

**Student Characteristics
and Achievement in 22 KIPP
Middle Schools**

Final Report

June 2010

Christina Clark Tuttle
Bing-ru Teh
Ira Nichols-Barrer
Brian P. Gill
Philip Gleason



MATHEMATICA
Policy Research, Inc.

EXECUTIVE SUMMARY

The Knowledge Is Power Program (KIPP) is a bold effort to create a network of charter schools designed to transform and improve the educational opportunities available to low-income families. KIPP schools seek to actively engage students and parents in the educational process, expand the time and effort students devote to their studies, reinforce students' social competencies and positive behaviors, and dramatically improve their academic achievement. Ultimately, the goal of KIPP is to prepare students to enroll and succeed in college. The KIPP Foundation is guiding this effort by selecting and training school leaders, promoting the program model, and supporting the KIPP network schools. KIPP's "Five Pillars" distinguish its approach:

- **High expectations** for all students to reach high academic achievement, regardless of students' backgrounds
- **Choice and commitment** on the part of students, parents, and faculty to a public, college preparatory education as well as the time and effort required to reach success
- **More time** on learning, both in academics and extra-curricular activities, each day, week, and year
- **Power to lead** for school principals, who are accountable for their school's budget and personnel
- **Focus on results**, by regularly assessing student learning and sharing results to drive continuous improvement and accountability

KIPP has grown from a core of two middle schools established in the mid-1990s to a nationwide network of 82 schools in 19 states and the District of Columbia. In the wake of this growth, the KIPP Foundation, its funders, and other stakeholders are eager to rigorously assess the effectiveness of the program and identify which school practices may be positively related to student outcomes. The Foundation is sponsoring the National Evaluation of KIPP Middle Schools, conducted by Mathematica Policy Research, to examine the impacts of KIPP on the achievement and attainment of its students.

This report presents preliminary findings from a matched, longitudinal analysis designed to estimate KIPP's effect on student achievement. Our preliminary work estimates effects in 22 KIPP middle schools—making this the first report that applies a rigorous (nonexperimental) methodological approach across a nationwide sample of KIPP schools. We selected schools for which we were able to collect longitudinal, student-level data, and that were established by the 2005-06 school year or earlier to ensure that a minimum of two entering cohorts of students per school would be observed for multiple years.

We find that students entering these 22 KIPP schools typically had prior achievement levels that were lower than average achievement in their local school districts. **For the vast majority of KIPP schools studied, impacts on students' state assessment scores in mathematics and reading are positive, statistically significant, and educationally substantial.** Estimated impacts are frequently large enough to substantially reduce race- and income-based achievement gaps within three years of entering KIPP. We describe these findings in more detail in the pages below.

This report is the first in a series of reports that will be produced by the study, which is currently scheduled to continue through 2014. The report concludes by describing our plans for the remainder of the study, which will expand the sample to more KIPP schools, make use of randomized experimental research designs in a subset of schools, and incorporate additional outcome measures beyond state test scores.

Student Characteristics in 22 KIPP Middle Schools

To examine the characteristics of the students who enter KIPP middle schools, we compared the fourth grade characteristics of future KIPP students to those of non-KIPP students in the same districts and in the same elementary feeder schools (in other words, elementary schools attended by students who later enrolled in KIPP). We found no evidence that KIPP middle schools are systematically enrolling more advantaged students from their districts. In Figures ES.1 and ES.2, we show the overwhelming percentage of students at each school in our sample who are identified as racial/ethnic minorities and eligible for the federal free or reduced-price lunch (FRPL) program. **Over 90 percent of students at 19 of 22 KIPP schools are black or Hispanic and over 70 percent of students are FRPL-eligible at 14 of 15 KIPP schools for which we had access to that measure. On average, KIPP middle schools have student bodies characterized by higher concentrations of poverty and racial minorities, but lower concentrations of special education and limited English proficiency (LEP) students, than the public schools from which they draw.**

