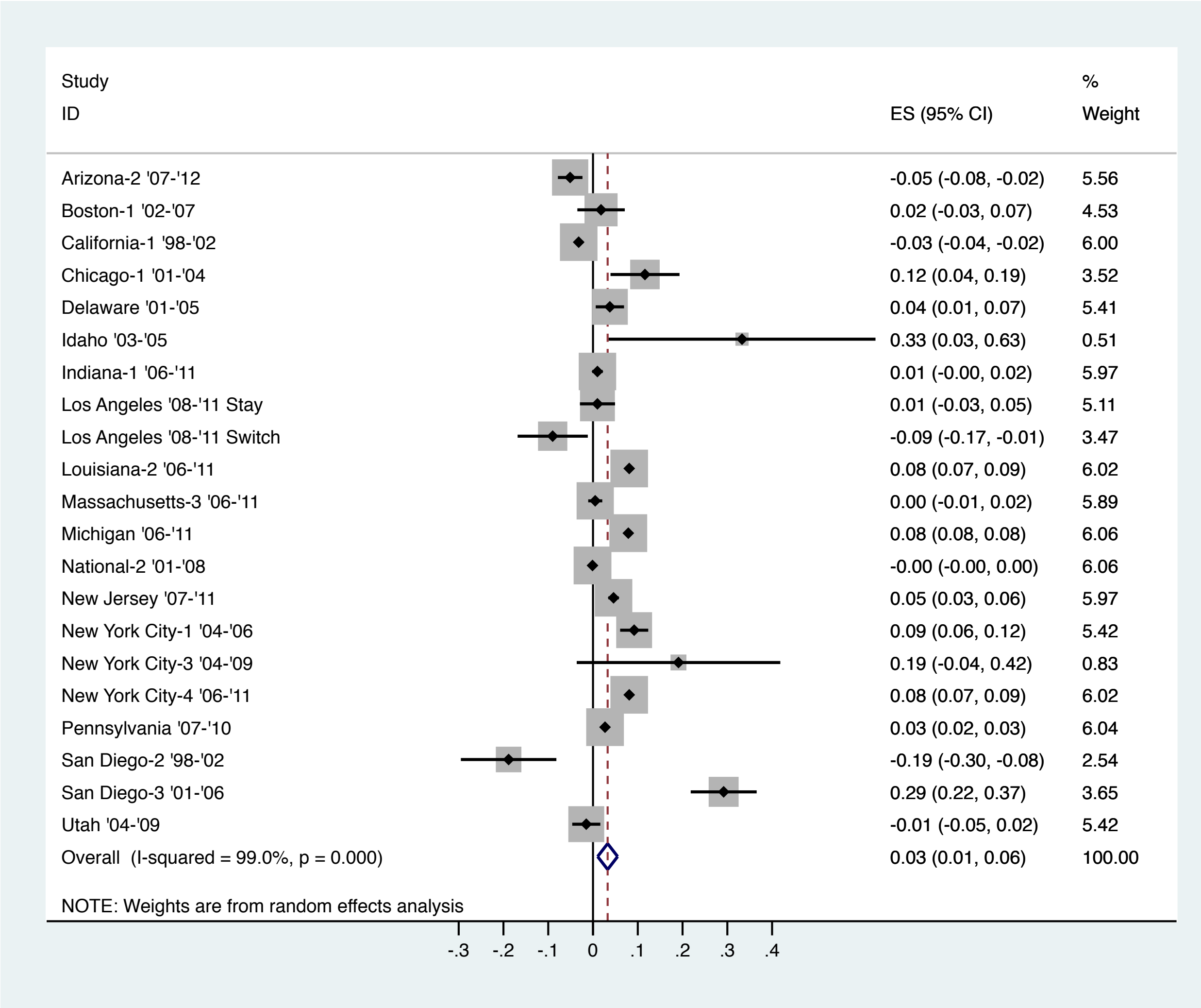
Figures 3 and 4 show the estimated effects in middle school studies for reading and math respectively. For reading, estimates lie in a fairly narrow band centered at just above zero, with roughly two-thirds of estimates being positive. Positive results from Los Angeles (for students staying in newly converted charter schools), Boston, and Massachusetts exhibit the largest effect size in these studies. Figure 4 shows that most studies of math achievement produced positive effect sizes, often statistically significant. Again, the biggest outlier is the result from Boston, with a positive effect size about double the size of the next biggest estimate (from New York City).

Figure 1. Elementary School Reading Effect Sizes by Study, Showing Weights Ascribed by Random-Effects Meta-Analysis to Each Study



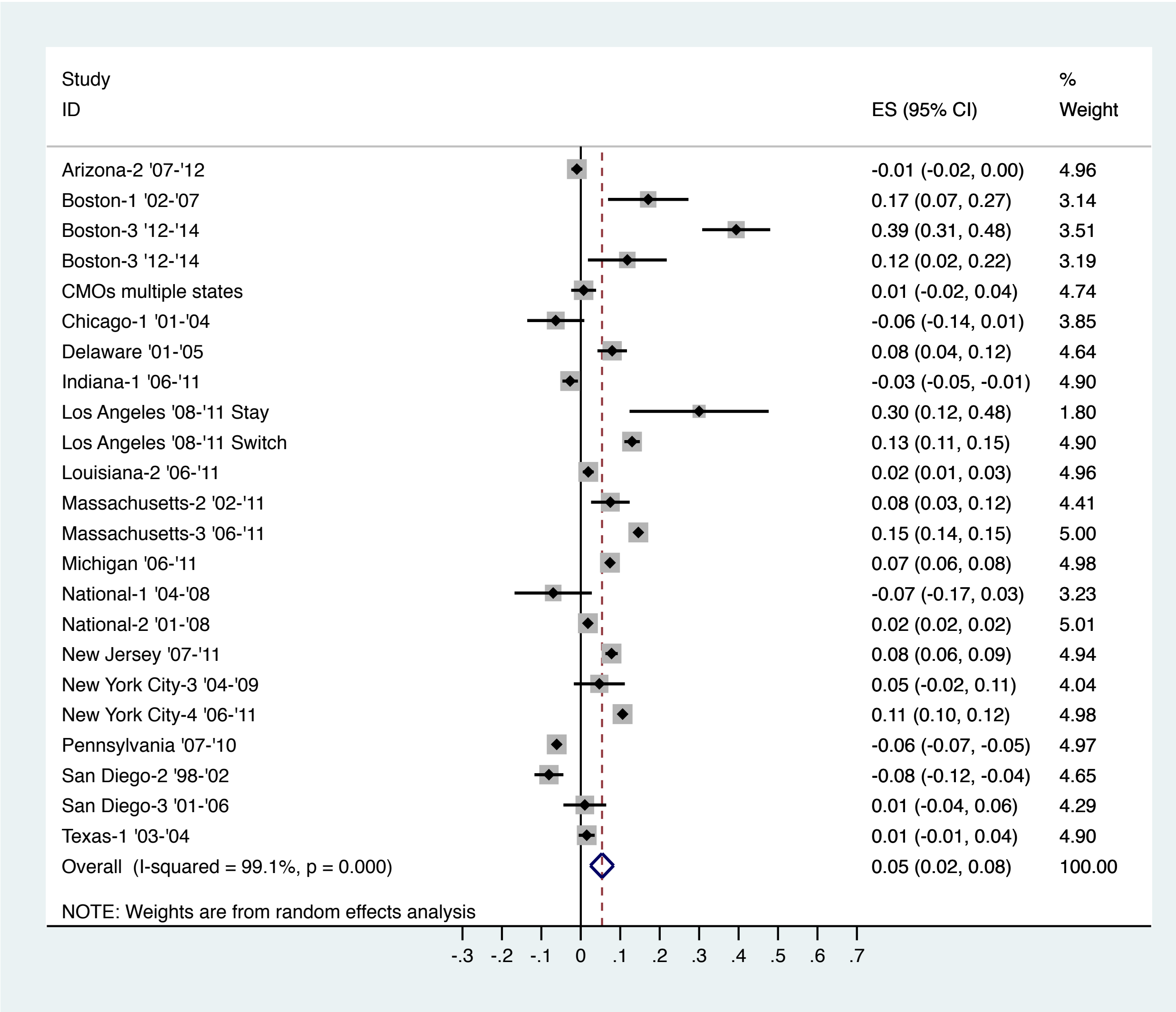
Notes: Weights are from random-effects analysis. The horizontal lines show the 95 percent confidence interval, which is also indicated in the second column from the right. The rightmost column shows the weight ascribed to each study, with the size of the square proportional to these weights. The overall effect size estimate is shown at the bottom. Geographic locations with estimates from multiple studies have unique numbers appended to their labels to distinguish between studies. Table 7, in the appendix, indicates the author and year of the study referenced by each study ID label.

