

Broader,
BOLDER
Approach
to Education

April 22, 2013

Market-oriented education reforms' rhetoric trumps reality

The impacts of test-based teacher evaluations, school closures, and increased charter school access on student outcomes in Chicago, New York City, and Washington, D.C.

BY ELAINE WEISS AND DON LONG

Table of contents

Executive summary 3

The “reformers” and their claims 7

A brief synopsis of the evidence regarding core components of the reform agenda 10

 Test-based accountability systems 10

 School closures 12

 Increasing the reach of charters 13

 Turnover and instability 15

 A “no excuses” take on the influence of poverty 16

The reforms have not strengthened school systems in the three cities 16

 Washington, D.C.: Firings and closures increased churn 19

 New York City: Merit pay failed while school reshuffling produced mixed results 26

 Chicago: Neighborhood school replacement increased instability but not performance 33

 The bottom line: Highlighted reforms in the three cities produced upheaval, overshadowing smaller promising programs 39

The reforms have not improved student outcomes 41

 Washington, D.C.: Race and income-based achievement gaps grew 42

 New York City: Nine years of market-based reforms failed to improve test scores or narrow achievement gaps 49

 Chicago: As achievement gaps grew, African American students fell further behind, and college readiness remained disappointing 54

The reforms have not saved money 62

 Washington, D.C.: Increased spending without public and parental input produced few benefits 62

 New York City: Reforms hinged on likely unsustainable spending increases 64

 Chicago: Spending increases mirrored those of other large cities, but donors’ policy influence outstripped their funding 65

 The bottom line: More spending, bolstered by private donations, depressed democratic input without helping students 67

Conclusion 68

About the authors 73

Appendix tables 73

Endnotes 76

References 83

Executive summary

Top-down pressure from federal education policies such as Race to the Top and No Child Left Behind, combined with organized advocacy efforts, is making a popular set of market-oriented education “reforms” look more like the new status quo than real reform. Reformers assert that test-based teacher evaluation, increased school “choice” through expanded access to charter schools, and the closure of “failing” and underenrolled schools will boost falling student achievement and narrow longstanding race- and income-based achievement gaps. The 2010 documentary *Waiting for Superman* presented these policies as sure fixes for education woes closely correlated with child poverty.

This report from the Broader, Bolder Approach to Education examines these assertions by assessing the impacts of these reforms in three large urban school districts: Washington, D.C., New York City, and Chicago. These districts were chosen for study because all enjoyed the benefit of mayoral control, produce reliable district-level test score data from the National Assessment of Educational Progress (NAEP), and were led by vocal proponents who implemented versions of this reform agenda. Indeed, former reform leaders in all three cities have become high-profile national proponents who disseminate the agenda across multiple districts and states.

The report finds that the reforms delivered few benefits and in some cases harmed the students they purport to help. It also identifies a set of largely neglected policies with real promise to weaken the poverty-education link, if they receive some of the attention and resources now targeted to the touted reforms. Specifically the report finds:

- Test scores increased less, and achievement gaps grew more, in “reform” cities than in other urban districts.
- Reported successes for targeted students evaporated upon closer examination.
- Test-based accountability prompted churn that thinned the ranks of experienced teachers, but not necessarily bad teachers.
- School closures did not send students to better schools or save school districts money.
- Charter schools further disrupted the districts while providing mixed benefits, particularly for the highest-needs students.
- Emphasis on the widely touted market-oriented reforms drew attention and resources from initiatives with greater promise.
- The reforms missed a critical factor driving achievement gaps: the influence of poverty on academic performance.
- Real, sustained change requires strategies that are more realistic, patient, and multipronged.

Test scores increased less, and achievement gaps grew more, in “reform” cities than in other urban districts. Leaders promised that the reforms would raise test scores, especially those of minority and low-income students, and close race- and income-based achievement gaps. Analysis of the most reliable, comparable data—National Assessment of Educational Progress scores—shows that the rhetoric did not match the reality. While test scores increased and achievement gaps shrank in most large urban districts over the past decade, scores stagnated for low-income and minority students and/or achievement gaps widened in the reform cities.

- Between 2005 and 2011, in large, urban districts, Hispanic eighth-graders gained an average six points in reading (from 243 to 249), black eighth-graders gained 5 points (from 240 to 245), and white eighth-graders gained 3 points (from 270 to 273) (NCES 2005b, 2011b). In District of Columbia Public Schools, however, Hispanic eighth-graders' scores fell 15 points (from 247 to 232), black eighth-graders' scores fell 2 points (from 233 to 231), and white eighth-graders' scores fell 13 points (from 303 to 290) (Levy 2012c).

Reported successes for targeted students evaporated upon closer examination. Reformers in all three cities claimed that they had boosted student achievement and closed achievement gaps. But when state test scores were recalibrated to make standards consistent, compared with NAEP scores, and disaggregated by race and income, gains vanished or turned out to have accrued only to white and high-income students.

- New York City Mayor Michael Bloomberg claimed to halve the white/Asian to black/Latino achievement gap in city schools from 2003 to 2011, but scores on state-administered tests, averaged across fourth and eighth grades in reading and math, show that the achievement gap had stagnated; it was 26.2 percentage points in 2003 and 25.8 percentage points in 2011 (a 0.01 standard deviation change). Columbia University professor Aaron Pallas, who calculated the 1 percent reduction, noted, “The mayor has thus overstated the cut in the achievement gap by a factor of 50” (Pallas 2012b).

Test-based accountability prompted churn that thinned the ranks of experienced teachers, but not necessarily bad teachers. Reformers said that using student test scores to evaluate teachers, to reward and fire teachers, and to target schools for closure would improve the quality of teachers in low-income schools. But instead, narrow, unreliable metrics turned off great teachers, increased churn, and drained experience from teacher pools, with no boost to student achievement.

- District of Columbia Public Schools' IMPACT system, which bases teacher evaluations (and dismissals) heavily on test scores, is associated with higher teacher turnover. The share of DCPS teachers leaving after one year increased from 15.3 percent in 2001–2007 (before IMPACT began in 2009) to 19.3 percent in 2008–2012; the share leaving after two years increased from 27.8 percent to 33.2 percent; the share leaving after three years increased from 37.5 percent to 42.7 percent; and after four years fully half (52.1 percent) of teachers left the system, up from 45.3 percent (Levy 2012g). Few teachers reach “experienced” status, generally considered at least five years and, by some experts, seven years or more.

School closures did not send students to better schools or save school districts money. Reformers closed schools deemed “failing” so students could transfer to “better-performing” schools. But most students whose schools were closed went to other schools with even lower test scores, and the disruption (some had multiple moves) was exacerbated by longer commutes and spikes in gang violence as established lines were crossed.

- Although Duncan closed Chicago public schools deemed “underperforming” in order to move students to better schools, the closings had almost no effect on student achievement because almost all displaced elementary school students transferred from one low-performing school to another, according to a study of 18 schools closed between 2001 and 2006 (de la Torre and Gwynne 2009). Only the 6 percent who moved to better schools with greater resources had improved outcomes.

Charter schools further disrupted the districts while providing mixed benefits, particularly for the highest-needs students. Reformers say charters offer better options and outcomes for students in “failing” public schools. But charter outcomes in these cities and across the country are uneven. Charters serve fewer of the highest-needs students than do regular public schools and can disrupt school districts logistically and financially. High-performing charters may also spend more per student than regular public schools.

- The Chicago Public School (CPS) system uses its own “value-added” metric to measure school performance, with schools scoring lower on the distribution identified for closure. By this measure, if students in the types of schools most likely to be closed moved to charters, they would move to lower-performing schools. Specifically, students who moved from high-poverty regular public schools at the 47th percentile in performance would go to charter schools at the 40th percentile, and those moving from intensely segregated schools at the 43rd percentile would end up in charters at the 33rd percentile. Random-lottery enrollment schools, which, unlike charter schools, do not “select out” students via a challenging application process, outperform their demographically comparable charter counterparts: Students who moved to charters would drop from the 52nd percentile to the 40th percentile (Caref et al. 2012).

Emphasis on the widely touted market-oriented reforms drew attention and resources from initiatives with greater promise. Less-publicized strategies for boosting student achievement were piloted in these cities but not widely replicated or expanded to scale because leaders and funders focused on the market-oriented reforms. These promising but overlooked reforms are more multifaceted and holistic than reforms that seek quick fixes and rely on narrow, unreliable metrics.