Figure ES.1. Percentage of Black and Hispanic Students in KIPP Schools

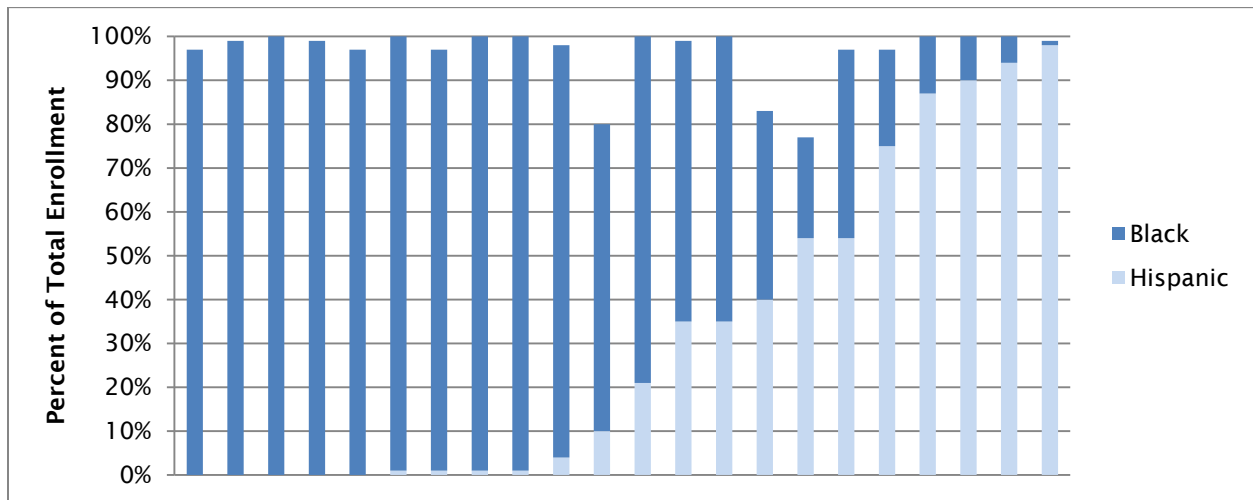
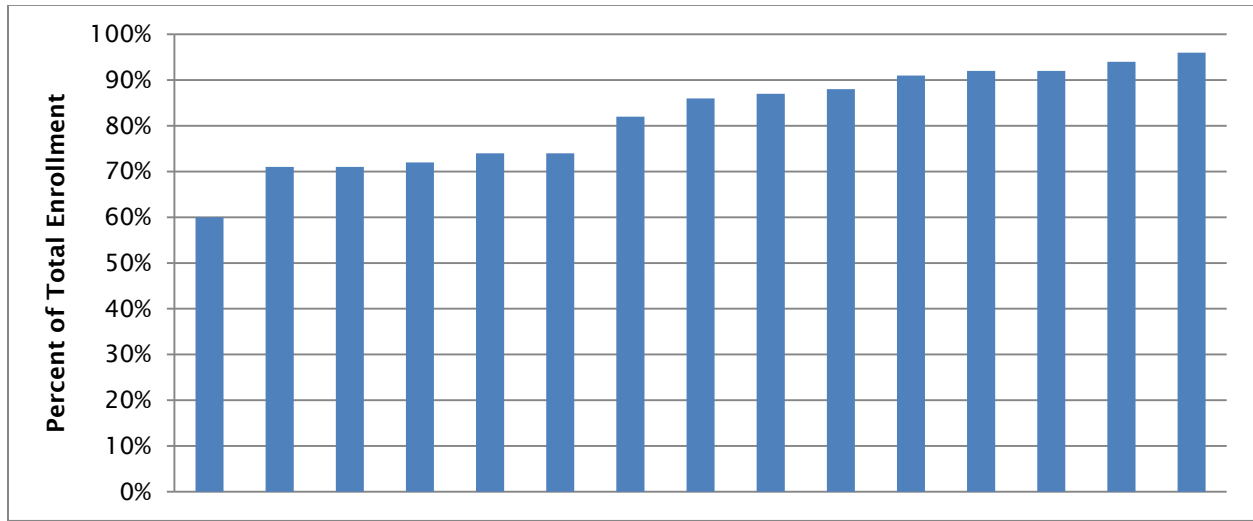
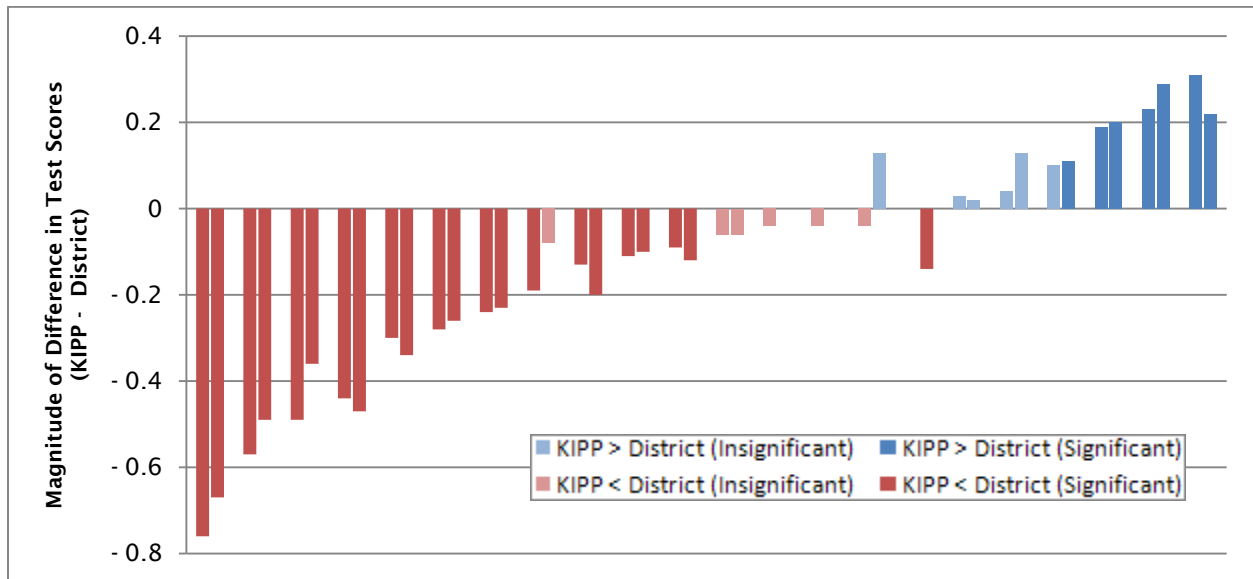