Figure 2. Elementary School Math Effect Sizes by Study, Showing Weights Ascribed by Random-Effects Meta-Analysis to Each Study



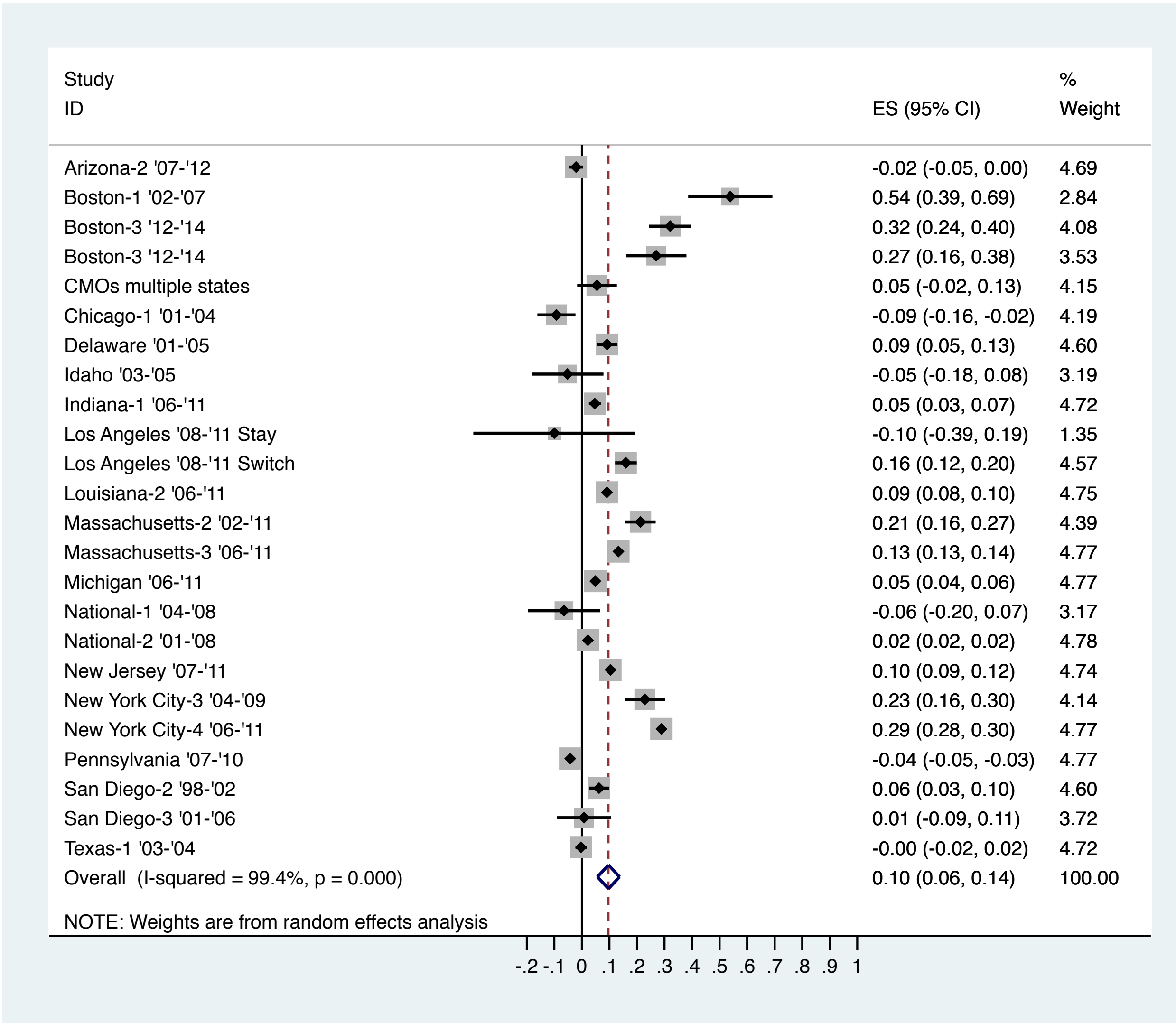
Notes: Weights are from random-effects analysis. The horizontal lines show the 95 percent confidence interval, which is also indicated in the second column from the right. The rightmost column shows the weight ascribed to each study, with the size of the square proportional to these weights. The overall effect size estimate is shown at the bottom. Geographic locations with estimates from multiple studies have unique numbers appended to their labels to distinguish between studies. Table 7, in the appendix, indicates the author and year of the study referenced by each study ID label.

Figure 3. Middle School Reading Effect Sizes by Study, Showing Weights Ascribed by Random-Effects Meta-Analysis to Each Study



Notes: Weights are from random-effects analysis. The horizontal lines show the 95 percent confidence interval, which is also indicated in the second column from the right. The rightmost column shows the weight ascribed to each study, with the size of the square proportional to these weights. The overall effect size estimate is shown at the bottom. Geographic locations with estimates from multiple studies have unique numbers appended to their labels to distinguish between studies. Table 7, in the appendix, indicates the author and year of the study referenced by each study ID label.

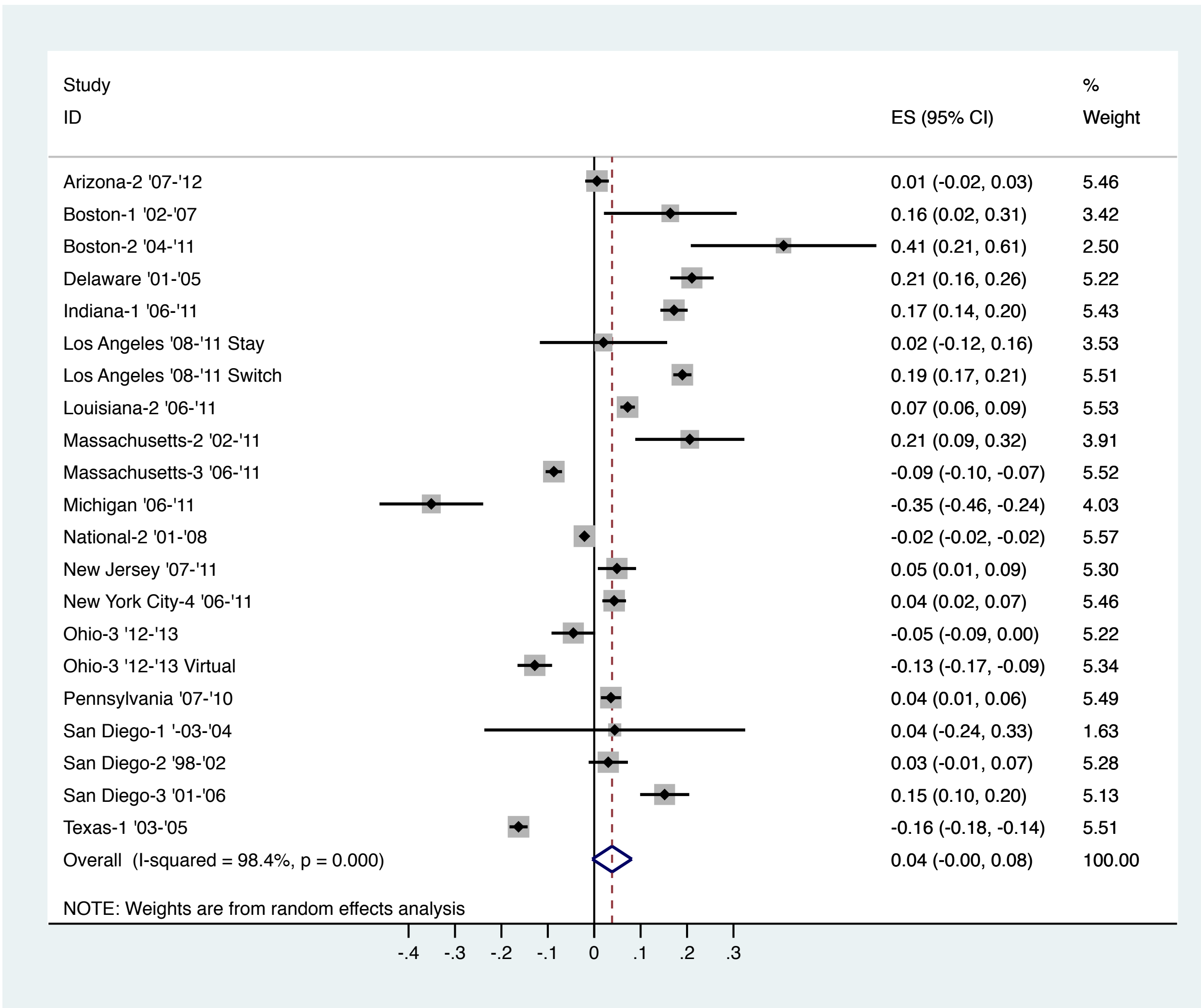
Figure 4. Middle School Math Effect Sizes by Study, Showing Weights Ascribed by Random-Effects Meta-Analysis to Each Study



Notes: Weights are from random-effects analysis. The horizontal lines show the 95 percent confidence interval, which is also indicated in the second column from the right. The rightmost column shows the weight ascribed to each study, with the size of the square proportional to these weights. The overall effect size estimate is shown at the bottom. Geographic locations with estimates from multiple studies have unique numbers appended to their labels to distinguish between studies. Table 7, in the appendix, indicates the author and year of the study referenced by each study ID label.