- Between 2003 and 2005, New York Mayor Michael Bloomberg created 100 new, small, and reportedly academically rigorous secondary schools across the city’s five boroughs (Goldstein 2010). These schools conferred many benefits for at-risk students, including a reported 10 percentage-point increase in the share of ninth-grade students on track to graduate (from 48.5 percent to 58.5 percent) and a 7.6 percentage-point increase in Regents exam scores indicating college readiness in English. Among the added resources small schools delivered were community partners providing students with relevant learning opportunities inside and outside the classroom and providing school faculty with additional staffing support and start-up resources. For example, a law firm offered internships to students of the Urban Assembly School for Law & Justice and helped staff integrate real-world examples into the curriculum. Another partner helped the Urban Assembly New York Harbor School grow one of the largest oyster beds in the region (Bloom and Unterman 2012, 2). Extra staff and other resources also enabled the schools to nurture the strong personal teacher-student interactions that likely contributed to the academic gains.
- Duncan worked to improve the college readiness of low-income and minority students in Chicago by increasing their access to AP courses, providing high schools with college counselors, and holding principals accountable for ensuring that students applied for financial aid. Scholarship aid nearly doubled in one year (2011–2012). Budget cuts have since removed counselors from almost all schools.
- Just before Rhee became chancellor, DCPS expanded its high-quality, full-day prekindergarten program to serve middle- and upper-income as well as low-income 3- and 4-year-olds whose parents requested seats for them. The district also adopted a holistic, hands-on curriculum designed to nurture healthy development of all early childhood domains: cognitive, emotional, physical, and behavioral. Data showed that third-graders who had participated in

the program had higher test scores than their nonparticipating peers. However, Rhee implemented a policy that required parents to enter a lottery to get a spot for the coming year, limiting access. And pre-K is not even on the agenda of StudentsFirst, the well-funded education reform organization founded by Rhee to promote the reforms she says worked in Washington, D.C.

The reforms missed a critical factor in achievement gaps: the influence of poverty on academic performance. In all three cities, a narrow focus on market-oriented policies diverted attention from the need to address socioeconomic factors that impede learning. In 2010, student eligibility rates for free- and reduced-price meals were 67 percent in Washington, D.C., 72 percent in New York City, and 77 percent in Chicago (NAF 2009–2010). Clearly child poverty has been a significant factor contributing to low student test scores and graduation rates in these three cities. Failing to provide supports that alleviate impediments posed by poverty ensures continued low student test scores and graduation rates, and large gaps between average test scores of white and affluent students and test scores of minority and low-income students.

Real, sustained change requires strategies that are realistic, patient, and multipronged. In each city, the initiatives showing more promise than the touted reforms demonstrate that achievement gaps can only be closed when the opportunity gaps driving them are addressed. The hands-on experiences and consistent, intensive teacher-student relationships of New York City's small schools must replace reform's test preparation, novice teachers, and churn. Heavy reliance on college- and career-readiness test metrics should give way to CPS-style college- and career-readiness supports: helping students choose courses and schools, access AP classes, and match their skills with career goals; and holding schools accountable for scholarship applications. DCPS's high-quality prekindergarten program, which is designed to nurture all aspects of children's development, should serve as a model for all cities and students, not be sidelined in "reform" agendas.

Districts that recognize the impact of poverty and address it head-on find the greatest success. Though it is higher-income, Montgomery County, Md., serves a student body that is as ethnically diverse as any of these urban districts, and has a large and growing share of low-income students. In contrast to the reformers, however, Montgomery County Public Schools (MCPS) Superintendent Joshua Starr staunchly opposes using test scores to evaluate teachers, instead employing a peer-assisted review system that focuses on teacher support, development, and collaboration (Strauss 2012). MCPS has no charter schools. Rather, it channels extra resources, including targeted professional development for qualified teachers, the smallest classrooms, and intensive literacy, to the neediest schools. It has developed a holistic, creative curriculum to nurture in-depth, critical thinking. This includes art, music, and physical education teachers in every school. MCPS also leverages the county's mixed-use housing policies to integrate schools (Schwartz 2012). Finally, it employs high-quality prekindergarten, health clinics, and afterschool enrichment to further close income-based opportunity gaps. As Starr highlights, all of this has produced some of the highest test scores among minority and low-income students of any district, smaller and shrinking achievement gaps, and high school graduation and college attendance rates that are the envy of the country (Starr 2012).

Every school district has unique needs and resources. But providing *all* students with the enriching experiences that already help *high-income* students thrive would represent a big step forward, and away from narrow reforms that miss the mark.

The “reformers” and their claims

A specific set of policies labeled “reforms” have dominated education policy discussions in recent years. Self-identified reformers such as former New York City schools chancellor Joel Klein, former Washington, D.C., school chancellor Michelle Rhee, and Secretary of Education Arne Duncan, emphasize three reforms, in particular, as the answer to the declining student achievement and widening race- and income-based achievement gaps that they assert represent the state of American education:

- 1) The use of student test score data to make decisions about hiring, firing, and rewarding teachers and principals.
- 2) The use of student test score data to target “failing” schools for “turnaround” and/or closure.
- 3) The expansion of student and parent choice about schools, in the form of increased access to charter schools (often as replacements for neighborhood schools closed for underenrollment).

Although student achievement is in fact increasing, and race-based achievement gaps have been shrinking, these market-oriented reforms have become embedded in federal policies and organized advocacy efforts to change how school systems and districts operate. Under the Obama Administration’s signature Race to the Top (RTTT) initiative, states must establish plans to enact these reforms in order to attain grants. States seeking waivers from some of the requirements of the federal No Child Left Behind (NCLB) program must undertake “dramatic efforts” to “turn around” the lowest-performing schools, and their plans, too, usually rely on these reforms.

These same reforms represent the basic policy platforms of a number of well-funded education reform organizations, including StudentsFirst (founded by Rhee), Stand for Children, 50CAN, and Democrats for Education Reform (DFER), all of which have state affiliates across the country. Combined, these efforts represent a concerted push to establish this set of reforms as the new norm. In its mission statement, StudentsFirst states its intent to “build a national movement” and “demand that legislators, courts, district administrators, and school boards create and enforce policies” consistent with these reforms (StudentsFirst 2012). And its recent “State Policy Report Cards” ranked every

This paper compares and contrasts reformers’ promises and claims of improvement with actual outcomes. Use of reformer claims as benchmarks does not indicate and should not be interpreted as endorsement or adoption by the Broader Bolder Approach to Education (BBA) of those benchmarks. For example, reformers assert that using student test scores to evaluate teachers and take action based on those evaluations can improve student outcomes, and even turn around struggling schools in relatively short periods of time. BBA does not believe that this is true, nor that truly troubled schools can “turn around” in a sustained manner in a short period or as the result of isolated reforms. Nor does BBA endorse any specific “college or career readiness” standard(s). Similarly, though student test scores are discussed in the report, because they are widely employed as a key measure of student achievement, as discussed below, BBA does not believe that they can be reliably used to measure student skill level or “growth” over time, but rather that NAEP scores are the appropriate benchmark for such analyses.

state and the District of Columbia on the degree to which it implemented them, finding the majority severely lacking (StudentsFirst 2013). Indeed, with StudentsFirst targeting races for statewide public office, 50Can mounting state policy campaigns, Stand for Children working at the local school board level as well as in statehouses, and DFER advocating local, state, and federal policy changes, policies consistent with these reforms are being promoted and adopted in districts and states across the country (Cavanagh and Sawchuk 2012). Further evidence of their growing hold is seen in the number of consulting companies that employ them as the focus of their advising efforts on behalf of these states, cities, and districts.

The persistence of problematic achievement gaps between students of different races and income levels—gaps particularly afflicting students in poor urban schools and districts—is not in dispute. Indeed, 50 years of rising income inequality has accompanied a rapidly widening income-based achievement gap, which has eclipsed the race-based gap (Reardon 2011). The question is how we close these gaps.

Some reformers position their policies as higher minded than the policies advocated by others. Rhee and Klein advance a “no excuses” response to those who say poverty is an impediment to education, and frequently label those with whom they disagree as “defending the status quo” (StudentsFirst 2011). Others, such as Duncan, acknowledge the impact of poverty and promote a larger range of policies, while still emphasizing the same core set of reforms. But the question most critical for the millions of at-risk students and their families—and the nation as a whole—is not whether one group or another is “reforming” or “making excuses,” but what works and what does not.

This paper addresses that question directly. First, it reviews the reformers’ claims against the body of academic research devoted to each of these reform policies. As with most areas of policy, the evidence is mixed. Some of the policies are more promising than others, but none is even close to securing definitive endorsement based on the data. Indeed, policies that are “tested” in the social science equivalent of a laboratory inevitably vary, depending upon the “laboratory” employed (and many other variables), resulting in different outcomes in different cases.