Figure ES.2. Percentage of Students Eligible for FRPL in KIPP Schools



The prior achievement of students entering KIPP schools varies, but KIPP schools most often enroll students whose average fourth-grade achievement is lower than the districtwide average. Figure ES.3 shows whether KIPP students were higher or lower achieving than other students in the district, separately for each KIPP school included in the study. In cases in which the bar—which reflects the difference in baseline test scores between KIPP and district students—is above zero, this implies that KIPP students were higher achieving than district students on average; when the bar is below zero, KIPP students were lower achieving than district students. The size of each bar represents the magnitude of the difference.

Figure ES.3. Distribution of Baseline Math and Reading Differences: KIPP vs. District



Notes: The bars on the left represent baseline differences in math scores, and the bars on the right represent differences in reading scores. The dark-red and dark-blue bars indicate differences from the district population that are considered statistically significant at the five percent level. Test scores were converted to standardized z-scores to allow comparisons across sites.

We also examined whether students leave KIPP schools prior to completion at unusually high rates. **Cumulative rates of attrition vary widely in different KIPP schools, but we did not find systematically higher (or lower) levels of attrition among these KIPP middle schools as compared with other schools within their districts.** In approximately one-third of cases, attrition was higher at KIPP than at other district schools by a statistically significant margin; in another third, attrition was lower at KIPP than at other district schools by a statistically significant margin.

Grade repetition rates, by contrast, are consistently elevated at KIPP middle schools as compared to district public schools, particularly in fifth and sixth grades. These differences likely reflect differences in policies toward student promotion rather than differences in student performance or achievement. In particular, the differences likely capture KIPP’s philosophy that students should be promoted to the next grade level only after they have demonstrated mastery of their current grade’s material.

Effects On Test Scores In 22 KIPP Middle Schools

The key outcomes in this matched, longitudinal analysis of the effectiveness of KIPP middle schools are students’ scores on the state-administered assessments in reading and math. Our approach to estimating KIPP’s effect involves examining the achievement trajectories of KIPP students before and after they enter KIPP schools, and comparing those trajectories to the trajectories of other students in their local school districts. Within this general framework, we employed a variety of different comparison groups and statistical models to explore the robustness of results to different assumptions.

Our preferred, benchmark approach uses propensity score matching techniques to identify other district students who, in elementary school, closely resemble the KIPP students in our sample but do not end up attending KIPP schools. After identifying the matched comparison group, our regression models then control statistically for any remaining differences in the characteristics of the two groups, as related to gender, race/ethnicity, poverty status, special education status, limited English proficiency, and baseline achievement (that is, mathematics and reading test scores for two years prior to KIPP entry). We examine the achievement trajectories of the KIPP students and comparison students in each of the first four years after KIPP entry (as KIPP middle schools typically serve four grades, from fifth through eighth).