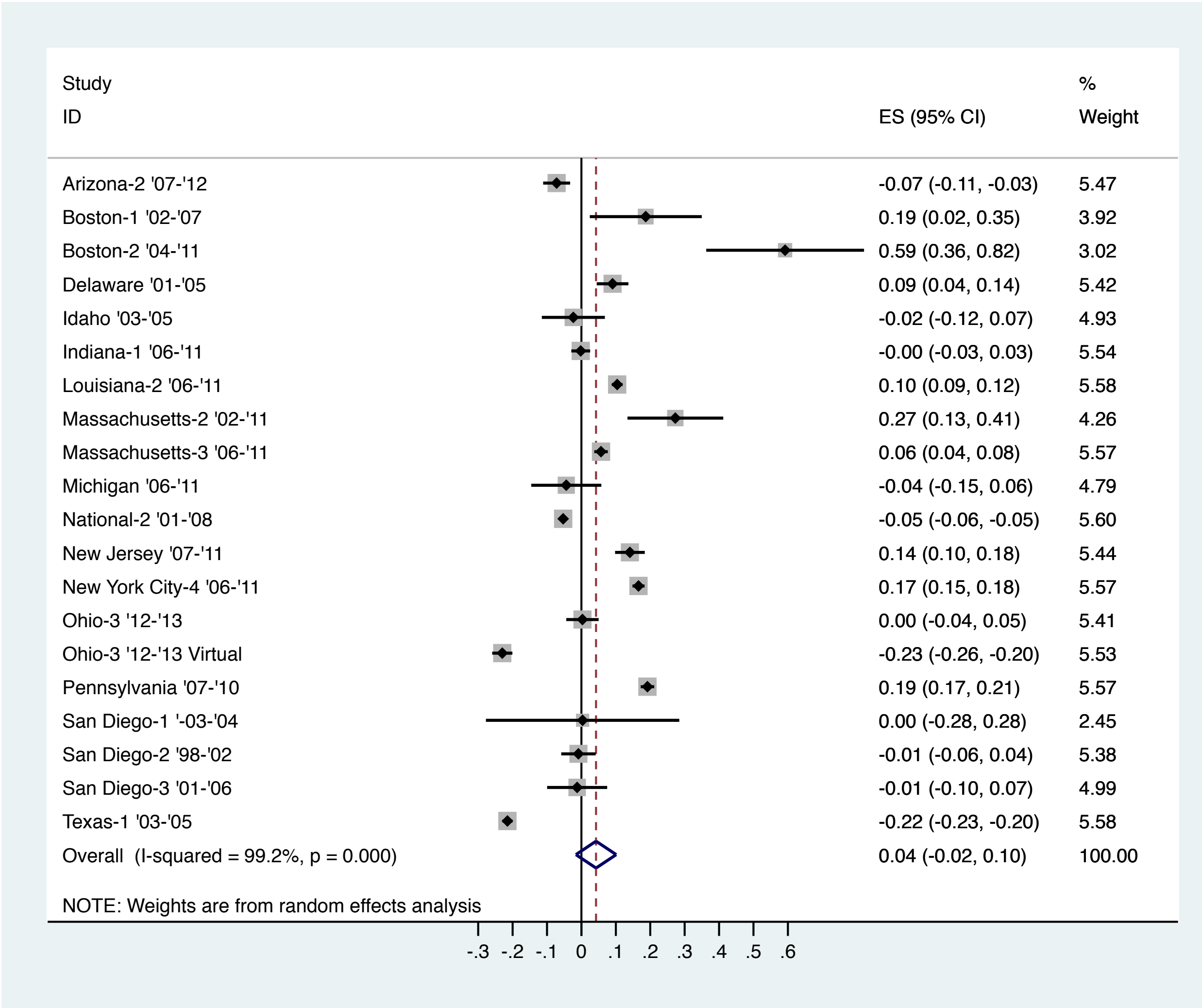
Figures 5 and 6 show corresponding figures for high school results. Although our overall estimates are insignificantly different from zero, there are a number of individual studies that find statistically significant positive and negative effects of attending a charter school. The overall estimated effect size for reading is positive, but four significant large negative estimates (from Michigan, Texas, Massachusetts, and virtual charter schools in Ohio) counteract a large number of small positive effect size estimates from other studies. A similar pattern emerges for math in figure 6, with far more positive and significant results than negative and significant results, but the overall estimated effect size of 0.042 is not statistically significant. As in high school reading, in high school math the Texas study and the study of virtual charter schools in Ohio contribute precisely estimated large negative effects.

Figure 5. High School Reading Effect Sizes by Study, Showing Weights Ascribed by Random-Effects Meta-Analysis to Each Study



Notes: Weights are from random-effects analysis. The horizontal lines show the 95 percent confidence interval, which is also indicated in the second column from the right. The rightmost column shows the weight ascribed to each study, with the size of the square proportional to these weights. The overall effect size estimate is shown at the bottom. Geographic locations with estimates from multiple studies have unique numbers appended to their labels to distinguish between studies. Table 7, in the appendix, indicates the author and year of the study referenced by each study ID label.

Figure 6. High School Math Effect Sizes by Study, Showing Weights Ascribed by Random-Effects Meta-Analysis to Each Study



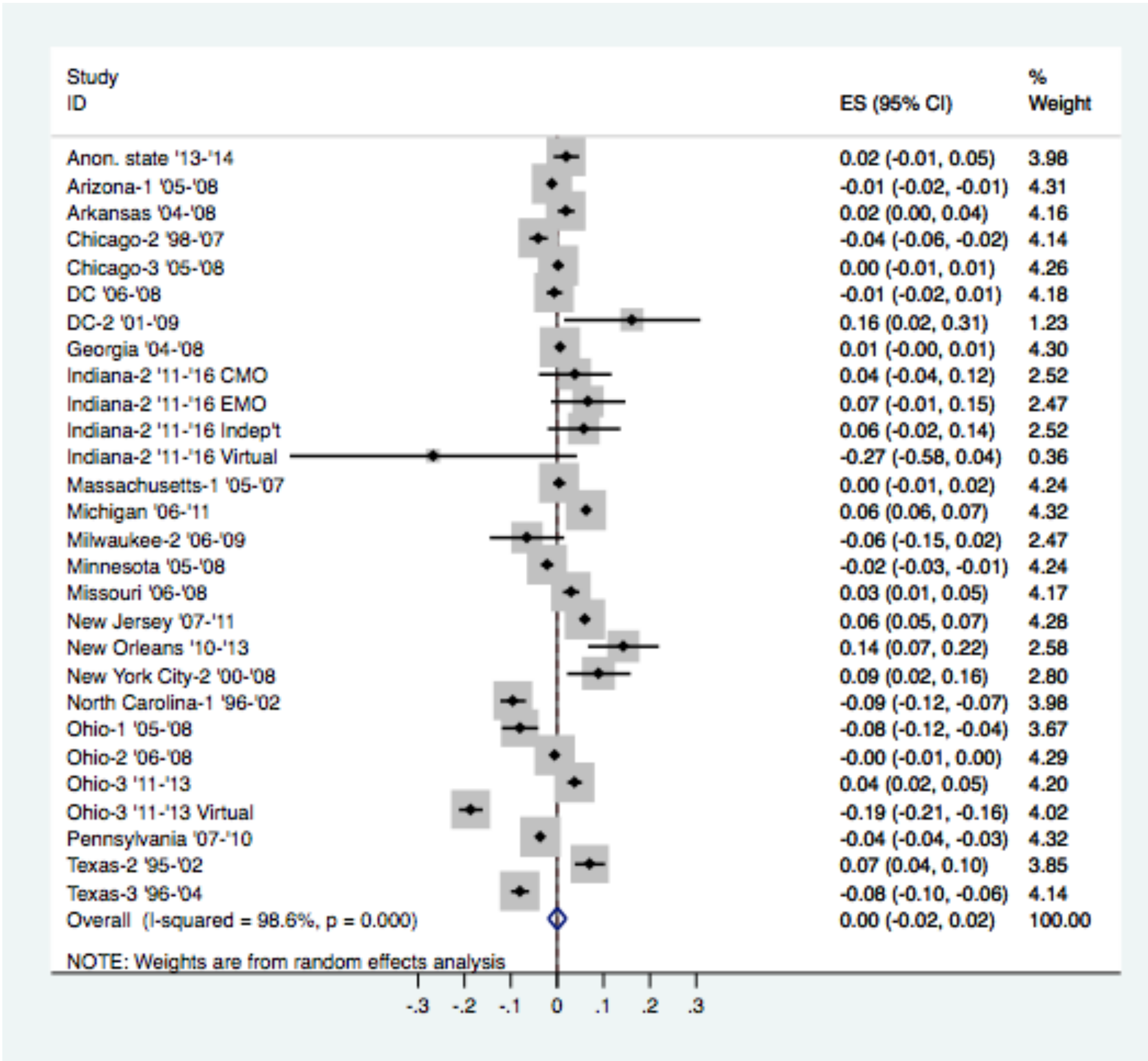
Notes: Weights are from random-effects analysis. The horizontal lines show the 95 percent confidence interval, which is also indicated in the second column from the right. The rightmost column shows the weight ascribed to each study, with the size of the square proportional to these weights. The overall effect size estimate is shown at the bottom. Geographic locations with estimates from multiple studies have unique numbers appended to their labels to distinguish between studies. Table 7, in the appendix, indicates the author and year of the study referenced by each study ID label.

Figures 7 and 8 show the results from the studies that combine elementary and middle schools, for which overall we find no significant effects. Considerable variation emerges for reading in figure 7, with studies of Washington, D.C., New Orleans, the Promise Academy in the Harlem Children's Zone in New York City, and Texas showing the largest positive effects, and two studies of virtual charter schools in Indiana and Ohio, and studies of all charter schools in North Carolina, Ohio, and Texas showing the largest negative effects.

Figure 8 shows estimates for math from studies that combine elementary and middle schools. Again, the insignificant overall estimate masks considerable variation. The studies with the largest estimated positive effects come from New Orleans, New York City, New Jersey, and Texas. Washington D.C. contributes a large positive effect with a large standard error making the estimate not significantly different from zero. The largest estimated negative effects come from studies of virtual charter schools in Indiana and Ohio, studies of all charter schools in Ohio, North Carolina, and Texas, and Indiana charters managed by Education Management Organizations (EMO's). (The same pair of Texas studies that produces the contradictory estimates outlined above for reading also produces the quite large contradictory results for math.)