Second, and of greater significance, the paper assesses these policies by reviewing their actual impact in three major cities. Over the past decade, different iterations of this set of reforms were implemented in Washington, D.C., New York City, and Chicago. The school districts in all three cities were controlled by their mayors, and thus had the potential to implement the reforms fully and flexibly. The outcomes in these three cities can therefore be seen as the best that a district enacting similar reforms could expect, with the fewest possible barriers to implementation.

Reliable National Assessment of Educational Progress data on changes in test scores, and thus achievement gaps, as well as somewhat comparable high school graduation rates, are available for all three city’s school districts, as are data from other large, urban districts that did not employ the same reforms, or that did so to a much lesser degree, providing a basis for comparison. Unlike city or state test scores, the NAEP data are comparable across districts in different states, thoroughly measure student learning and knowledge, and are sampled, and thus not subject to the system-gaming that can affect state test score data. Importantly too, NAEP data are available for a representative group of all students, and for minority and low-income students, the subgroups purported to benefit from these reforms. Specifically, the report explores how the reforms in each city have influenced:

- Teacher turnover and the stability of students’ educational contexts

- Student outcomes as measured both by disaggregated NAEP scores in math and reading and by the cities' or states' own assessments
- High school graduation rates
- School budgets

Although cost savings was not a promise, or goal, of reformers, tracking changes in school budgets is important, because public officials in many of the other cities and states that have recently endorsed similar reforms—including Cleveland, Detroit, and Philadelphia—say these policies will not only improve student outcomes but also save money and, thus, help close major deficits. For example, the Boston Consulting Group's plan, commissioned by the Philadelphia School Reform Commission, states, "Faced with both a financial crisis and persistently unsafe schools with low student achievement, the Philadelphia public schools are undergoing a transformation in order to ensure that all students have access to safe, high-quality schools, while bringing the budget back into balance." A spreadsheet detailing the financial aspects of the plan asserts that the budget will be balanced by 2014 (School District of Philadelphia 2012, 2 and 28).¹

The findings from the three cities do not show that the reforms have succeeded in fulfilling either their backers' promises or their claims. While in each city the mayor or chancellor reported large increases in test scores and/or decreases in achievement gaps since reforms were implemented, the improvements were often refuted soon after they were announced. In these districts, most test scores have not increased any more than in other, nonreform districts or than in the period predating the new policies; some test scores have fallen; and achievement gaps by race and income are as wide or wider than they were prior to the reforms. New teacher evaluation systems based heavily on student test scores seem to have increased teacher turnover, rather than improved teacher quality. Graduation rates among low-income and minority students have improved somewhat, narrowing citywide gaps in those rates, though it is not entirely clear that these cities have made gains in graduation rates that are greater than those in other cities.

Moreover, these reforms have had significant costs. All three cities received substantial private donations and enacted major budget increases, with New York City virtually doubling education spending (in real dollars) in just a decade. While the reforms are not the main cause of the increases, some aspects have introduced new costs. Moreover, the costs raise questions about the viability of the reforms today in other cities facing major budget shortfalls, and the sustainability of the reforms in these three cities.

At the same time, officials in these cities have implemented other new policies that produced positive results. New schools with a focus on real-life experiences and strong teacher-student relationships have improved outcomes for high-risk students. Smarter teacher recruitment efforts have brought qualified educators to hard-to-serve schools. Enhanced and expanded prekindergarten (pre-K) programs may be boosting low-income students' elementary school achievement. However, these promising policy changes have largely been ignored. Because of the excessive focus on prominent "reforms," less visible programs with real potential to narrow achievement gaps haven't had the kind of support or resources they need to succeed or expand.

In short, some reform elements are working, but most are not. And policies that address students' noncognitive needs are lacking. But multidimensional, nuanced efforts with a long-term trajectory and mixed outcomes can be difficult to communicate when the urgency to implement quick fixes prevails. As such, reformers have mostly advanced simpler messages that fail to tell the real story. Districts, states, and the federal government should take a hard look at the grow-

ing evidence to ensure that what actually works gets replicated, and that popular but ineffective or counterproductive “reform” policies do not become the new status quo.

A brief synopsis of the evidence regarding core components of the reform agenda

An American University forum in Washington, D.C., in March 2012, “Education Now: Cities at the Forefront of Reform,” promoted the basic tenets of the reform movement. Representing the federal government and two of the three cities featured in this report were U.S. Education Secretary Arne Duncan, Chicago Mayor Rahm Emanuel, and New York City Mayor Michael Bloomberg, along with their respective superintendents. While Washington, D.C., leaders did not participate, Los Angeles Mayor Antonio Villaraigosa and Los Angeles’s schools superintendent were there, representing a district without mayoral control (American University 2012). The participants attributed a decline in U.S. global competitiveness to unions’ protection of poor teachers, an insufficient emphasis on STEM (science, technology, engineering, and math) instruction, and a failure to prepare students for at least a two-year college. To correct these deficiencies and others, they advocated for more school-performance data to be used to eliminate what they termed “low-quality” schools (and to enable parents to make informed choices among regular district schools and an expanded number of charter schools) and greater use of test scores to assess teachers and principals’ effectiveness and make hiring and firing decisions.

All three leaders asserted that this set of changes would improve student achievement and narrow race- and income-based achievement gaps.² As the literature review below illustrates, however, these claims are exaggerated in some respects and false in others. There is no sound theoretical basis for the premise that market-oriented reforms can fix income-correlated gaps in educational attainment.

Test-based accountability systems

Bloomberg was perhaps the strongest proponent among the panelists of test-based accountability. Whereas Duncan acknowledged that No Child Left Behind had led to excessive reliance on testing, Bloomberg defended the practice of “teaching to the test” as reflecting state standards and curriculum. All the mayors, along with Duncan, emphasized firing “bad” teachers as a key means of improving quality.

“If you care about education, you have to care about the quality of the teachers and get those out of the system who cannot do the job,” Bloomberg said (American University 2012).

Research does not show that it is effective or reliable to rely on student test scores to evaluate, fire, or provide merit bonuses to teachers (Baker et al. 2010). First, test scores are influenced by myriad factors, both in and out of school, most of which are beyond the ability of teachers to influence, but which nevertheless lead to judgments regarding teachers’ effectiveness (Rothstein, Jacobsen, and Wilder 2008).

Even “value-added” methods (VAM), which try to account for the impact of student background characteristics on test scores, inevitably fail to capture many factors beyond teacher control. Though they use a complex statistical technique called regression analysis, which incorporates multiple factors, value-added measures still have many flaws. The initial development of value-added systems in the 1990s enabled evaluators to move beyond the use of raw student scores,

which only reflect a student's knowledge at a given point and do not capture change by taking into account where he or she began. Value-added methods correct for that problem by using two tests, one from the current year and another from the prior year, or, one from the beginning and another from the end of the school year, thus capturing student growth over a period. Also, some versions of the regression analyses control for easier-to-measure individual student characteristics that are associated with learning and achievement, such as use of free-and-reduced-price lunch, and race and/or ethnicity. Nonetheless, value-added methods still employ narrow tests as the only measure of student knowledge and skill, incorporate only a very limited number and range of out-of-school factors, and cannot capture many of the other in-school influences, such as tutors, team teaching, or counseling, that can influence student scores as much as, or even more than, the individual teacher being assessed (Weiss 2011).³

As a result, value-added scores are problematic. After New York City published value-added scores for every teacher in the city in 2012, veteran teachers considered by parents and principals as among the city's best were depicted as poor performers, scores shot up and down from year to year, and many of the scores were found to be wildly inaccurate and based on incorrect data (see, for example, Pallas 2012a and Winerip 2012). According to the Board on Testing and Assessment of the National Research Council of the National Academy of Sciences, "VAM estimates of teacher effectiveness should not be used to make operational decisions because such estimates are far too unstable to be considered fair or reliable" (BOTA 2009).