Our benchmark estimates reflect the effect of having ever enrolled at KIPP, even if a student subsequently withdraws. This approach—under which all students who spend a year in a KIPP school continue to count as part of the “treatment group,” even if they withdraw from KIPP prior to completing eighth grade—is necessary to ensure that impact estimates are not inflated by relying exclusively on students who remain enrolled, given that those who stay are likely to be doing better than those who leave. Although this approach is necessary to avoid an upward bias in impact estimates, it means that the impact estimates implicitly hold KIPP accountable not only for students who stay, but also for students who spend a year at a KIPP school and subsequently leave. If KIPP in fact has positive impacts, then this is a conservative approach that will underestimate the full effect of actually attending KIPP for students who stay.¹

¹ We also estimated the effect of KIPP on students who remain enrolled, by year, but these estimates (reported in the Appendix) are likely to be biased upward.

Our impact estimates suggest two key results across 22 KIPP schools:

Key Finding One: Impacts for large majorities of the 22 KIPP middle schools included in the study are positive in both reading and math in all four years after students enter KIPP schools.

The overall pattern of estimates of the effect of KIPP is consistently positive. This is reflected in Figures ES.4a and ES.4b, which show the proportions of KIPP schools in the study with impacts on math and reading achievement that are positive, negative, and not statistically distinguishable from zero. Within two years after entry, students are experiencing statistically significant, positive impacts in 18 of 22 KIPP schools in math and 15 of 22 KIPP schools in reading. Meanwhile, only two KIPP schools register a significant negative impact in reading in any year of treatment. The single school showing a significant negative impact in mathematics (in year 1) actually reverses that impact by year 3, when it becomes significantly positive.

These findings are consistent across a variety of alternative specifications. Strong positive estimates of KIPP's impact are evident in models using districtwide comparisons (with statistical controls) as well as those using matched comparisons; in models using imputation for missing data as well as those that eschew imputation; and in models that make different assumptions about the test scores of grade repeaters.

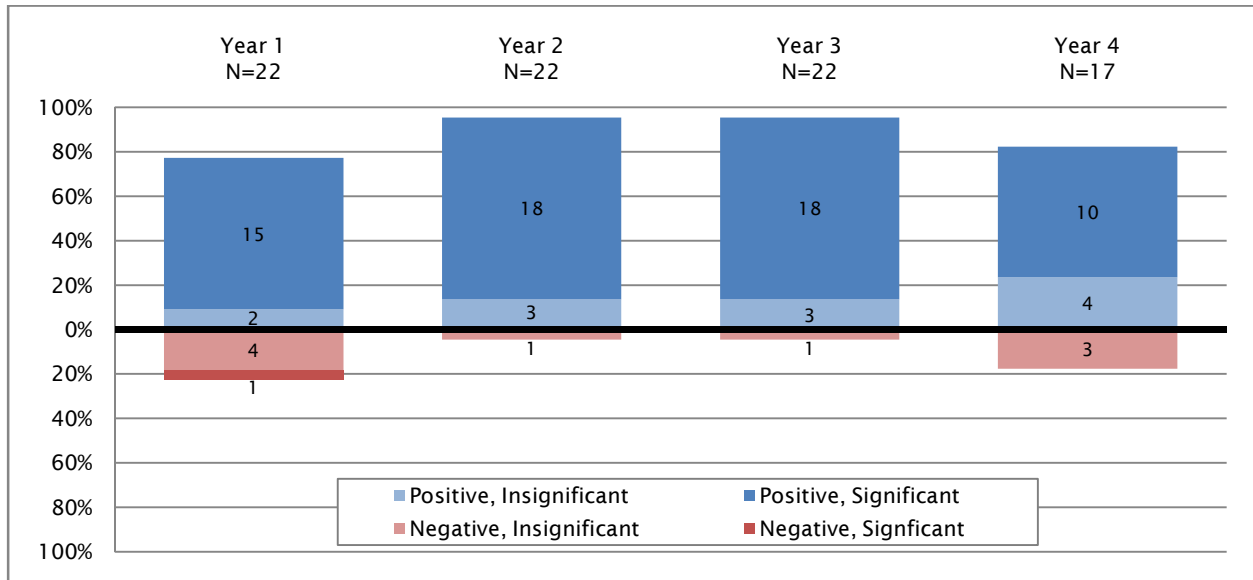
Key Finding Two: The magnitude of KIPP impacts is often substantial.