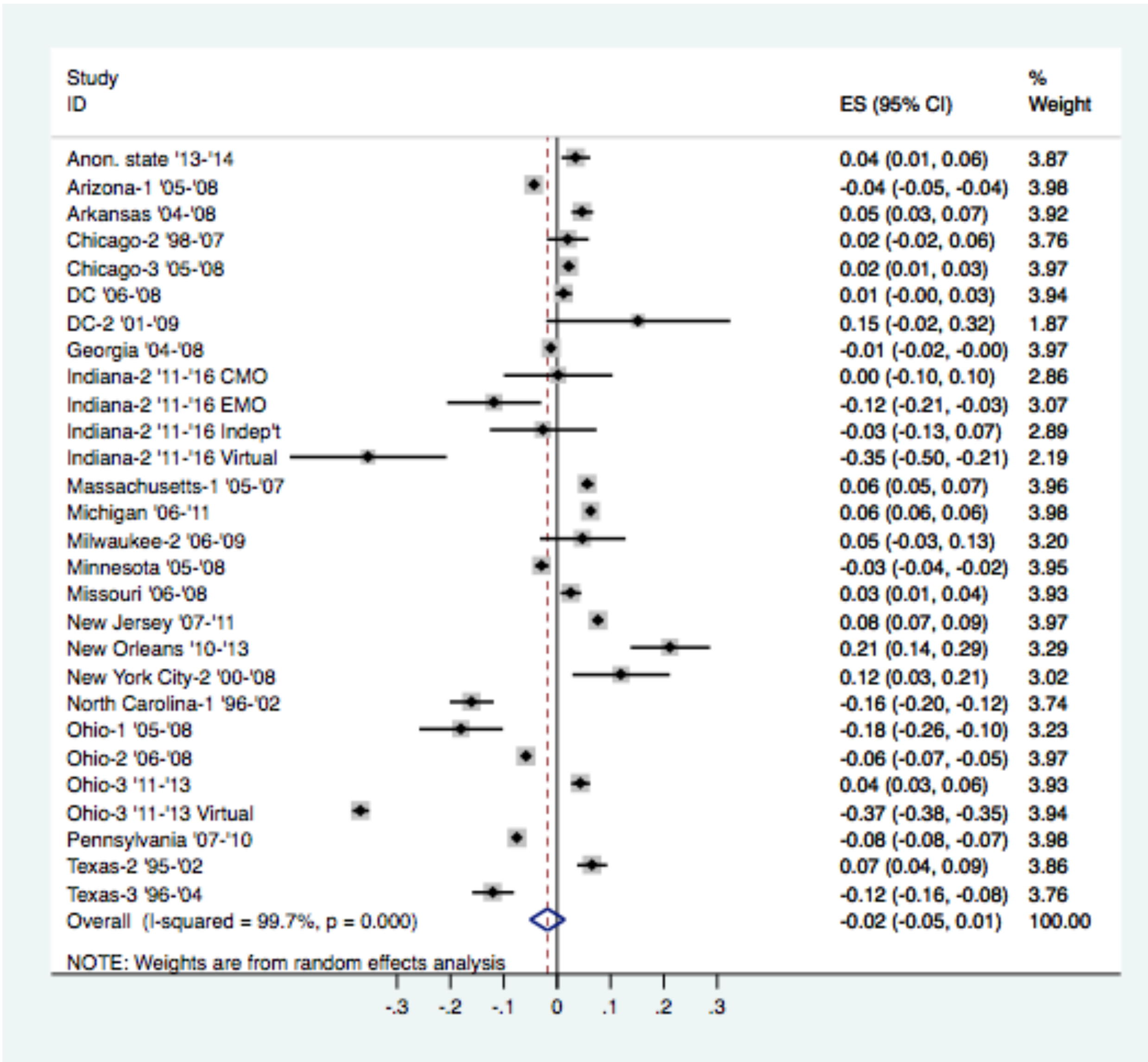
Figures 9 and 10 show reading and math results for the "All" grade span studies, which in the case of reading and math produced an overall positive effect size that was insignificantly different from zero. For reading, as shown in figure 9, most of the effect sizes are clustered in a narrow band on either side of zero. The main exceptions are positive effect sizes found for Delaware and Louisiana. For math, as shown in figure 10, the overall estimate is positive but not statistically significant at the 5 percent level. There are four large positive effect size estimates, for New York City, Indianapolis, Denver, and Idaho, but the latter three of these receive a small weight in the overall estimate because they are estimated quite imprecisely compared to the other studies that mostly have effect sizes near zero. Two studies of North Carolina generated the most negative estimates.

Figure 7. Reading Effect Sizes for Studies That Combine Elementary and Middle Schools by Study, Showing Weights Ascribed by Random-Effects Meta-Analysis to Each Study



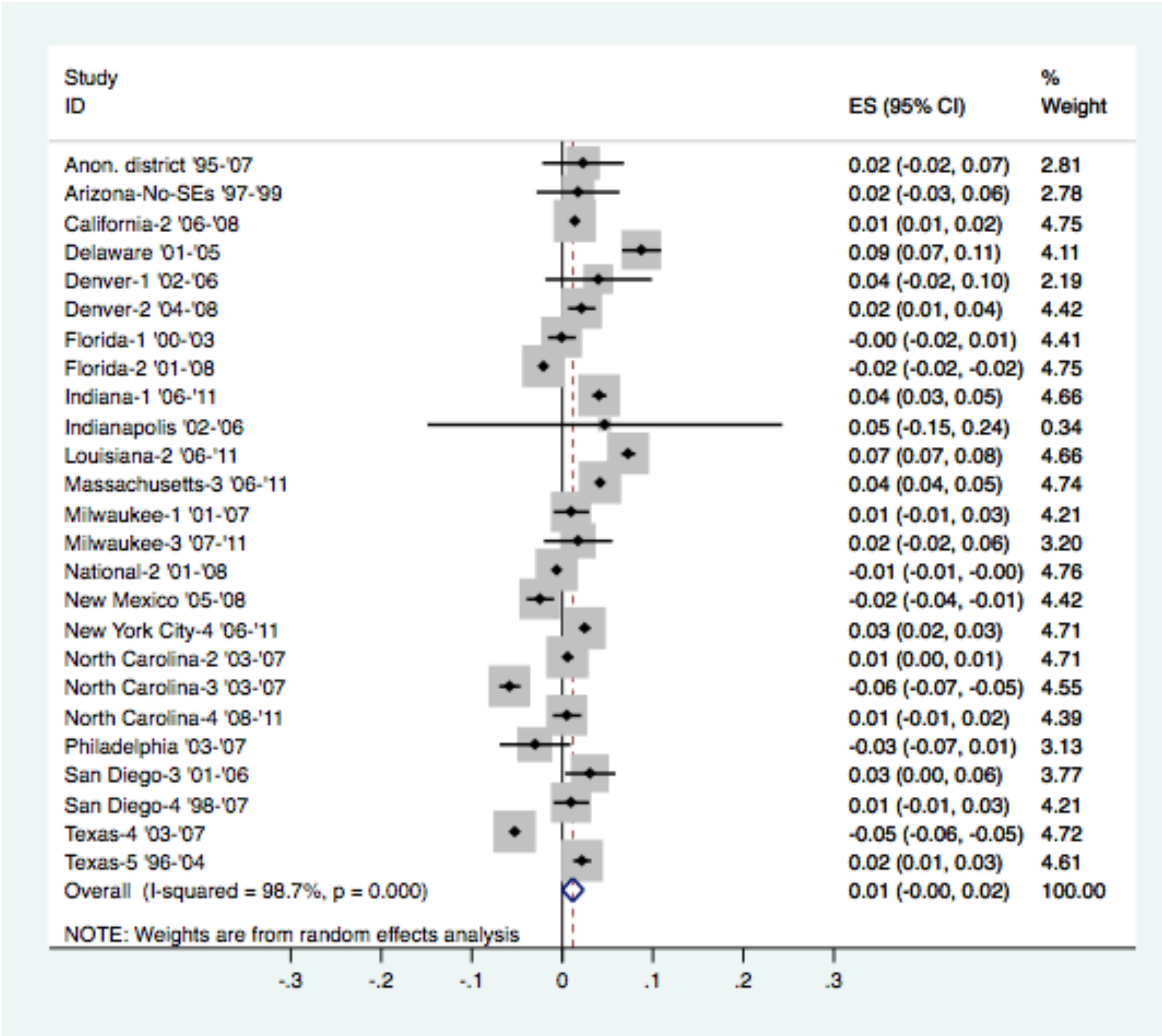
Notes: Weights are from random-effects analysis. The horizontal lines show the 95 percent confidence interval, which is also indicated in the second column from the right. The rightmost column shows the weight ascribed to each study, with the size of the square proportional to these weights. The overall effect size estimate is shown at the bottom. Geographic locations with estimates from multiple studies have unique numbers appended to their labels to distinguish between studies. Table 7, in the appendix, indicates the author and year of the study referenced by each study ID label. The studies in this figure are referred to in table 1 as "Combined Elementary/Middle Studies."