Second, basing important decisions on student test scores has negative consequences. Teachers are driven to "teach to the test," adjusting instruction to cover material on the test or even teaching test-taking skills rather than content in order to produce the higher scores demanded (Baker et al. 2010; Scherrer 2011; Rothstein, Jacobsen, and Wilder 2008). Test-based accountability systems also give teachers incentives to choose classes and students who are likely to score higher, and/or to gain more from one year to the next, and, perhaps most troubling, to avoid serving students who are less likely to do well on tests (Peneston et al. 2011).⁴ As such, some experts caution that value-added methods and other test-based teacher evaluation systems may lead to the loss of qualified and strong teachers in hard-to-staff schools and districts (Weiss 2011). At the extreme, some observers note that the increasingly high stakes attached to raising student test scores have led to a growing number of district-wide cheating scandals (FairTest 2012).⁵

Proponents of using student test scores to award merit-pay bonuses to teachers assert that current salary systems provide teachers with insufficient incentive to focus on improving student outcomes, and that bonuses for raising test scores would thus improve teaching and student outcomes. The same shortcomings of VAM in other contexts apply to test-driven merit-pay systems, in particular the failure of VAM to account for factors influencing test scores that are beyond teacher control, and the added incentive to narrow instruction to test content in order to raise scores.

Experiments assessing test-driven merit-pay outcomes have been conducted in several sites, but none has yet found these systems effective. One large pilot conducted in the Metropolitan Nashville Public Schools over three school years, 2006–2009, was designed to avert the inadequate pay and rigid strategies blamed for failures in other, prior pilots. It offered large bonuses, from \$5,000 for a small test-score bump to \$15,000 for raising scores to the very top of the continuum (Springer et al. 2010). The structure of the experiment also allowed teachers to employ any strategy they chose to raise scores, from obtaining more professional development or coaching to collaborating with others. Despite the thoughtful design and major investment, however, the bonuses produced no improvement in student test scores (Springer et al. 2010). Moreover, teachers surveyed did not believe that those who received bonuses were better teachers.

These findings were supported by a RAND evaluation that found that New York City’s experimental program had no impact on student test scores. According to the authors, “the results of this evaluation add to a growing body of research from the United States that finds no effects on student achievement of narrow pay-for-performance policies that focus only on financial incentives without other features, such as targeted professional development or revised teacher evaluations” (Marsh et al. 2011).

A different model funded by a federal grant and employed in Chicago downplayed individual teacher rewards and spread bonuses throughout the schools’ staff, instead focusing on other aspects of teacher support to boost student achievement and teacher retention. Nonetheless, results at the two-year mark were disappointing: Neither test scores nor teacher retention had improved (Sawchuk 2010).⁶ At four years, there was some increase in school-level teacher retention, but still no impact on student achievement (Glazerman and Seifullah 2012).

In his 2009 best-seller *Drive*, labor policy expert Daniel Pink contends that merit pay programs do not induce teachers to improve their instruction because, unlike corporate executives, teachers’ intrinsic motivation for entering the profession, and for improving student outcomes, is not fundamentally rooted in salaries or compensation. Moreover, because the field requires conceptual and creative thinking, a simplistic reward-for-higher-performance mechanism does not apply (Pink 2009).

Research confirms Pink’s suggestion that recruiting good teachers to low-income and struggling schools and retaining them relies much more heavily on their working conditions than on salaries, or students’ income or background. The findings from a recent Harvard Graduate School of Education study show that “guaranteeing an effective teacher for all students—especially minority students who live in poverty—cannot be accomplished simply by offering financial bonuses or mandating the reassignment of effective teachers” (Moore Johnson, Kraft, and Papay 2012, 18).

School closures

Another facet of the school reform agenda is the intentional closure of neighborhood schools. School closures can be driven by the desire either to reduce district spending (by shuttering an underutilized building) or to transfer students from a “failing” school to one that is “higher-performing,” and thus able to provide better educational opportunities.⁷

Studies of school closures show that both premises are flawed. First, the determination that a school is underenrolled is, itself, subject to manipulation, as explored in the section of this paper reviewing real-life outcomes. Also, as discussed in more detail later in this paper, purportedly underutilized neighborhood schools are often converted to charter schools that serve even fewer students than before conversion—charter schools which are not deemed underenrolled.

The premise for closing failing schools, often targeted for closure based on student test scores, is that many of the teachers, the principal, and other aspects of the school are too problematic to remedy, so improvement must start from scratch.⁸ Here, too, the research suggests that the policy does not improve student outcomes overall and may even exacerbate existing problems. A 2009 study of students from 18 Chicago elementary schools closed between 2001 and 2006 found no overall effect on those students’ academic performance, but raised a concern that since most students transferred to very low-performing schools, while only 6 percent went to high-performing schools, the disruption offered no benefits (de la Torre and Gwynne 2009, 16). Using data from an unnamed high school in a large urban school district in the western United States, a 2010 study assessed the “academic performance and experiences of Latino

and African American high school students in the year following” their school’s closure (Kirshner, Gaertner, and Pozzoboni 2010). Among the findings were declines in academic performance and “added stressors [for] students who were already contending with challenges associated with urban poverty.” And a recent study of a district with declining enrollment that used student achievement to target schools for closure found adverse effects on test scores and attendance, though the study noted that these effects could be minimized by transferring students to higher-performing schools (Engberg et al. 2011).

In all, the studies point to a critical limitation: Districts that use low student test scores to label schools as “failing” and close them are unlikely to have a sufficient number of high-performing schools to serve transferred students. As such, even if students could, hypothetically, benefit from such moves, beneficial moves may not be possible in many, if not most, current contexts.

Increasing the reach of charters

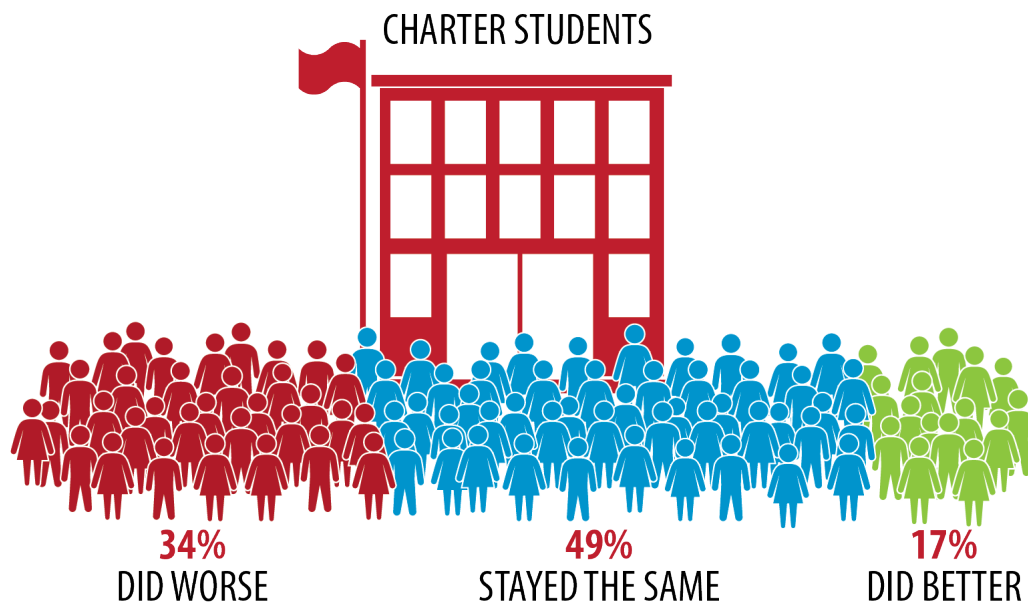
Expanding access to charter schools—framed as key to parental and student choice—was another major priority for the panelists of the March 2012 American University forum, and a common focus of reforms in the three cities assessed in this paper. However, assertions that charter schools improve educational outcomes are not supported by rigorous studies. Research findings on the effectiveness of charter schools in boosting achievement for low-income students and in narrowing achievement gaps are mixed at best. The largest and most comprehensive study to date was conducted in 2009 by the Center for Research on Education Outcomes (CREDO) at Stanford University. *Multiple Choice: Charter School Performance in Sixteen States* analyzed outcomes of charter school students in 70 percent of all charter schools in the United States by using simulations to compare them to their “virtual” peers in the traditional public schools they would otherwise have attended. The report found “wide variation in performance.” In particular, math test scores were better for 17 percent of charter school students, were the same for roughly half of charter students, and were worse for 34 percent of charter students. In other words, twice as many students lost as gained from being in a charter school, a discovery that the researchers call “sobering” (CREDO 2009b, 3).⁹ Reading scores in charter schools versus traditional public schools were similarly mixed, with gains differing in elementary versus middle and high schools, across ethnic groups, and in first versus later years in the charter school.