To illustrate the cumulative effects of KIPP schools, we show estimated impacts after three years in Figures ES.5a and ES.5b. The observed effects are especially large in math (which often shows larger impacts in schoolwide interventions; see, e.g., Hoxby, Murarka, and Kang 2009; Dobbie and Fryer 2009; Abdulkadiroglu et al. 2009). By year three, half of the KIPP schools in our sample are producing math impacts of 0.48 standard deviations or more, equivalent to the effect of moving a student from the 30th percentile to the 48th percentile on a typical test distribution. Compared to national norms during this grade span, a 0.48 effect size after three years represents 1.2 years of accumulated extra growth in mathematics over the three year period (Bloom et al. 2008). For comparison, the black-white test score gap in math is typically estimated as approximately one standard deviation at fourth grade and eighth grade (Bloom et al. 2008). Half of these KIPP schools are producing impacts large enough to cut that gap in half within three years.

Three-year impacts in reading are also large in many KIPP schools, if not as large as the effects in math. Half of the KIPP schools in our sample show three-year reading effects of 0.28 standard deviations or more. An effect size of 0.28 standard deviations represents an estimated 0.9 years of additional instruction, or about one-third of the black-white gap in reading achievement in fourth grade and eighth grade (Bloom et al. 2008).

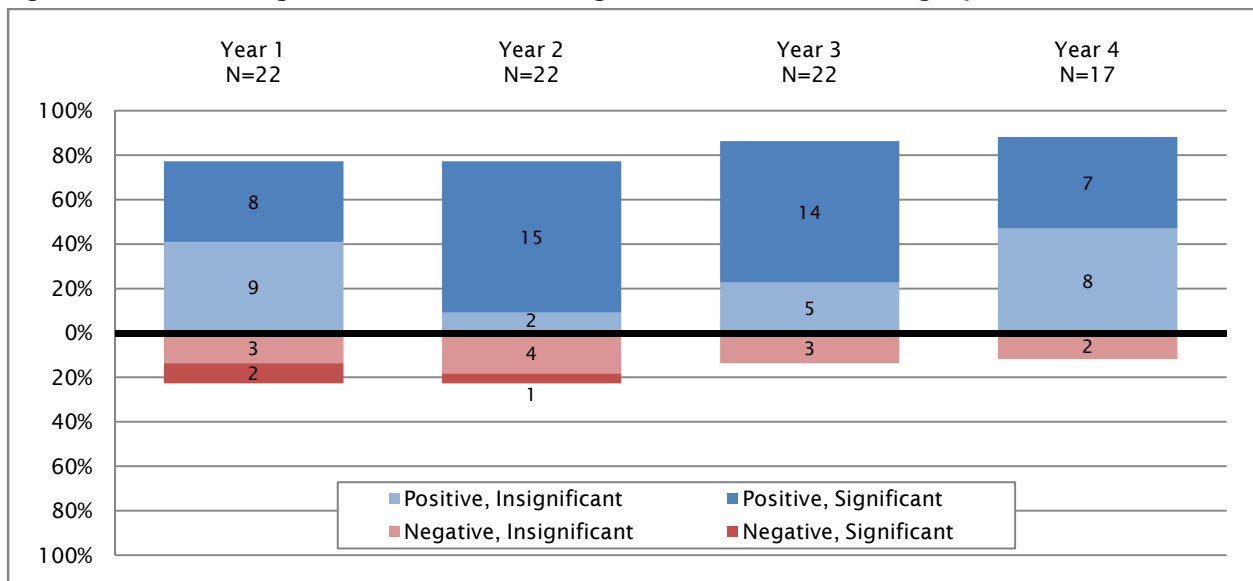
These effect sizes also compare favorably to the effects of other prominent and successful educational interventions. The achievement effects of class-size reduction are often used as a benchmark for other educational interventions. After three years of treatment (grades K-2) in classes one-third smaller than typical, average student gains amounted to 0.20 standard deviations in math and 0.23 standard deviations in reading (U.S. Department of Education, 1998)—less than the effect size achieved in 17 of 22 KIPP schools in math and 13 of 22 KIPP schools in reading.