Figure 8. Math Effect Sizes for Studies That Combine Elementary and Middle Schools by Study, Showing Weights Ascribed by Random-Effects Meta-Analysis to Each Study



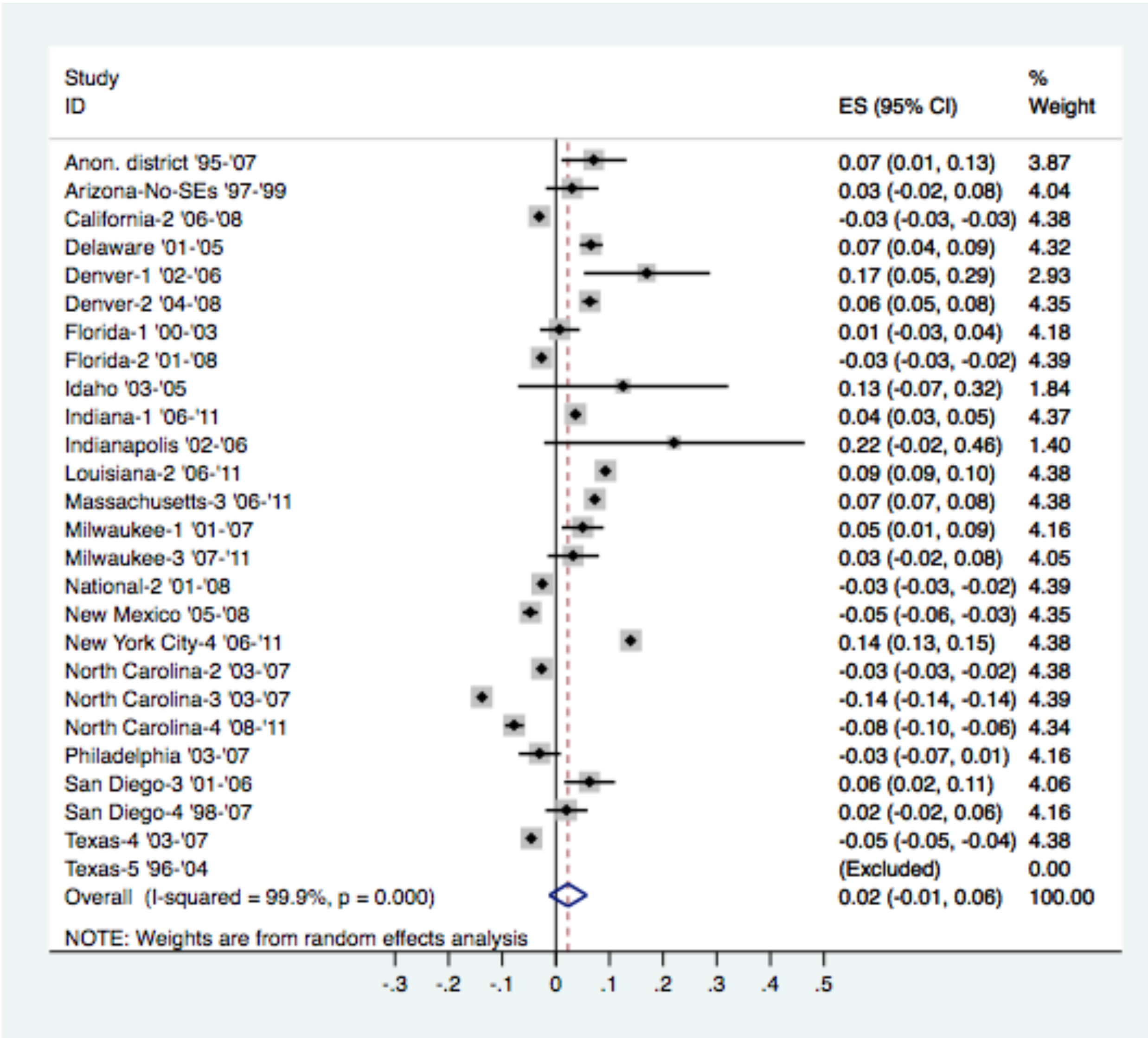
Notes: Weights are from random-effects analysis. The horizontal lines show the 95 percent confidence interval, which is also indicated in the second column from the right. The rightmost column shows the weight ascribed to each study, with the size of the square proportional to these weights. The overall effect size estimate is shown at the bottom. Geographic locations with estimates from multiple studies have unique numbers appended to their labels to distinguish between studies. Table 7, in the appendix, indicates the author and year of the study referenced by each study ID label. The studies in this figure are referred to in table 1 as "Combined Elementary/Middle Studies."

Figure 9. Reading Effect Sizes for Studies That Combine Elementary, Middle, and High Schools by Study, Showing Weights Ascribed by Random-Effects Meta-Analysis to Each Study



Notes: Weights are from random-effects analysis. The horizontal lines show the 95 percent confidence interval, which is also indicated in the second column from the right. The rightmost column shows the weight ascribed to each study, with the size of the square proportional to these weights. The overall effect size estimate is shown at the bottom. Geographic locations with estimates from multiple studies have unique numbers appended to their labels to distinguish between studies. Table 7, in the appendix, indicates the author and year of the study referenced by each study ID label. The studies in this figure are referred to in table 1 as "All" grade spans.

Figure 10. Math Effect Sizes for Studies That Combine Elementary, Middle, and High Schools by Study, Showing Weights Ascribed by Random-Effects Meta-Analysis to Each Study



Notes: Weights are from random-effects analysis. The horizontal lines show the 95 percent confidence interval, which is also indicated in the second column from the right. The rightmost column shows the weight ascribed to each study, with the size of the square proportional to these weights. The overall effect size estimate is shown at the bottom. Geographic locations with estimates from multiple studies have unique numbers appended to their labels to distinguish between studies. Table 7, in the appendix, indicates the author and year of the study referenced by each study ID label. The studies in this figure are referred to in table 1 as “All” grade spans.

Estimated Effects for KIPP Middle Schools Are Far Higher Than for Other Charter Middle Schools

The middle school results presented in tables 1, 2a and 2b and in figures 3 and 4 exclude the many estimates for individual KIPP schools. Table 3 shows the results of a meta-analysis that includes only the KIPP schools. This can be thought of as the second meta-analysis of the KIPP literature, following up on the similar analysis in Betts and Tang (2011). KIPP schools appear to have a statistically significant and positive influence on both reading and math achievement, with the effect size for math being twice as large as for reading.[5]

Table 3. KIPP School Estimates: Effect Sizes and Significance by Grade Span and Subject Area

Grade Span	Reading Tests (# estimates-# locations), % true variation	Math Tests (# estimates-# locations), % true variation
Including Only KIPP Estimates		
Middle	0.174* (8-5), 85.2%	0.374* (8-5), 94.2%

Note: Asterisks indicate effect size significantly different from zero at the 5 percent level or less. The numbers in parentheses indicate the number of estimates included in the associated estimate of effect size, and the number of locales, which in the case of KIPP schools is unknown due to the shielding of charter school identities in one study.

The estimates for KIPP middle schools are far higher than our average estimates in table 1, with estimated effect sizes for reading and math of 0.174 and 0.374 respectively. These effect sizes are enough to move a student initially at the 50th percentile to percentiles 56.9 and 64.6 in a single year of attendance at a KIPP school. These are very large effects, by any standard.

5. The effect sizes for KIPP are materially higher than the effects estimated by Betts and Tang (2011), which were 0.096 and 0.223 respectively. Since our earlier literature review, the preliminary report from a national study (Tuttle et al. 2010) has been replaced by a final report (Tuttle et al. 2013), and we have substituted the single effect size by subject in the later report for the many school-level estimates in the earlier report on the advice of one of the authors (Brian Gill, personal communication, February 2014).

4. Understanding Sources of Variation across Locations, Types of Charters and Types of Students

The earlier results show that most variation across studies is real. Betts and Atkinson (2012) argue that there should not be a single estimate of the gains from school choice, because not only do schools of choice differ in the quality of education they provide, but so do the traditional public schools against which schools of choice compete. A recent paper by Chabrier, Cohodes and Oreopoulos (2016) shows this latter source of heterogeneity clearly, with the estimated impact of charter schools falling sharply as average test scores increase at the traditional public schools which charter applicants would otherwise attend.

Variations across types of charters surely matter too. First, our analysis suggests that charters operating at the middle school level produce larger gains in achievement, relative to traditional public schools, than charters at other grade spans. Second, the literature on KIPP schools, and the broader literature on "No Excuse" charter schools, including the KIPP schools, which feature extended school time, uniforms and parent contracts, suggests that these types of charters consistently outperform traditional public schools. A smaller literature examines "virtual" charter schools, also known as online charter schools. A multi-state report by CREDO (2015) and studies in Indiana and Ohio suggest that on average these schools produce learning gains much below that of traditional public schools. (Our study includes only the latter two results due to a lack of information on standard errors in the first report.)

Results for Student Subgroups

Turning to variations by students, a key question is whether charter schools benefit students from some groups more than others. A warning about the subgroup estimates in this section is that only a subset of studies have performed these analyses, meaning that they may not apply to the locations studied nationally. For this reason, comparing the results below to the overall results for the same grade span in Table 1 is likely to be misleading.

Table 4 shows estimated effect sizes from meta-analyses for three at-risk populations: students in special education, English Language Learners (ELLs), and students eligible for federal meal assistance, the last of which is a commonly used proxy for poverty, for grade spans with multiple estimates available.

Results for the three types of at-risk populations are mixed, and difficult to summarize simply. No subpopulation appears to do worse when attending charter schools, but beyond that the impact of charter schools varies by subject and grade span. For students eligible for meal assistance, results are positive and statistically significant for all three grade spans for reading and for two of three grade spans for math.