Charter schools are public schools, and the vast majority of their funding comes from public sources. As such charter schools, like other public (and private) schools, vary tremendously in terms of resources, teacher qualifications, school day and year, services provided, and a host of other factors. In the case of charters, these factors include state policies authorizing, funding, and monitoring them. In addition, unlike regular public schools, some charters, like KIPP schools, employ established, tested models, while others are newly minted and employ practices that have not been assessed.¹⁰ While most are nonprofits, some are for-profits. Given this variation across a range of factors, it is not surprising that studies of specific types and brands of charters, and of charters in particular cities and states, have found substantial variation in their success, relative to comparable traditional public schools.

New York and Ohio illustrate many of the substantial divides in state-level findings. A study of New York City charter schools by the same Stanford researchers found significant benefits for most groups of students who enroll, including low-income, minority, and special-education students, though none for English language learners (ELL) nor for those who were retained in grade (CREDO 2010). Conversely, the Stanford study found that charter schools in Ohio, a national leader in the authorization and opening of new charter schools, performed significantly worse than comparable

CHARTER SCHOOLS OFFER MIXED RESULTS, NOT UNIVERSAL SUCCESS, FOR TARGETED STUDENTS

Reformers in DC, NYC, and Chicago depicted charter schools as solutions to “failing” schools and “dropout factories.” City-level outcomes painting a much more complex picture—with some of the neediest students left out altogether—are mirrored by a national study of charter students’ performance relative to their regular school equivalents.



Source: Center for Research on Education Outcomes, *Multiple Choice: Charter School Performance in Sixteen States*, 2009. The study compared charter school students with their “virtual” counterparts in comparable regular public schools.

public schools (CREDO 2009a). A second in-depth study of the history of charter laws and schools in Ohio found that the state’s charter schools were plagued by mismanagement, financial misconduct, and lack of accountability (Dingerson 2008, 17).¹¹ These disparities mirror the broader findings in the multi-state 2009 CREDO study, which discovered substantial variation across states “over and above existing differences among states in their academic results” (CREDO 2009b, 3).¹²

These findings, in turn, should inform broader discussions regarding the role of extended learning time, enrichment activities, prekindergarten (pre-K) and other early childhood programs, health clinics, the mix of students served, and other features of charter schools that may differ from their traditional public school counterparts and which are known to influence student outcomes. Such discussions are critical both to understanding why and which charter schools “out-perform” their regular public school counterparts and to guiding changes in education policy irrespective of the type of school.

One consequence of the expansion of charter schools in a school district is a corresponding loss of students in the district’s regular or neighborhood public schools. Of course, districts lose students for many reasons; for example, many rural districts have lost students in recent decades as their populations declined. The district public-to-charter school

shift, too, is often described as a decline in student enrollment, though it often does not, as do rural losses, reflect an overall reduction in the number of students who are served in the district.

The shift nonetheless presents problems for school districts, as if a real decline had taken place. Because the number of public school students determines district budgets, schools that lose students to charters may have similar overhead expenses but fewer dollars to pay for them (Dingerson 2008, 31).¹³ The way in which charters operate also can complicate planning, hiring, and other logistical matters for traditional public schools. In Ohio, for example, charter recruitment and closures during the school year mean more student mobility, leading to disruption and lack of capacity. In Dayton, where public schools begin two weeks earlier than most charter schools, the district must staff the public schools to accommodate all enrolled students for the first weeks; when students enrolled in charter schools move to those schools two weeks into the school year, the school district has no way to pay the teachers who are no longer needed. In the longer term, districts known to fire or attempt to hire teachers midyear in response to outflows of students to or influxes from charter schools have an increasingly hard time recruiting high-quality teachers. A description of how parents in Washington, D.C., scramble to get their children into the “right” school at the start of the year, and reserve slots at multiple schools until they choose one, illustrates much the same problem posed by some aspects of charter-based choice (Brown 2012b).

Turnover and instability

The asserted goal of test-based teacher evaluations, elimination of bad teachers and principals, and closures of “failing” schools is improved teacher quality and better student outcomes (Rhee and Klein 2010; StudentsFirst 2011). As Anthony Bryk and other scholars have found, however, the “cake mix” required to turn around troubled schools is a complex one, and requires, among other ingredients, stable and cohesive teacher and leadership efforts that evolve and cohere over time (Bryk et al. 2010). This makes sense; with so much else unstable and uncertain in their lives—housing, meals, parental attention—low-income students need their schools to provide that stability. New schools and new teachers, especially when they become a pattern repeated over several years, can further disrupt their already chaotic lives (Hanushek, Kain, and Rivkin 2004).

If these reforms lead to instability (and the next section of this paper shows that they do), they are producing a negative, unintended consequence.

This is not to say that bad teachers should not be fired, that teachers cannot do their jobs well after only a few years, or that there is never a reason to close a school. However, a large body of research explains the advantages of experienced teachers over lower-paid novices, and of the importance of continuity and stability in improving student outcomes (e.g., Haycock 2006; Holzman 2012).¹⁴ As Tom Carroll, president of the National Commission on Teaching and America’s Future, pointed out, “research clearly shows that with each year of experience, teachers improve their proficiency and effectiveness during the first seven years. National Board for Professional Teaching Standards (NBPTS) certification demonstrates that many teachers are still gaining in proficiency and improving their effectiveness after an average of 11 years of teaching” (Carroll and Foster 2010, 12). In fact, a recent study on the impact of teacher turnover concluded that, distinct from the relative quality of teachers who may be brought in to replace those who leave, teacher turnover itself harms a school.¹⁵ Turnover lowers school morale and professional culture, depletes the staff’s store of knowledge about students and the community, and impedes the collegiality, professional support, and trust that teachers need to

improve student achievement (Loeb, Ronfeldt, and Wyckoff 2012). In order for increased teacher turnover to improve student outcomes, then, teachers who leave must be replaced by teachers who are substantially more effective.

A “no excuses” take on the influence of poverty

A common refrain underpinning espousal of the reforms discussed above is the dismissal of the impacts of child, family, and community poverty on student achievement and school success. When discussing their reform agendas, leaders such as Michelle Rhee have employed a “no excuses” rhetoric, asserting that any teacher who cites poverty as a driver of student or school failure is excusing him or herself from the responsibility to teach well. Indeed, Rhee alleges that, “what often happens when we start to talk about wraparound services [that acknowledge the need to address poverty’s impacts] is a lot of people start to give up responsibility” (Maxwell 2012). Former New York City schools leader Joel Klein has gone further, explicitly rejecting decades of research that demonstrate the benefits of alleviating poverty-related obstacles to effective learning.

No single impediment to closing the nation’s shameful achievement gap looms larger than the culture of excuse that now permeates our schools. Too many educators today excuse teachers, principals, and school superintendents who fail to substantially raise the performance of low-income minority students by claiming that schools cannot really be held accountable for student achievement because disadvantaged students bear multiple burdens of poverty. The favored solution du jour to minority underachievement is to reduce the handicap of being poor by establishing full-service health clinics at schools, dispensing more housing vouchers, expanding preschool programs, and offering after-school services like mental health counseling for students and parents. America will never fix education until it first fixes poverty—or so the argument goes. In fact, the skeptics of urban schools have got the diagnosis exactly backward. The truth is that America will never fix poverty until it fixes its urban schools. (Klein 2009)

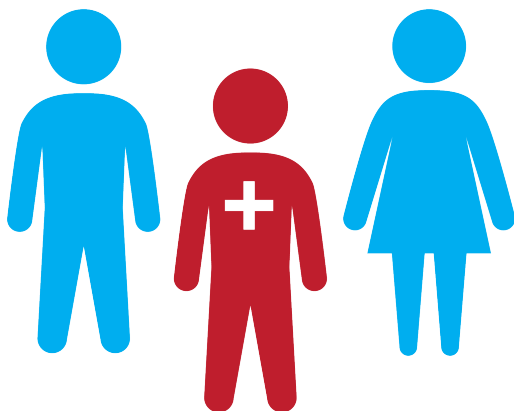
But rigorous studies have revealed the many ways in which growing up with limited resources and being raised by poorly educated parents impede students’ development and academic achievement (Rothstein 2004; Barton and Coley 2007; Berliner 2009; Ladd 2011; and Yoshikawa, Aber, and Beardslee 2012). These include the large achievement gaps that emerge long before children enter kindergarten (Galinsky 2006; Isaacs 2012; Rolnick and Grunewald 2007). These achievement gaps are exacerbated by gaps in nutrition and health status and care that similarly divide along lines of economic class and race (Rothstein 2004 and Berliner 2009). In addition, loss of learning in afterschool hours and over the summer account for large shares of income- and race-based achievement gaps (Alexander, Entwisle, and Olson 2007).

