Injecting Successful Charter School Strategies into Traditional Public Schools:
Evidence from Houston[*]

Roland G. Fryer, Jr.
Harvard University

October 2013

Abstract

Starting in the 2010-2011 school year, we implemented five strategies gleaned from practices in successful charter schools – increased instructional time, a more rigorous approach to building human capital of teachers and administrators, high-dosage tutoring, frequent use of data to inform instruction, and a culture of high expectations – in twenty of the lowest performing schools in Houston, Texas. We show that the average impact of these changes on student achievement is 0.161 standard deviations in math and 0.021 standard deviations in reading, which is comparable to reported impacts of attending a high-performing charter school.

"If an unfriendly foreign power had attempted to impose on America the mediocre education performance that exists today, we might well have viewed it as an act of war." A Nation at Risk (1983)

American public schools are in need of reform. Data from the National Assessment of Educational Progress (NAEP) – a set of assessments administered every two years to a nationally representative group of fourth, eighth, and twelfth graders – reveal that 33 percent of eighth graders are proficient in reading and 34 percent are proficient in math. Data for fourth and twelfth graders are similar. In 2010, roughly one in five high school graduates did not score high enough on the United States Army's Armed Services Vocational Aptitude Battery (ASVAB) to meet the minimum standard necessary to enlist in the Army (Theokas 2010). According to a Center for Education Policy report, forty-eight percent of American schools did not meet the standards set out by the No Child Left Behind Act of 2001 (Usher 2011) – ranging from 11% in Wisconsin to 89% in Florida.[1] There are approximately 5,000 "chronically low performing" schools in America – 5% of all public elementary and secondary education in the country – roughly half are in large cities, one-third are in rural areas, and the remaining are in suburbs (Duncan 2009).[2]

There has been no paucity of effort aimed at increasing achievement and closing racial achievement gaps in the past few decades: lowering class size, increasing spending, and providing incentives for teachers are only a few of the dozens of ambitious policy prescriptions in education reform.[3] Moreover, school districts have taken a variety of targeted approaches to cope with "chronically low performing" schools. Between 2001 and 2006, Chicago closed 44 schools and reassigned students to other schools. In New York City, the city closed 91 public schools between 2002 and 2010 – converting most of them to charter schools. In November 2005, 102 of the lowest performing public schools in New

---

[1] There have been many other attempts to close the achievement gap, none of which significantly or systematically reduce racial disparities in educational achievement (see Fryer 2011a, Jacob and Ludwig 2008).
[2] A school is designated "chronically low performing" if it fails to make "adequate yearly progress" for three consecutive years. High schools can also be deemed chronically low performing if their graduation rate is lower than sixty percent for three consecutive years.

Orleans were turned over to the Recovery School District (RSD), which is operated at the state level; some of these schools are currently run directly by the RSD while others are run by charter school operators. Tennessee created the Tennessee Achievement School District, which takes control of the lowest-performing schools across the state from the home district and centralizes the governance for these schools under this school turn-around entity. Despite these reforms to increase student achievement, measures of academic success have been largely constant over the past thirty years (Fryer 2011a). This lack of progress has caused some to argue that schools alone cannot increase achievement among the poor (Coleman 1966, Ravitch 2010).

Yet, due to new evidence on the efficacy of certain charter schools demonstrating that some combination of school policies and procedures can significantly increase achievement among poor black and Hispanic students, there may be room for optimism. Using data from the Promise Academy in the Harlem Children's Zone – a 97-block area in central Harlem that provides myriad social programs along with achievement-driven charter schools – Dobbie and Fryer (2011) show that middle school students gain 0.229 standard deviations (hereafter σ) in math per year and 0.047σ in reading on state standardized test. Thus, after four years, students in these schools have erased the achievement gap in math (relative to the average white student in NYC) and halved it in reading. Perhaps more important, Dobbie and Fryer (2013) demonstrate that the same sample of students are significantly more likely to attend college, less likely to be pregnant (girls) and less likely to be incarcerated (boys). Dobbie and Fryer (2011, 2013) provide evidence that it is the school policies – not social programs – that are responsible for the achievement gains, though one cannot rule out important interactions. Consistent with these findings, others have shown similar results with larger and more diverse samples of charter schools that are not coupled with community programs (Abdulkadiroglu et al. 2011, Angrist et al. 2010, Angrist et al. 2013).