Table 4. Effect Sizes for Studies of Selected Subsamples of Student Populations and Significance From Meta-Analysis, by Grade Span and Subject Area

	Grade Span		
Student Population	**Combined E/M (# estimates-# locations), % true variation**	**E, M and Combined E/M (# estimates-# locations), % true variation**	**All (# estimates-# locations), % true variation**
	READING TESTS		
Students in Special Education	-0.002 (12-12), 79.9%	-0.002 (12-12), 79.9%	0.025* (10-10), 82.7%
English Language Learners	0.005 (12-12), 73.5%	0.010 (13-12), 76.5%	0.032 (10-10), 87.7%
Students Eligible for Federal Meal Assistance	0.020* (16-13), 86.7%	0.021* (18-15), 84.9%	0.028* (10-10), 93.8%
	MATH TESTS		
Students in Special Education	0.002 (12-12), 79.4%	0.002 (12-12), 79.4%	0.017* (10-10), 0.0%
English Language Learners	0.027 (12-12), 81.3%	0.027 (13-12), 80.3%	0.015 (10-10), 58.6%
Students Eligible for Federal Meal Assistance	0.013 (16-13), 89.6%	0.023* (18-15), 90.4 %	0.022* (10-10), 93.1%

Note: Asterisks indicate effect size significantly different from zero at the 5 percent level or less.

Table 5 shows results from grade spans with multiple studies by race/ethnicity.[6] Interestingly, in most but not all cases, results are negative and significant for both reading and math for white and Asian students. This contrasts with results for black students, where coefficients are always positive and significant in two cases. For Hispanic students results are never significant. These results are not necessarily representative of all the geographic locations studied in the literature: the number of locations studied is about one third to one half the number of locations studied for the results for all students in Table 1. But the results provide at least weak evidence that black students may gain more than white or Asian students. This could of course reflect the relative effectiveness of traditional public schools available to the various racial/ethnic groups.

6. The table does not separately show results for studies of E, M, and H grade spans as there are only one or two studies at these grade spans.

Table 5. Effect Sizes for Studies of Racial/Ethnic Subsamples of Student Populations and Significance From Meta-Analysis, by Grade Span and Subject Area

Note: Asterisks indicate effect size significantly different from zero at the 5 percent level or less.

	Grade Span		
Race/Ethnicity	**Combined E/M (# estimates-# locations), % true variation**	**E, M and Combined E/M (# estimates-# locations), % true variation**	**All (# estimates-# locations), % true variation**
	READING TESTS		
White	-0.032* (19-14), 97.3%	-0.040* (22-17), 96.9%	-0.020* (14-13), 97.5%
Black	0.031* (20-14), 93.3%	0.027* (23-16), 92.4%	0.006 (14-13), 97.2%
Hispanic	-0.021 (20-14), 93.3%	-0.023 (23-16), 88.6%	-0.001 (14-13), 95.7%
Native American	-0.142 (9-9), 95.4%	-0.142 (9-9), 95.4%	-0.054 (9-9), 45.8%
Asian	-0.026 (12-12), 59.4%	-0.033* (14-13), 61.1%	-0.051* (10-10), 95%
	MATH TESTS		
White	-0.081* (19-14), 99.0%	-0.082* (22-16), 98.6%	-0.012 (14-13), 98.7%
Black	0.021 (20-14), 96.1%	0.024 (23-16), 95.6%	0.024 (14-13), 98.6%
Hispanic	-0.011 (20-14), 93.7%	-0.008 (23-16), 93.5%	0.019 (14-13), 98%
Native American	-0.034 (7-7), 67.6%	-0.034 (7-7), 67.6%	-0.077* (9-9), 57.9%
Asian	-0.46* (12-12), 78.3%	-0.058* (14-13), 77.8%	-0.037* (10-10), 64.2%

Urban Districts and Schools

Table 6 shows the results when we focus on studies of urban districts or on individual schools in urban areas. In all but one case the effect sizes are positive and statistically significant. In contrast, in the overall sample (Table 1) effects were positive and significant for five of 12 cases. The math and effect size estimates are higher in the urban subsample shown in Table 6 than in the overall sample shown in Table 1 for all cases except reading in the combined elementary/middle school grade span. As in Table 1, we continue to exclude KIPP schools from the analysis of urban charter schools, even though KIPP schools typically locate in urban settings. Including them would have increased the effect sizes in table 6 considerably.

Table 6. Effect Sizes for Studies of Urban Districts and Schools, by Grade Span and Subject Area

Grade Span	Reading Tests (# estimates-# locations), % true variation	Math Tests (# estimates-# locations), % true variation
Elementary	0.034* (9-5), 77.8%	0.054* (9-5), 91.5%
Middle	0.098* (10-5), 95.2%	0.176* (10-5), 97.2%
High	0.121* (8-4), 93.8%	0.132* (6-3), 92.8%
Combined Elementary/Middle	-0.002 (6-4), 82.6%	0.023* (6-4), 40.6%)
Elementary, Middle, and Combined Elementary/Middle	0.050* (25-7), 96.0%	0.102* (25-7), 98.8%
All	0.019* (10-7), 21.8%	0.064* (10-7), 95.9%

Note: Asterisks indicate effect size significantly different from zero at the 5 percent level or less.

There could be multiple reasons for the larger effects in urban settings. One obvious possibility is that charter schools have more value to add in large urban districts if the traditional schools in these areas are underserving their students to a greater extent than are their nonurban counterparts. Angrist et al. (2013) attributes the success of urban charter schools in Massachusetts to the "No Excuses" approach to education, which the authors describe as emphasizing "discipline and comportment, traditional reading and math skills, instruction time, and selective teacher hiring."

7. Conclusion

When studying charter schools' effect on achievement, one can ask several related questions. First, on *average* are charter schools boosting achievement relative to traditional public schools? Second, are there exceptions where we can find some charters outperforming and others underperforming? Third, what have social scientists learned about the types of charter schools, the types of students, or the school settings, which can predict the relative effectiveness of charter schools?

On the question of the average overall effect, for no grade span or subject tested did we find a negative average effect of charter schools. For five out of 12 combinations of grades and subject areas, we found a positive and significant overall effect of charter schools on achievement. (For six of the remaining seven cases, the estimated average effect was positive but not statistically significant.) The results are more compelling for math than reading, both in terms of the number of grade spans for which we found significant effects, and the magnitude of those significant effects. In cases of significant average effects, students are often predicted to gain about one half to one percentile a point per year, but with larger estimates for charter schools at the middle school level.

On the second question, we find considerable heterogeneity in effect sizes across studies. Overall, our findings confirm that the impact of the charter sector on student outcomes varies considerably—especially across geographic areas. This likely reflects variations not only in charter schools' effectiveness but also in that of the traditional public schools against which the charters are compared. There is little doubt that charters outperform traditional public schools in some areas and underperform them in others -- most of the variation across studies is real and does not reflect statistical "noise".

On the third question, the factors related to charter effectiveness, social scientists are beginning to make some progress. Charter schools in urban areas on average produce strong positive effects, especially in math. Charter schools operating as middle schools produce quite high predicted effects for both math and reading relative to charter schools in other grade spans. Only a fraction of studies break down the charter schools' effects by student characteristics, but the existing papers tentatively suggest that black students and those receiving federal meal assistance are among the students most likely to benefit from attending a charter school. A number of studies point to preliminary evidence that "No Excuse" charter schools, such as the KIPP schools, produce unusually large gains. Conversely, a handful of recent studies suggests that virtual (online) charter schools are markedly less effective than traditional public schools.

Looking forward, we expect that new studies will continue to appear, and over time some of the above conclusions will evolve as we gain a fuller picture of charter schools' effects. Of the three sets of results, on average effects, the degree of variation in effects, and factors related to variation, the result most likely to stand the test of time is that the relative performance of charter schools and traditional public schools is highly variable. A challenge for the charter school movement will be to ensure going forward that those charters that underperform either improve or, over time, shut down.

Appendix. Details on Meta-Analysis

We assume that the effect of charter schools on achievement is not fixed across studies. Given that charter schools are afforded considerable freedom to experiment, and that the regulatory framework for charter schools varies across states, and surely across individual districts as well, it would seem untenable to make the alternative assumption that there is a single fixed impact of charter schools on achievement.[7]

In a random effects meta-analysis, we take a weighted average of the effect sizes across studies. If Y_i is the effect size for the i[th] of k studies and W_i is the weight for each study, our overall estimated effect size *M* is:

$$(1) \qquad M = \frac{\sum_{i=1}^{k} W_i Y_i}{\sum_{i=1}^{k} W_i}$$

The weight for each study is the inverse of the sum of the within-study variance (based on the standard error) and an estimate of the true between-study variance, T^2:

$$(2) \quad W_i = \frac{1}{V_{Y_i} + T^2}$$

The between-studies variance estimate T^2 is based on a method of moments estimate of the variance of true effect sizes. Note that as T^2 becomes large relative to the average within-study variance estimate, then we will tend towards equal-weighting across studies, whereas as T^2 becomes relatively small the weights can become highly unequal with heavier weight given to studies with the lowest sampling variance.

We report the I^2 statistic introduced by Higgins et al. (2003), which provides an estimate of the percentage of the variation in effect sizes that reflects true underlying variation.

---

7 For a review of the random-effects approach to meta-analysis and measures of heterogeneity, see Borenstein et al. (2009), chapters 12 through 16.