Figure ES.4a: Percentage of KIPP Schools with Significant Effects in Math, by Year



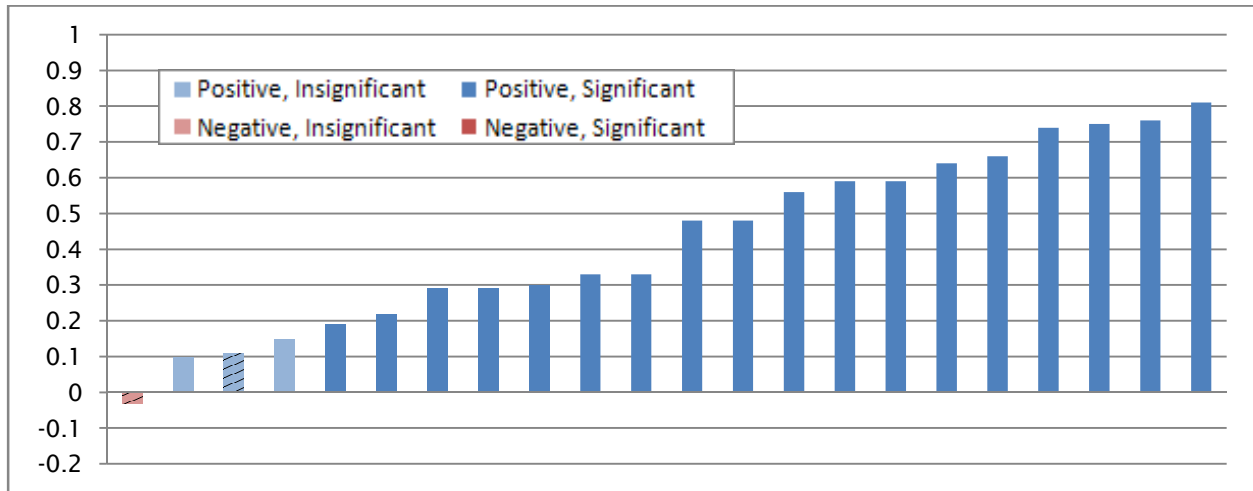
Note: Each bar represents the percentage of schools in the sample where the magnitude of the impact is positive versus negative in a given year. Dark-blue and dark-red colors indicate results that are considered statistically significant at the five percent level (where blue is positive and red is negative).

Figure ES.4b: Percentage of KIPP Schools with Significant Effects in Reading, by Year



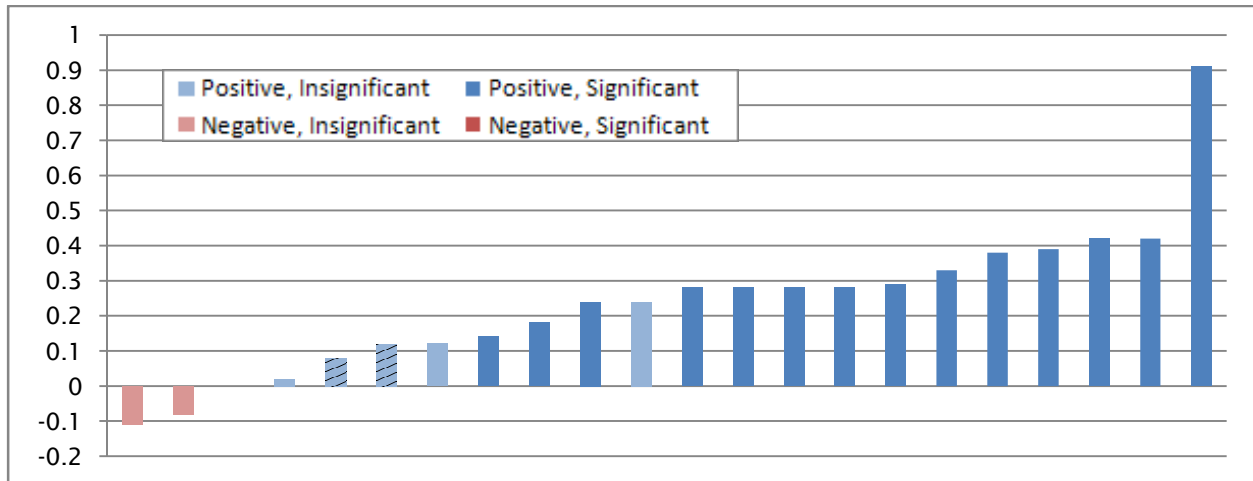
Note: Each bar represents the percentage of schools in the sample where the magnitude of the impact is positive versus negative in a given year. Dark-blue and dark-red colors indicate results that are considered statistically significant at the five percent level (where blue is positive and red is negative).