A strategy to increase achievement and combat the racial achievement gap, yet to be tested, is to infuse the school policies exemplified in the most successful charter schools into traditional public schools with their standard hierarchies and bureaucracy, local politics, school boards, and collective bargaining agreements. Theoretically, introducing school policies and procedures typified by successful charter schools in traditional public schools could have one of three effects. If the policies most correlated with charter school

effectiveness are general lessons about the education production function, then these strategies may yield significant increases in student achievement. If, however, a large part of the success of the achievement-increasing charter schools we emulate can be attributed to selective attrition of unmotivated students out of these schools, the tendency of highly involved parents to enroll their children in charter school lotteries, or school policies that cannot be easily replicated in a traditional public school, then an attempt to create public schools in this image is likely futile.[4] Third, some argue that major reform efforts are often more disruptive than helpful, can lower teacher morale, or might be viewed by students as punishment for past performance, any of which may have a negative impact on student achievement (Campbell, Harvey, and Hill 2000). Which one of the above effects will dominate is unknown. The estimates in this paper may combine elements from these and other channels.

Starting in the 2010-2011 school year, we implemented five correlates of effective charter schools described in Dobbie and Fryer (2013) – increased time, better human capital, more student-level differentiation, frequent use of data to alter the pace of classroom instruction, and a culture of high expectations – in twenty of the lowest performing schools (containing more than 16,000 students) in Houston, Texas. Houston is the largest school district in Texas and the seventh largest in the country. It is a microcosm of public systems across the country—large, ethnically and linguistically diverse, governed by a school board, and boasts a substantial achievement gap between rich and poor, black and white. **The racial achievement gap in Houston elementary schools is 0.6σ – or roughly 8 months of school behind – and 0.8σ in secondary schools (e.g. 10 months behind).**

To increase time on task, the school day was lengthened one hour and the school year was lengthened ten days in the nine secondary schools. This is twenty-one percent more time in school than students in these schools obtained in the year pre-treatment and roughly the same as achievement-increasing charter schools in New York City.[5] In addition, students were strongly encouraged and even incentivized to attend classes on Saturday. In the eleven elementary schools, the length of the day and the year were not changed, but non-

---

[4] Throughout the text, I depart from custom by using the terms "we," "our," and so on. Although this is a sole-authored work, it took a large team of people to implement the experiments. Using "I" seems disingenuous.

[5] Using the data set constructed by Dobbie and Fryer (2011b), we label a charter school "successful" if its treatment effect on combined math and reading achievement is above the median in the sample, according to their non-experimental estimates.

instructional activities (e.g. twenty minute bathroom breaks) were significantly reduced. In an effort to significantly alter the human capital, nineteen out of twenty principals were removed before the experiment began. To enhance student-level differentiation, we supplied all fourth, sixth and ninth graders with a math tutor and provided extra reading or math instruction to students in other grades who had previously performed below grade level. This model was adapted from the MATCH school in Boston – a charter school that largely adheres to the methods described in Dobbie and Fryer (2013). In order to help teachers use interim data on student performance to guide and inform instructional practice, we required schools to administer interim assessments every three to four weeks and provided schools with three cumulative benchmarks assessments, as well as assistance in analyzing and presenting student performance on these assessments. Finally, to instill a culture of high expectations and college access, we started by setting clear expectations for school leadership. Schools were provided with a rubric for the school and classroom environment and were expected to implement school-parent-student contracts. Specific student performance goals were set for each school and the principal was held accountable and provided financial incentives based on these goals.

Such invasive changes were possible, in part, because twelve of the twenty schools (nine secondary and three elementary) were either "chronically low performing" or on the verge of being labeled as such and taken over by the state of Texas. Thus, despite our best efforts, random assignment was not a feasible option for these schools. To round out our sample of twenty schools and provide a way to choose between alternative non-experimental empirical specifications, we randomly selected eight additional elementary schools from sixteen low – but not chronically low – performing elementary schools.

In the sample of sixteen elementary schools in which treatment and control were chosen via matched-pair random assignment, providing estimates of the impact of injecting charter school practices in traditional public schools is straightforward and transparent. In the remaining set of schools, we use four separate statistical approaches to adjust for pre-intervention differences between treatment and comparison school attendees. Treatment is defined as being zoned to attend an Apollo school and "comparison school" attendees are all other students in HISD not zoned for Apollo. We begin by using district administrative data on student characteristics, most importantly previous year achievement, to fit least squares models. This approach may not account for important student level unobservables,

potential mean reversion, or measurement error in previous year test score, so we also estimate a difference-in-differences specification that can partially account for these concerns. Houston has a widely used choice program that allows students to attend any public school they want, subject to capacity constraints, which introduces the potential for selection into (or out of) treatment. Following Cullen et al. (2005), our third and fourth empirical models instrument for a student's enrollment in a treatment school with an indicator for whether or not they are zoned to attend a treatment school.