Table 7 Details on the Studies Used in Any of Our Approaches

Authors	Year Published	Name of State or City	First Year of Data	Final Year of Data	Study Label in Meta-Analysis Plots	Grade spans Included in Meta-Analysis of Effect Size
Abdulkadiroglu et al.	2009	Boston	2002	2007	Boston-1 '02-'07	E, M, H
Abdulkadiroglu et al.	2016	Boston, New Orleans	2012 2010	2014 2013	New Orleans '10-'13	M EM
Ahn and McEachin	2017	Ohio	2011 2013	2013 2013	Ohio-3 '11-'13 and Ohio-3 '11-'13, virtual	EM, virtual and other charters H, virtual and other charters
Angrist et al.	2012	Boston (1 KIPP school)	2006	2009	kipp-lynn	M
Angrist, Pathak, and Walters	2013	Massachusetts	2002	2011	Massachusetts-2 '02-'11	M
Angrist et al.	2016	Boston	2004	2011	Boston-2 '04-'11	H
Ballou et al.	2006	Idaho	2003	2005	Idaho '03-'05	E, M, H, A
Betts et al.	2005	San Diego	1998	2002	San Diego-2 '98-'02	E, M, H
Betts, Tang, and Zau	2010	San Diego	2001	2006	San Diego-3 '01-'06	E, M, H, A
Bifulco and Ladd	2006	North Carolina	1996	2002	North Carolina-1 '96-'02	EM
Booker et al.	2007	Texas	1995	2002	Texas-2 '95-'02	EM
Buddin and Zimmer	2003	California	1998	2002	California-1 '98-'02	E
Chingos and West	2015	Arizona	2007	2012	Arizona-2 '07-'12	E, M, H
CREDO	2009	National	2001	2008	National-2 '01-'08	E, M, H
CREDO	2009	Arizona, Arkansas, California, Chicago,	varies	varies	Arizona-1 '05-'08 Arkansas '04-'08 California-2 '06-'08 Chicago-3 '05-'08	EM (9 locations), A (7 locations)

		Denver, DC, Florida, Georgia, Massachusetts, Minnesota, Missouri, New Mexico, North Carolina, . Ohio, Texas			Denver-2 '04-'08 DC '06-'08 Florida-2 '01-'08 Georgia '04-'08 Massachusetts-1 '05-'07 Minnesota '05-'08 Missouri '06-'08 New Mexico '05-'08 North Carolina-2 '03-'07 Ohio-2 '06-'08 Texas-4 '03-'07	
CREDO	2013	Louisiana, Massachusetts, Michigan, New York City	varies	varies	Louisiana-2 '06-'11 Massachusetts-3 '06-'11 Michigan '06-'11 New York City-4 '06-'11	E, M, H, A (LA, MA, NYC) E, M, EM (MI)
CREDO	2012	Indiana, New Jersey	varies	varies	Indiana-1 '06-'11 New Jersey '07-'11	E, M, H, A (IN) E, M, EM (NJ)
CREDO	2011	Pennsylvania	varies	varies	Pennsylvania '07-'10	E, M, H, EM)
Dobbie and Fryer	2011	NYC (1 school, Promise Academy in Harlem Children's Zone)	2004	2009	New York City-3 '04-'09	E, M
Dobbie and Fryer	2016	Texas	1996	2004	Texas-5 '96-'04	A
Ferrare, Waddington and Berends	2017	Indiana	2011	2016	Indiana-2 '11-'16 (with indications for school type: CMO, EMO, Independent, Virtual)	EM
Foreman et al.	2017	Anon. state	2013	2014	Anon. state '13-'14	EM
Furgeson et al.	2012	CMOs multiple states	varies	varies	CMOs multiple states	M
Gleason et al.	2010	National (29 schools)	2004	2008	National-3 '04-'08	M

Gronberg and Jansen	2005	Texas	2003	2004	Texas-1 '03-'04	M, H
Hoxby and Murarka	2007	NYC	2004	2006	New York City-1 '04-'06	E
Hoxby, Murarka, and Kang	2009	NYC	2000	2008	New York City-2 '00-'08	EM
Hoxby and Rockoff	2004	Chicago	2001	2004	Chicago-1 '01-'04	E, M
Imberman	2010	Anonymous district	1995	2005	Anon. district '95-'07	A
Ladd Clotfelter and Holbein	2017	North Carolina	2003 2008	2007 2011	North Carolina-3 '03-'07 North Carolina-4 '08-'11	A A
McClure et al.	2005	San Diego	2003	2004	San Diego-1 '-03-'04	H
Miron et al.	2007	Delaware	2000	2005	Delaware '01-'05	E, M, H, A
Ni and Rorrer	2012	Utah	2004	2009	Utah '04-'09	E
Nichols and Özek	2010	DC	2001	2009	DC-2 '01-'09	EM
Nicotera, Mendiburo, and Berends	2011	Indianapolis	2002	2006	Indianapolis '02-'06	A
Nisar	2012	Milwaukee	2006	2009	Milwaukee-2 '06-'09	EM
Sass	2006	Florida	2000	2003	Florida-1 '00-'03	A
Shin, Fuller and Dauter	2017	Los Angeles (separate analyses for movers and stayers)	2008	2011	Los Angeles '08-'11 Switch, Los Angeles '08-'11 Stay	E, M, H
Tuttle et al.	2013	KIPP multiple states	2002	2011	KIPP multiple states '02-'11	M
Witte, Wolf, Carlson, and Dean	2010, 2011, 2012 (average of one year gains from each	Milwaukee	2007	2011	Milwaukee-3 '07-'11	A

	of four years)					
Woodworth et al.	2008	Bay Area (3 KIPP schools)	2003	2005	kipp-bayarea-A, -B and -C	M
Zimmer et al.	2009 (sub-group estimates) 2012 (all students)	Chicago, Denver, Milwaukee, Ohio, Philadelphia, San Diego, Texas	varies	varies	Chicago-2 '98-'07, Denver-1 '02-'06 Milwaukee-1 '01-'07 Ohio-1 '05-'08 Philadelphia '03-'07 San Diego-4 '98-'07 Texas-3 '96-'04	EM (3 locations), A (4 locations)

Note: E, M, H and A stand for analyses of elementary, middle, high schools, and all grades, respectively, and EM stands for combined elementary and middle.

References

Abdulkadiroğlu, Atila, Joshua D. Angrist, Sarah R. Cohodes, Susan M. Dynarski, Jon Fullerton, Thomas J. Kane, and Parag A. Pathak. 2009. *Informing the Debate: Comparing Boston's Charter, Pilot and Traditional Schools*. Boston, MA: The Boston Foundation.

Abdulkadiroğlu, Atila, Joshua D. Angrist, Peter D. Hull, and Parag A. Pathak. 2016. "Charters without Lotteries: Testing Takeovers in New Orleans and Boston." *American Economic Review* 106 (7): 1878–1920.

Ahn, June, and Andrew McEachin. 2017. "Student Enrollment Patterns and Achievement in Ohio's Online Charter Schools." *Educational Researcher* 46 (1): 44–57. https://doi.org/10.3102/0013189X17692999.

Angrist, Joshua D., Sarah R. Cohodes, Susan M. Dynarski, Parag A. Pathak, and Christopher R. Walters. 2016. "Stand and Deliver: Effects of Boston's Charter High Schools on College Preparation, Entry, and Choice." *Journal of Labor Economics* 34 (2): 275–318.

Angrist, Joshua D., Susan M. Dynarski, Thomas J. Kane, Parag A. Pathak, and Christopher R. Walters. 2012. "Who Benefits from KIPP?" *Journal of Policy Analysis and Management* 31 (4): 837–60.

Angrist, Joshua D., Parag A. Pathak, and Christopher R. Walters. 2013. "Explaining Charter School Effectiveness." *American Economic Journal: Applied Economics* 5: 1–27.

Ballou, Dale, Bettie Teasley, and Tim Zeidner. 2006. "Charter Schools in Idaho." Paper prepared for the National Conference on Charter School Research, National Center on School Choice. Nashville, TN: Vanderbilt University. http://www.vanderbilt.edu/schoolchoice/conference/papers/Ballouetal_2006-DRAFT.pdf.

Betts, Julian R., and Richard C. Atkinson. 2012. "Better Research Needed on the Impact of Charter Schools." *Science* 335 (6065): 171–172.

Betts, Julian R., Lorien A. Rice, Andrew C. Zau, Y. Emily Tang, and Cory R. Koedel. 2005. *Does School Choice Work? Effects on Student Integration and Achievement*. San Francisco: Public Policy Institute of California.

Betts, Julian R., and Y. Emily Tang. 2008. "Value-Added and Experimental Studies of the Effect of Charter Schools on Student Achievement: A Literature Review." Seattle, WA: National Charter School Research Project, Center on Reinventing Public Education.

———. 2011. "The Effect of Charter Schools on Student Achievement: A Meta-Analysis of the Literature." Seattle, WA: National Charter School Research Project, Center on Reinventing Public Education.

———. 2014. "A Meta-Analysis of the Literature on the Effect of Charter Schools on Student Achievement." Bothell, WA: Center on Reinventing Public Education. http://www.crpe.org/sites/default/files/CRPE_meta-analysis_charter-schools-effect-student-achievement_workingpaper.pdf.

Betts, Julian R., Y. Emily Tang, and Andrew C. Zau. 2010. "Madness in the Method? A Critical Analysis of Popular Methods of Estimating the Effect of Charter Schools on Student Achievement." In *Taking Measure of Charter Schools: Better Assessments, Better Policymaking, Better Schools*, edited by Julian R. Betts and Paul T. Hill. Lanham, MD: Rowman & Littlefield.

Bifulco, Robert, and Helen F. Ladd. 2006. "The Impacts of Charter Schools on Student Achievement: Evidence from North Carolina." *Education Finance and Policy* 1: 50–90.

Booker, Kevin, Scott Gilpatric, Timothy Gronberg, and Dennis Jansen. 2007. "The Impact of Charter School Attendance on Student Performance." *Journal of Public Economics* 91: 5–6.

Borenstein, Michael, Larry V. Hedges, Julian P.T. Higgins, and Hannah R. Rothstein. 2009. *Introduction to Meta-Analysis*. Chichester: United Kingdom: John Wiley and Sons.