The reforms have not strengthened school systems in the three cities

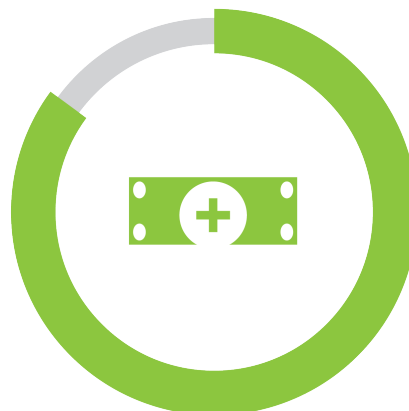
As briefly summarized in the previous section of this report, extensive research has assessed the effectiveness of the three main components of the “reform” agenda: using standardized test scores to make decisions regarding schools and educators; closing schools that are labeled failing and/or underenrolled; and expanding access to charter schools to give parents and students greater choice. Results have been inconclusive, and policymakers have used the findings to make cases both for and against these policies. As such, these studies have led to renewed debate, rather than to consensus, regarding which policies best improve outcomes for underachieving students. Indeed, education policy has become increasingly politicized, and discussions hostile, as advocates of these reforms and their opponents cleave to their respective positions (Bushaw and Lopez 2012).

REFORMS FAIL WHEN THEY IGNORE THE POVERTY-RELATED CAUSES OF ACHIEVEMENT GAPS

Lack of consistent physical and mental health care is a major driver of the opportunity gaps associated with growing up in poverty. Low-income children miss many more days of school due to preventable illnesses, relative to their wealthier peers—a reality largely dismissed in reform agendas.



1 IN 3 CHILDREN
LIVING IN POVERTY VISIT THE
EMERGENCY ROOM EVERY
YEAR FOR PREVENTABLE
MEDICAL CONDITIONS



85%
DECREASE IN HOSPITALIZATION
COSTS FOR CHILDREN IN
CINCINNATI WITH
SCHOOL-BASED HEALTH CLINICS

Source: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics, *Summary Health Statistics for U.S. Children: National Health Interview Survey, 2010*, Table 10; National Assembly on School-Based Health Care, *Cost-Savings of School-Based Health Centers*, 2010.

Cities looking to implement similar plans will thus find it more informative to understand how these reform efforts have actually worked in practice.

The main sections of the paper, below, assess the reforms implemented by Michelle Rhee during her tenure as DC Public Schools (DCPS) chancellor, with many of the same policies in place under her successor, Kaya Henderson; by Joel Klein when he was New York City Schools chancellor; and by Education Secretary Arne Duncan when he was Chicago Public Schools chief education officer (CEO), with his policies largely followed by his successors Ron Huberman, Jean-Claude Brizard,¹⁶ and, it appears, new Chicago Public Schools CEO Barbara Byrd-Bennett.

Rhee was appointed by Mayor Adrian Fenty in 2007, and she left after he failed to win reelection in 2010. Klein served from 2002, when he was appointed by Mayor Michael Bloomberg, to 2010, when he left to become an advisor to media mogul Rupert Murdoch. As CEO of Chicago Public Schools (CPS), Duncan worked with Mayor Richard Daley from 2001 through 2008 to implement “Renaissance 2010”—market-oriented reforms that emphasized school closings and consolidations and charter schools (Lipman 2009).¹⁷

Chicago, New York, and Washington, D.C., are, to a large degree, pioneers in implementing the three specific policies assessed in this paper. While many other cities have instituted some of the same policies or are beginning to do so, none enacted them as comprehensively or with as much mayoral backing, as these three cities. Indeed, Rhee, Klein, and Duncan all cite mayoral control as one key to effective education reform.

Thus, we can make a valid comparison between the “reform” cities of Chicago, New York, and Washington, D.C., with “nonreform” cities. For the purposes of this paper, the latter include other large, low-income, urban districts for which similar data—in particular, test scores from the National Assessment of Educational Progress, and graduation rates that are comparable across districts—have been collected in recent years.

The problems driving achievement gaps are complex, and policies designed to narrow them go beyond those explored in this paper. Any leader of a major school district promotes and/or institutes a variety of policies and programs, and makes changes to many others. This paper does not assess all of the policies and programs that these three districts enacted or implemented during this period. It focuses on the three that were core to these leaders’ tenures, and that they and other reformers are most widely promoting and replicating elsewhere, for that reason. It also, however, reviews certain smaller policy changes that, while overshadowed by the big three, may have contributed to the outcomes reported, and/or have shown promise in boosting future achievement.

Note that this report tracks gains (and losses) based on NAEP scores across the three cities during their periods of major reform, and compares changes with trends in other large, urban districts and nationwide. Data for large, urban districts are from NAEP Trial Urban District Assessment (TUDA) reports unless otherwise noted. TUDA uses NAEP scores from a representative sample of students in large, urban districts with high proportions of low-income and minority students. TUDA began in 2002 with six districts and had 21 as of 2011.¹⁸

As mentioned earlier, the use of NAEP, versus city or state test scores, allows for apples-to-apples comparisons across cities in different states, and makes these comparisons more valid and reliable than many that are commonly cited; NAEP is a more thorough measure of student learning and knowledge than many state tests, and it is sampled, and thus not subject to the system-gaming that can affect test scores. The report also notes scores on other tests, since they are often used as the basis for claims of gains in specific schools or types of schools. That non-NAEP tests should be viewed with caution was underscored in a study of CPS reforms by the Consortium on Chicago School Research.

“The publicly reported statistics used to hold schools and districts accountable for making academic progress are not accurate measures of progress.”

Many of the findings in this report contradict trends that appear in publicly reported data. For instance, publicly reported statistics indicate that CPS has made tremendous progress in elementary math and reading tests, while this analysis demonstrates only incremental gains in math and almost no growth in reading. The discrepancies are due to myriad issues with publicly reported data—including changes in test content and scoring—that make year-over-year comparisons nearly impossible without complex statistical analyses, such as those undertaken for this report. This leads to another key message in this report: The publicly reported statistics used to hold schools

and districts accountable for making academic progress are not accurate measures of progress. (Luppescu et al. 2011, 5)

As this report details, when the market-based policies at the center of the reform agenda play out in a comprehensive manner across many years, the results, as captured in reliable data, are not encouraging. This evidence should give pause to supporters of this popular set of reforms, including policymakers who are implementing similar policies. Reforms that produce a lack of progress on improving test scores or closing achievement gaps are no different from the “status quo” that they purport to break.

This section of the report also examines data on school district budgets over the period of the reformers’ tenures. While the reformers in the three cities that are the focus of this paper have not claimed such savings as a benefit of reforms, other reform advocates make this promise. It is therefore important to understand not only the potential of these reform strategies to improve student outcomes, but the degree to which they can, indeed, provide such savings without doing further harm to students.

Washington, D.C.: Firings and closures increased churn

During Michelle Rhee’s three-and-a-half years as chancellor of the DC Public Schools, several hundred teachers were fired. Many more were laid off or left voluntarily, along with dozens of principals, increasing the rate of turnover in the schools, especially those serving the District’s lowest-income students. Closing underenrolled and low-performing schools further increased churn for those students without visible benefits or budget savings.