Figure ES.5a: Distribution of Test Score Effect Sizes After Three Years in Math



Note: Bars with crosshatching represent schools that had their KIPP affiliation removed. Dark-blue and dark-red colors indicate results that are considered statistically significant at the five percent level (where blue is positive and red is negative).

Figure ES.5b: Distribution of Test Score Effect Sizes After Three Years in Reading



Note: Bars with crosshatching represent schools that had their KIPP affiliation removed. Dark-blue and dark-red colors indicate results that are considered statistically significant at the five percent level (where blue is positive and red is negative).

More directly relevant are three recent studies of charter schools that found positive impacts. An experimental study of New York City charter schools estimated annual achievement impacts of 0.09 standard deviations in math and 0.06 standard deviations in reading (Hoxby, Murarka, and Kang 2009). If accumulated over three years, these would amount to 0.27 standard deviations in math and 0.16 standard deviations in reading—substantially less than most of these KIPP schools are producing. The KIPP results more closely resemble the results from studies of Boston charter schools and the Harlem Children’s Zone’s Promise Academy charter schools—results that have been viewed by some observers as spectacular (see, e.g., Brooks 2009). Charter middle schools in Boston are estimated as producing annual achievement impacts of 0.09 in reading and 0.18 in math (Abdulkadiroglu et al 2009), which, if accumulated over three years, would put them squarely in the middle of the KIPP performance distribution (using a method that is less conservative than ours,

and likely to produce larger impact estimates). The Promise Academy middle school produced three-year impacts estimated as 0.27 standard deviations in reading—near the middle of the KIPP distribution—and 0.77 standard deviations in math—near the high end of the KIPP distribution (Dobbie and Fryer 2009).

Finally, as noted above, our estimates are likely to underestimate KIPP’s full impact on students who remain enrolled, because we use a conservative approach in which students who have withdrawn from KIPP schools after a year enrolled remain classified in the KIPP “treatment group.”

Additional Findings

To better understand the effects of KIPP on student achievement, we conducted some preliminary analysis of how these effects vary across years after enrolling in KIPP, for KIPP schools that have closed, and across student subgroups. Key findings that have emerged from this preliminary analysis include:

- In most KIPP schools, cumulative positive effects increase for at least the first three years after KIPP entry. In math, 18 of 22 KIPP schools show larger cumulative effects in year 3 than year 1, and in reading, 19 of 22 show larger cumulative effects in year 3 than year 1. But the largest single-year impacts are often in the first year, especially in math.
- Of only three schools that never demonstrate a statistically significant positive impact in either mathematics or reading in any year, two are schools from which the KIPP Foundation withdrew the KIPP affiliation. Both schools subsequently closed.
- We find no evidence that KIPP impacts are higher or lower for specific subgroups of students. We examined impacts for the following subgroups: higher versus lower-performing students on test scores at baseline; LEP students; male students; black students and black male students; and Hispanic students and Hispanic male students. We did not find clear patterns suggesting that KIPP impacts for any of these subgroups differed systematically from average impacts for all KIPP students.

Next Steps

The ultimate goal of this evaluation, when completed, is to produce the best possible estimate of the average impact of KIPP middle schools on their students’ academic outcomes. Achieving this goal requires using multiple analytic methods in a coordinated way to produce an estimate that has the greatest possible causal rigor while also representing the largest possible sample of operating KIPP middle schools (that is, to produce an estimate that is both internally and externally valid). The overall study was designed to achieve this aim by complementing the strong causal rigor of randomized experiments—also known as randomized controlled trials, or RCTs—based on the admissions lotteries conducted in a subset of KIPP schools, with the greater comprehensiveness that can be achieved using nonexperimental methods in the full population of KIPP middle schools. In future reports, we will expand our analyses by estimating lottery-based experimental impacts and using them to attempt to validate the nonexperimental methods.

In addition to incorporating experimental impact estimates, future reports will also expand the scope of the evaluation by including a larger population of KIPP middle schools; incorporating additional student outcomes beyond state test scores; and exploring aspects of the operation of KIPP schools that may be related to producing larger impacts on students.