Buddin, Richard, and Ron Zimmer. 2003. "Appendix C: Academic Outcomes." In *Charter School Operations and Performance: Evidence from California*. Santa Monica, CA: RAND Corporation.

Chabrier, Julia, Sarah R. Cohodes, and Philip Oreopoulos. 2016. "What Can We Learn from Charter School Lotteries?" *Journal of Economic Perspectives* 30 (5): 57–84.

Charter School Achievement Consensus Panel. 2006. *Key Issues in Studying Charter Schools and Achievement: A Review and Suggestions for National Guidelines*. Vol. 2. National Charter School Research Project White Paper Series. Seattle, WA: Center on Reinventing Public Education.

Chingos, Matthew M., and Martin R. West. 2015. "The Uneven Performance of Arizona's Charter Schools." *Educational Evaluation and Policy Analysis 37* 1: 120–134.

Clotfelter, Charles T., Helen F. Ladd, and and Jacob L. Vigdor. 2007. *How and Why Do Teacher Credentials Matter for Student Achievement?* Cambridge, MA: National Bureau of Economic Research. http://www.nber.org/papers/w12828.

CREDO. 2009. *Multiple Choice: Charter School Performance in 16 States*. Stanford, CA: CREDO. http://credo.stanford.edu.

———. 2011. *Charter School Performance in Pennsylvania*. Stanford, CA: CREDO: Available. http://credo.stanford.edu.

———. 2012a. *Charter School Performance in Indiana*. Stanford, CA: CREDO: Available. http://credo.stanford.edu.

———. 2012b. *Charter School Performance in New Jersey*. Stanford, CA: CREDO. http://credo.stanford.edu.

———. 2013a. *Charter School Performance in Louisiana*. Stanford, CA: CREDO: Available. http://credo.stanford.edu.

———. 2013b. *Charter School Performance in Massachusetts*. Stanford, CA: CREDO: Available. http://credo.stanford.edu.

———. 2013c. *Charter School Performance in Michigan*. Stanford, CA: CREDO: Available. http://credo.stanford.edu.

———. 2013d. *Charter School Performance in New York City*. Stanford, CA: CREDO. http://credo.stanford.edu.

———. 2015. *Online Charter School Study*. Stanford, CA: CREDO.

Dobbie, Will, and Roland G. Fryer Jr. 2011. "Are High-Quality Schools Enough to Increase Achievement among the Poor? Evidence from the Harlem Children's Zone." *American Economic Journal: Applied Economics* 3 (3): 158–87.

Dobbie, Will S., and Roland G. Fryer Jr. 2016. *Charter Schools and Labor Market Outcomes*. Cambridge, MA: National Bureau of Economic Research.

Ferrare, Joseph J., Joseph R. Waddington, and Mark Berends. 2017. "Virtual Illusion: Longitudinal Effects of Charter School Types on Student Achievement in Indiana." Center for Research on Educational Opportunity, Notre Dame University. Notre Dame, IN.

Foreman, Leesa M., Kaitlin P. Anderson, Gary W. Ritter, and Patrick J. Wolf. 2017. "Using 'Broken' Lotteries to Check the Validity of Charter School Evaluations Using Matching Designs." *Educational Policy*, November.

Furgeson, Joshua, Brian Gill, Joshua Haimson, Alexandra Killewald, Moira McCullough, Ira Nichols-Barrer, Bing-ru Teh, et al. 2012. "Charter-School Management Organizations: Diverse Strategies and Diverse Student Impacts." Princeton, NJ & Seattle, WA: Mathematica Policy Research and Center for Reinventing Public Education.

Gleason, Philip, Melissa Clark, Christina Clark Tuttle, and Emily Dwoyer. 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.

Gronberg, Timothy J., and Dennis W. Jansen. 2005. *Texas Charter Schools: An Assessment in 2005*. Austin: Texas: Public Policy Foundation. http://www.texaspolicy.com/pdf/2005-09-charterschools-rr.pdf.

Higgins, J., S.G. Thompson, J.J. Deeks, and D. G. Altman. 2003. "Measuring Inconsistency in Meta-Analyses." *British Medical Journal 327* 7414: 557–560.

Hoxby, Caroline M., and Sonali Murarka. 2007. *New York City's Charter Schools Overall Report*. Cambridge, MA: New. http://users.nber.org/.

Hoxby, Caroline M., Sonali Murarka, and Jenny Kang. 2009. *How New York City's Charter Schools Affect Achievement*. Cambridge, MA: New. http://users.nber.org/.

Hoxby, Caroline M., and Jonah E. Rockoff. 2004. *"The Impact of Charter Schools on Student Achievement." Unpublished Paper*. Cambridge, MA: New. http://www.rand.org/content/dam/rand/www/external/labor/seminars/adp/pdfs/2005hoxby.pdf.

Imberman, Scott A. 2010. "Achievement and Behavior in Charter Schools: Drawing a More Complete Picture." *The Review of Economics and Statistics* 93 (2): 416–35.

Ladd, Helen F., Charles T. Clotfelter, and John B. Holbein. 2017. "The Growing Segmentation of the Charter School Sector in North Carolina." *Education Finance and Policy* 12 (4): 536–63.

McClure, Larry, Betsy Strick, Rachel Jacob-Almeida, and Christopher Reicher. 2005. "The Preuss School at UCSD: School Characteristics and Students' Achievement." University of California, San Diego, The Center for Research on Educational Equity, Assessment and Teaching Excellence. http://create.ucsd.edu/Research_Evaluation/PreussReportDecember2005.pdf.

Miron, Gary, Anne Cullen, Brooks Applegate, and Patricia Farrell. 2007. "Evaluation of the Delaware Charter School Reform: Year One Report." Kalamazoo, MI: The Evaluation Center, Western Michigan University. http://www.doe.k12.de.us/files/pdf/sbe_decseval.pdf.

National Alliance for Public Charter Schools. 2017. "Estimated Charter Public School Enrollment, 2016-17." https://www.publiccharters.org/sites/default/files/migrated/wp-content/uploads/2017/01/EER_Report_V5.pdf.

Ni, Yongmei, and Andrea K. Rorrer. 2012. "Twice Considered: Charter Schools and Student Achievement in Utah." *Economics of Education Review* 31 (5): 835–849.

Nichols, Austin, and Umut Özek. 2010. "Public School Choice and Student Achievement in the District of Columbia." Working Paper 53. Washington, DC: National Center for Analysis of Longitudinal Data. http://www.edweek.org/media/publicschoolchoiceandstudentachievement-calder.pdf.

Nicotera, Anna, Maria Mendiburo, and Mark Berends. 2011. "Charter School Effects in Indianapolis." In *School Choice and School Improvement*, edited by Mark Berends, Marisa Cannata, and Ellen B. Goldring, 35–50. Cambridge, Mass: Harvard Education Press.

Nisar, Hiren. 2012. "Do Charter Schools Improve Student Achievement?" In . New York: National Center for the Study of Privatization in Education. http://www.ncspe.org/publications_files/OP216.pdf.

Sass, Tim R. 2006. "Charter Schools and Student Achievement in Florida." *Education Finance and Policy* 1 (1): 91–122.

Shin, Hyo Jeong, Bruce Fuller, and Luke Dauter. 2017. "Heterogeneous Effects of Charter Schools: Unpacking Family Selection and Achievement Growth in Los Angeles." *Journal of School Choice* 11 (1): 60–94.

Tuttle, Christina Clark, Brian Gill, Philip Gleason, Virginia Knechtel, Ira Nichols-Barrer, and Alexandra Resch. 2013. *KIPP Middle Schools: Impacts on Achievement and Other Outcomes. Final Report*. Washington, DC: Mathematica Policy Research.

Tuttle, Christina Clark, Bing-ru Teh, Ira Nichols-Barrer, Brian P. Gill, and Philip Gleason. 2010. *Student Characteristics and Achievement in 22 KIPP Middle Schools*. Washington, DC: Mathematica Policy Research.

Witte, John F., Patrick J. Wolf, Deven Carlson, and Alicia Dean. 2012. "Milwaukee Independent Charter Schools Study: Final Report on Four-Year Achievement Gains." SCDP Evaluation Report #31. Fayetteville, AR: School Choice Demonstration Project. http://www.uaedreform.org/milwaukee-parental-choice-program-evaluation/.

Witte, John F., Patrick J. Wolf, Alicia Dean, and Deven Carlson. 2010. "Milwaukee Independent Charter Schools Study: Report on One Year of Student Growth." SCDP Evaluation Report #21. Fayetteville, AR: School Choice Demonstration Project. http://www.uaedreform.org/downloads/2011/03/report-21-milwaukee-independent-charter-schools-study-report-on-one-year-of-student-growth-2010-2011-update-and-policy-options.pdf

———. 2011. "Milwaukee Independent Charter Schools Study: Report on Two- and Three-Year Achievement Gains." SCDP Evaluation Report #21. Fayetteville, AR: School Choice Demonstration Project. http://files.eric.ed.gov/fulltext/ED518594.pdf.

Woodworth, Katrina R., Jane L. David, Roneeta Guha, Haiwen Wang, and Alejandra Lopez-Torkos. 2008. *San Francisco Bay Area KIPP Schools: A Study of Early Implementation and Achievement: Final Report*. Menlo Park, CA: SRI International.

Zimmer, Ron, Brian Gill, Kevin Booker, Stéphane Lavertu, Tim R. Sass, and John Witte. 2009. *Charter Schools in Eight States: Effects on Achievement, Attainment, Integration, and Competition*. Santa Monica, CA: Rand.