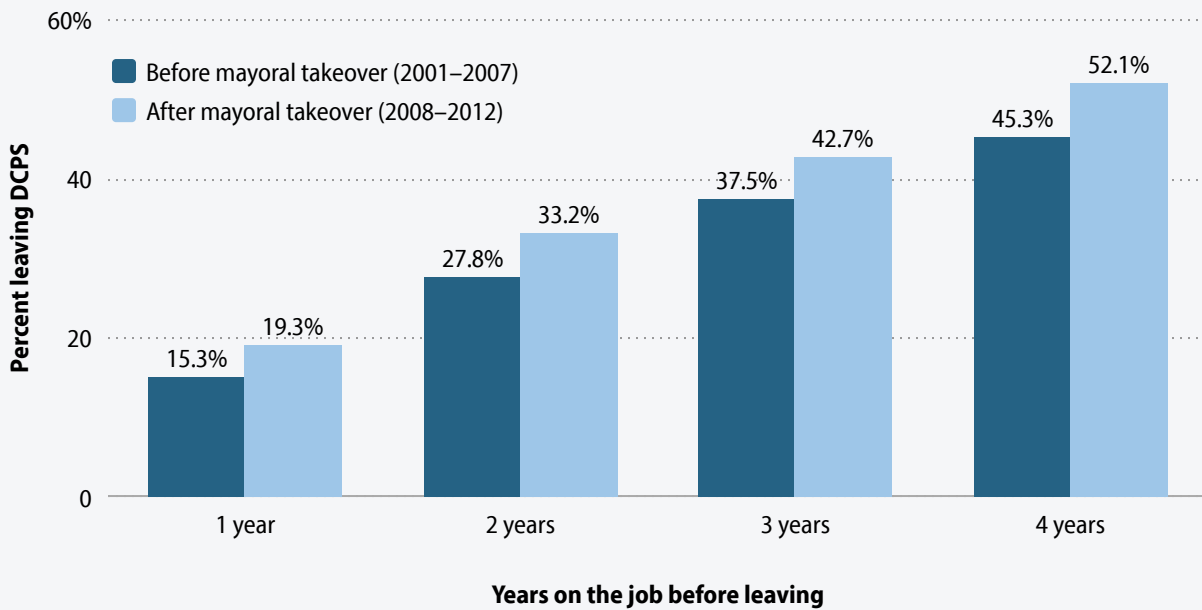
Test-based teacher evaluations and turnover. In 2009, DCPS instituted a system, called IMPACT, to use student test scores as a basis for evaluating individual teachers, in order to reward those found to be “effective” or “highly effective” and fire those deemed “ineffective” or “minimally ineffective.” Under Rhee, 50 percent of a teacher’s IMPACT score was based on his or her students’ test scores on the DC Comprehensive Assessment System (DC CAS), the district’s standardized test, which is given each April. The remaining 50 percent was based heavily on classroom observations, with a small portion devoted to “collaboration” and “professionalism” (DCPS 2011). Complaints about the system from teachers and others led to several changes under Rhee’s successor, Kaya Henderson, including allowing some teachers who received “minimally effective” ratings to keep their jobs, and exempting teachers with two consecutive “highly effective” ratings from three of five annual observations (Huffington Post 2011). More recently, Henderson reduced the weight accorded to value-added scores from 50 to 35 percent of the total IMPACT rating (Gartner 2012b).

Prior to IMPACT, DCPS had in place a rarely used “90-day plan” that allowed principals to place teachers in a semi-probationary status that required that teachers develop a remedial plan, be observed by the principal while they taught, and fulfill other requirements or face rapid dismissal (Turque 2008). Rhee employed this plan to eliminate a small number of teachers before IMPACT was implemented.

Turnover, which generally is high among urban and low-income schools and has been identified as a key impediment to student success, increased sharply after mayoral takeover and Rhee’s arrival in 2007, and the subsequent institution of market-oriented reforms, as illustrated in **Figure A**. While the high rate of turnover was to some extent intentional—Rhee announced when she was hired that many weak teachers would have to leave in order to improve the

FIGURE A

District of Columbia Public Schools' teacher attrition rates before and after mayoral takeover in 2007



Source: DCPS teacher retention data, 2001-2012 (Levy 2012g)

ECONOMIC POLICY INSTITUTE

system—it increased the share of the teaching pool consisting of novice teachers, who lack the experience needed to boost instructional capacity.

The share of teachers leaving after one year in the system increased from 15.3 percent in 2001–2007 to 19.3 percent in 2008–2012; the share leaving after two years increased from 27.8 percent to 33.2 percent; the share leaving after three years increased from 37.5 percent to 42.7 percent; and the share leaving after four years increased from 45.3 percent to 52.1 percent (Levy 2012g). Among the district’s charter school teachers and new teacher hires, the rates were much higher: Over half of new teacher hires left in their first *two years*, and eight of 10 were gone by the end of their sixth year (Simon 2012). While there are no perfectly comparable data from other cities, these numbers suggest that DCPS has one of the nation’s highest new teacher turnover rates. Scholars estimate that, on average, 30 percent of new U.S. teachers leave the profession after five years (Ronfeldt et al. 2011). That rate is as high as 45 percent in the highest-poverty school systems (Hunt and Carroll 2003).¹⁹

While Rhee made a point of firing teachers deemed ineffective, the higher attrition rate cannot be attributed largely to teacher terminations. Comparisons of data on teacher termination with reports on the changes in DCPS and news articles all suggest that most teachers left of their own volition.

At the end of Rhee’s first year, 2007–2008, 350 of the system’s roughly 4,000 teachers, or about 8 percent, were terminated, for reasons that were not clearly specified (Levy 2012h).²⁰ Then, in 2009, Rhee fired 131 teachers for poor

appraisals under the “90-day plan” that preceded IMPACT (Levy 2012h). She also fired 117 teachers for lack of licensure. In July 2010, Rhee fired 165 more teachers who had received ineffective ratings under the new IMPACT system and announced that hundreds of others rated “minimally effective” had one year to improve their performance before facing dismissal (Turque 2010c).

In 2011, Rhee fired 65 teachers who had been rated ineffective and another 141 who had been rated minimally effective for two years in a row. These terminations for ratings of ineffective or minimally effective accounted for roughly 3 percent, 4 percent, and 5 percent of the total DCPS base teaching force in 2009, 2010, and 2011, respectively (Levy 2012h). With attrition rates of 20 percent or more for first-year teachers and higher among the overall pool of teachers, the vast majority of exits, then, were due to other factors.

Altogether, around 1,000 educators were fired during Rhee’s three-and-a-half-year tenure, about half for ineffective ratings, and many of the rest ostensibly for budgetary reasons. For example, in July 2009, the D.C. Council cut the DCPS budget, directing that most of the money be taken out of summer school, which would not take effect until the next year. Rhee disobeyed the council’s line item authority and instead laid off 266 teachers. Rhee later suggested that she had taken advantage of the budget problem to fire teachers who had hit children, had sex with students, or had otherwise seriously broken the rules. When questioned by policymakers about why such teachers had not already been fired, Rhee offered some details regarding what she said were eight of the 266 teachers, but did not provide either back-up information or names (Perdido Street School 2010). At the same time, DCPS continued hiring; between March 2009 and January 2010, DCPS hired 2,211 new employees, including 1,034 teachers. Three quarters of these new hires took place in August, just before Rhee fired the 266 teachers (Levy 2012e). Then-Council Chair Vincent Gray pointed to DCPS’s decision to hire hundreds of new teachers just before firing the nearly 300 others, suggesting that Rhee and Fenty used the budget as an excuse to replace experienced teachers with lower-paid novices (Turque 2009).

Other data show that turnover affected some students more than others. Levy has calculated individual DCPS school attrition data, which show 30 percent, 40 percent, and even higher percentages of teachers leaving in a single year, such that in some schools, the *majority* of teachers are new at the start of the school year. For example, at Anacostia High School, where almost the entire staff (88 percent) was new, having been specially selected in fall 2010, one third left the following spring (Levy 2012f). This happened despite Rhee’s appointment of a charter management organization to run the school and improve its performance. Another distressed high school, Ballou, lost two-thirds of its teachers over just two years (2010–2012). While not as astronomical, attrition rates in several less-distressed schools were still quite high. Bancroft Elementary in Mount Pleasant, which serves a high proportion of Latino students, had a two-year teacher turnover rate of 40 percent, and Barnard Elementary School, housed in a beautiful new building in the rapidly gentrifying Petworth area, lost 45 percent of its teachers in two years (Levy 2012f).

These high turnover rates harm students, especially the low-income students concentrated in most DCPS schools. Mary Levy, a former school finance litigator who has been a consultant for the district’s city council, noted the negative impact of these high rates on both students and the system’s budget: “We lose half of our new teachers, our new hires ... within two years of their hiring. This is, of course, expensive because it costs money to recruit. We have a lot of teachers in their first three years of teaching. We have a lot at the high end, [too], and very few in between. This does not bode well for the future” (Levy 2011a).²¹

The National Commission on Teaching and America's Future attached a rough price tag to this high level of churn, using data on the actual cost to the district of time and money spent on activities associated with teacher leavers including: "recruitment and advertising; special incentives; administrative processing, training for new hires, first-time teachers, and all teachers; 'learning curve,' and transfer" (Barnes, Crowe, and Schaefer 2007, 13–14) As Tom Carroll, president of the commission, notes, this figure ignores "what may in fact be the largest costs of teacher turnover: lost teaching quality and effectiveness" (Carroll 2011, 4).