All statistical approaches lead to the same qualitative conclusions. Injecting strategies and best practices from achievement-increasing charter schools into low performing traditional public schools can significantly increase student achievement. Students in treatment elementary schools gain almost 0.2σ in math per year, relative to comparison samples. **Taken at face value, this is enough to eliminate the racial achievement gap in Houston elementary schools in three years. Students in treatment secondary schools gain 0.140σ per year in math, decreasing the gap by one-half over the length of the demonstration project.** Figure 1 demonstrates these results. Surprisingly, the impacts on reading for both elementary and secondary schools are small and positive, but statistically indistinguishable from zero.

Importantly, in the grade/subject areas in which we implemented all five policies described in Dobbie and Fryer (2013) – fourth, sixth, and ninth grade math – the increase in student achievement is substantially larger. Relative to students who attended comparison schools, fourth graders in treatment schools scored 0.242σ (0.104) higher in math; sixth and ninth grade math scores increased 0.320σ (0.073) relative to students in comparison schools.

The results are robust across identification strategies, model specification, construction of comparison schools, alternative student assessments, sample attrition, sample re-weighting to account for potential negative selection into treatment. Moreover, an almost identical (non-random assignment) experiment in Denver, Colorado and data from the Academy of Urban School Leadership (AUSL) – which uses the five best practices in Dobbie and Fryer (2013) as a core strategy to turnaround chronically low performing schools in Chicago – yield similar point estimates.

We conclude our main statistical analysis by estimating heterogeneous treatment effects on test scores across a variety of pre-determined subsamples, and investigating the impact of treatment on student attendance. Surprisingly, the treatment was most effective in

high school and elementary school. All other subsamples of the data yield consistent impacts, though there is evidence that Hispanic students gained significantly more than black students. For instance, in secondary schools, the impact of treatment on black students is 0.022σ (0.031) and 0.160σ (0.035) for Hispanic students – the p-value on the difference is 0.000. Put differently, the turnaround strategy implemented in partnership with HISD and EdLabs worked for all racial groups, but worked particularly well for Hispanic students. Finally, the treatment administered in Apollo schools was particularly effective for economically disadvantaged students. In secondary schools, the impact of treatment on economically disadvantaged students is 0.114σ (0.035) versus a treatment effect of 0.028σ (0.061) for their more-advantaged peers - the p-value on the difference is 0.052. The same is true in elementary schools, where the impact of treatment on economically disadvantaged students is 0.208σ (0.104) versus a treatment effect of 0.033σ (0.117) for their more-advantaged peers, with a 0.060 p-value on the difference.

Treatment effects on attendance in elementary school were small and statistically insignificant, potentially due to the high baseline attendance rate (97%). The impact of the treatment on attendance in the secondary schools was approximately half a percentage point, per year (1.5 percentage points total over the length of the experiment).

Figure 2 puts the magnitude of these estimates in perspective. Jacob and Ludwig (2008), in a survey of programs and policies designed to increase achievement among poor children, report that only three reforms pass a simple cost-benefit analysis: lowering class size, bonuses for teachers for teaching in hard-to-staff schools, and early childhood programs. The effect of lowering class size from 24 to 16 students per teacher is approximately 0.073σ per year (e.g. 0.22σ over three years) on combined math and reading scores (Krueger 1999). The effect of Teach for America, one attempt to bring more skilled teachers into poor performing schools, is 0.15σ in math and 0.03σ in reading (Decker et al. 2004). The effect of Head Start is 0.147σ (.103) in applied problems and 0.319σ (.147) in letter identification on the Woodcock-Johnson exam, but the effects on test scores fade in elementary school (Currie and Thomas 1995, Ludwig and Phillips 2007).

All these effect sizes are a fraction of the impact of our fully-loaded treatment that includes tutoring. The effects closest to the ones reported here are from a series of papers on achievement-increasing charter schools in which the impacts range from 0.229σ to 0.364σ in

math and 0.120σ to 0.265σ in reading (Abdulkadiroglu et al. 2011, Angrist et al. 2010, Curto and Fryer 2012).

The difference in achievement effects between math and reading, while striking, is consistent with previous work on the efficacy of charter schools and other educational interventions. Abdulkadiroglu et al. (2011) and Angrist et al. (2010) find that the treatment effect of attending an oversubscribed charter school is four times as large for math as ELA. Dobbie and Fryer (2011a) demonstrate effects that are almost 5 times as large in middle school and 1.6 times as large in elementary school, in favor of math. In larger samples, Hoxby and Murarka (2009) reports an effect size 2.5 times as large in New York City charters, and Gleason et al. (2010) show that an average urban charter school increases math scores by 0.16σ with statistically zero effect on reading.

There are many theories that may explain the disparity in treatment effects by subject area.[6] Research in developmental psychology has suggested that the critical period for language development occurs early in life, while the critical period for developing higher cognitive functions extends into adolescence (Hopkins and Bracht 1975, Newport 1990, Pinker 1994, Nelson 2000, Knudsen et al. 2006). This theory seems inconsistent with the fact that the elementary school reading estimates are similar in magnitude to the secondary school estimates. Another leading theory posits that reading scores are influenced by the language spoken when students are outside of the classroom (Charity et al. 2004, Rickford 1999). Charity et al. (2004) argue that if students speak non-standard English at home and in their communities, increasing reading scores might be especially difficult. This theory is consistent with our findings and could explain why students at an urban boarding school make similar progress on ELA and math (Curto and Fryer 2012).

---

[6] It is important to remember that our largest treatment effects were in grades with two-on-one tutoring in math – it is worth considering whether similar interventions for reading could have a sizeable impact on reading outcomes. This experiment is in progress in twenty low-performing middle schools in NYC.
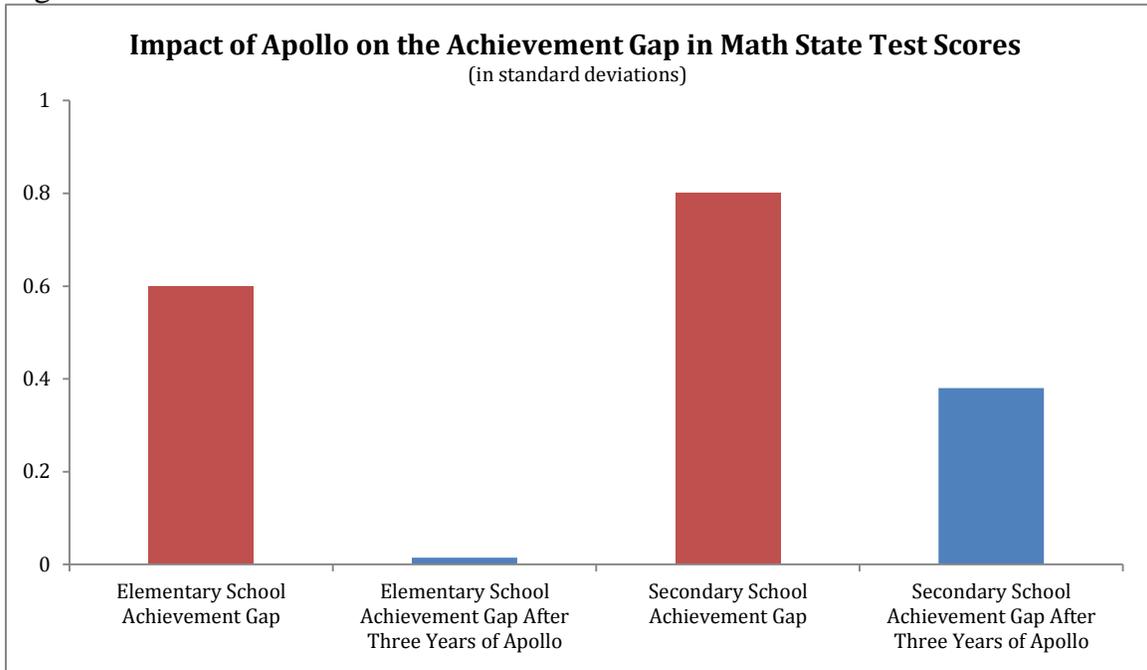
Figure 1

Impact of Apollo on the Achievement Gap in Math State Test Scores
(in standard deviations)

Figure 2

Treatment Effects in Context