In estimating the high cost to DCPS, Carroll explained that

teaching is no different than any other profession—experience matters. Researchers have found that teachers reach peak effectiveness with about seven years of experience. But 80 percent of the teachers hired by D.C. this year will be gone before they get there. ... [As a result,] the District is burning about \$12 million a year on teacher churn—\$12 million that is spent hiring and replacing teachers with no gain in school performance. (Simon 2012)

For a teacher replacement strategy to effectively improve student performance, then, the system must be losing mostly bad teachers. But it may rather be losing good teachers, who are leaving voluntarily (or by accident). The termination data provided above make clear that the majority of teachers who left DCPS during Rhee's tenure had not been rated ineffective, and that even among those who were terminated, only half had "ineffective" ratings. One likely reason that teachers who were not fired left anyway is an atmosphere of mistrust that intensified under Rhee. As Levy noted, DCPS has "never been a nice place to work, now it's an awful place to work" (Levy, personal communication 2012). Indeed, a recent report on rising attrition asserts that "DCPS has become a teacher turnover factory" that has a hard time retaining teachers who are committed to their school and the community it serves (Simon 2012).

It could also be that teachers who knew they were likely to get a bad rating retired or resigned to avoid that stigma. If the ratings accurately identified ineffective teachers, this might be a good thing. Because they are based heavily on student test scores, however, DCPS teacher evaluations may not be very accurate. One highly publicized example suggesting inaccuracy is seen in *Washington Post* education reporter Bill Turque's depiction of the termination of MacFarland Middle School teacher Sarah Wysocki, who had been praised by her principal, fellow teachers, and especially parents of her students as one of the best (Turque 2012a). Turque recounted the flaws with the value-added scoring method that failed to account for Wysocki's stellar performance in classroom observations and led to her firing. The article also noted Wysocki's concerns that many of her students' starting test scores—against which the "growth" in her classrooms was measured—were impossibly high, because students came from a school in which widespread cheating was under investigation. The suggested inaccuracy of her "ineffective" score was borne out by the fact that, immediately upon being fired by DCPS, Wysocki was hired by Fairfax County Public Schools, one of the nation's wealthiest and best districts, where she is considered a superb teacher (Brown 2011).²² While this is just one anecdote, it illustrates several flaws with IMPACT that are potentially applicable in many cases.

A lack of ratings reliability has not been the only problem. Veteran teacher Mary Sutton, who taught for 11 years at Malcolm X Elementary School in Anacostia and was one of few teachers to receive an "effective" rating in a school with a large share of low-income students, was "excessed" out of the system in 2011. While DCPS asserted the need to eliminate even effective "people we don't have an express need for," Nathan Saunders, president of the Washington Teachers

Union, claimed that “a push for younger, lower-cost hires—some recruited from programs such as Teach for America and D.C. Teaching Fellows—has wrongfully forced out seasoned practitioners” (Turque 2011a).

Saunders said Henderson’s position would be understandable if the District were not hiring hundreds of teachers every spring and summer, many of them in fields in which the excessed teachers worked. They include math, reading and other core subjects. Payroll records show that the city generally hires 300 to 400 teachers a year, many of them younger and less costly than the veterans let go. A recent independent study showed that the proportion of first- and second-year teachers has grown sharply in five years, especially in high-poverty communities east of the Anacostia River. (Turque 2011a)

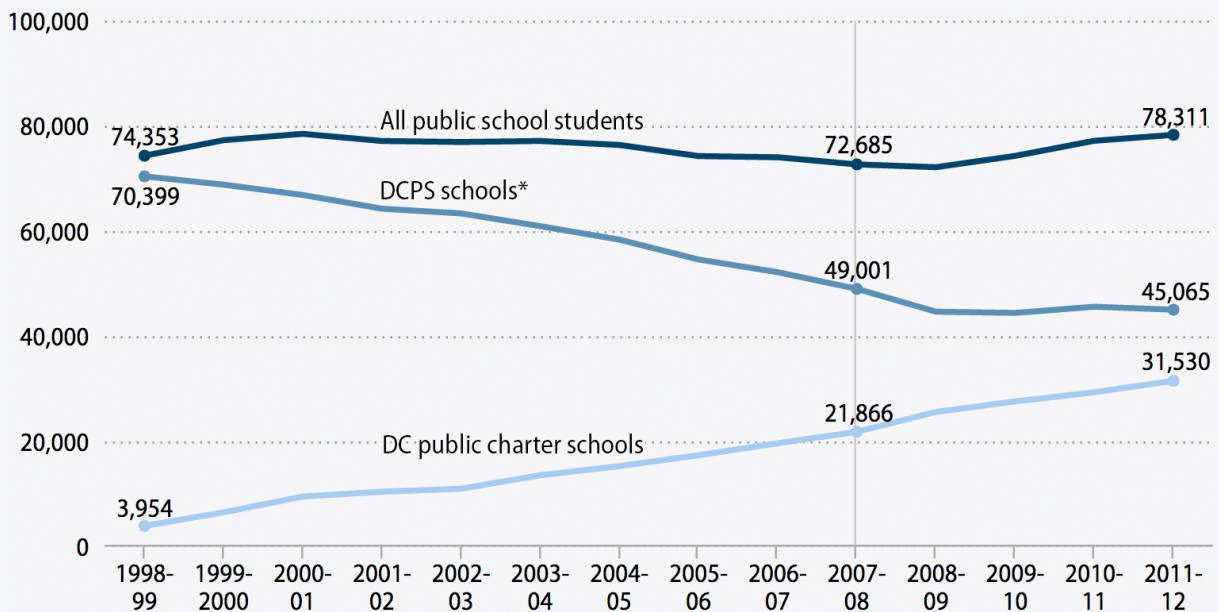
A 2009 Government Accountability Office (GAO) report, offered as testimony for Congress on the DCPS reforms under Rhee, provided other, data-driven reasons for concern. It found that “DCPS focused on a workforce replacement strategy to strengthen teacher and principal quality” (Ashby 2009, 8). As discussed above, without a strong plan to ensure high-quality new hires and avoid destabilization, such a strategy can backfire. The report found systemic flaws and inconsistencies from the start of Rhee’s term that could negatively affect teachers’ desire to stay in the district: The district implemented too many new “separate” programs or initiatives to boost student performance and, thus, was forced to refocus; it “failed to link employee expectations and performance evaluations to organizational goals”; though it “allowed principals to request changes to the staffing model based on their school’s needs, it did not establish or communicate clear guidance or criteria on how such requests would be treated,” resulting in city council and parent allegations of unfairness and lack of transparency; and it began hiring teacher coaches who were central to the 2009–2010 reforms later than expected, such that one in five positions was not filled on time, and “teacher coaches were often uncertain about their responsibilities and how to work with teachers, and received some conflicting guidance from principals.” In all, the report highlighted a lack of coordinated planning, follow-through, and transparency, the failure to communicate with relevant parties, and even exclusion of key constituencies, including parents, from much of the decision-making process (Ashby 2009, 10).²³

Principal turnover is no higher than before Rhee and Fenty controlled the system. Principals leave DC schools at a rate of about 20 percent each year, and 25 percent of schools open with a new principal (Levy 2012d). However, there are two respects in which Rhee’s execution of reform policies may have exacerbated the lack of strong leadership in high-needs schools. The new principals she brought in had particularly high attrition rates; of the 90 she recruited, the majority were already gone by the time she left in October 2010, and most lasted no more than two years (Levy 2012d). And a few particularly hard-to-staff and -serve schools had substantially greater turnover. Anacostia High School, transferred by Rhee to Friendship Public Charter Schools to manage, had six principals in six years (Brown 2012a). While such rapid turnover is high, it is unfortunately not unique; Dunbar High School, Hardy Middle School, MacFarland Middle School, Kelly Miller Middle School, and Powell Elementary School each had four principals in five years (2007–2012), and Johnson Middle School and Patterson Elementary School each had five principals in six years (Levy 2012d).

School closures. In addition to firing 1,000 teachers, Rhee closed 24 regular public schools and three special education centers during her three-and-a-half years as chancellor, asserting the need to save money in areas of declining enrollment. Overall student enrollment in the district did not fall during her tenure; rather, as in other “reform” cities and elsewhere, it continued to shift from regular public to charter schools. Charter schools enrolled 38 percent of D.C. public school students by the time Rhee left in the 2010–2011 school year, up from 30 percent when she arrived (Levy 2012a). At

FIGURE B

District of Columbia Public Schools' enrollment; all, traditional public, and public charter schools, from 1998–1999 to 2011–2012



*Excludes tuition students

Source: Authors' analysis of audited enrollment data for DCPS traditional and public charter schools, 1998–2012 (Levy 2012a)

ECONOMIC POLICY INSTITUTE

the same time, the decline in regular public school enrollment that had been the trend prior to Rhee's arrival leveled off. The result was a net total increase of over 4,000 students during Rhee's tenure, as illustrated in **Figure B**.

In alignment with the GAO report's finding that key DCPS decisions were made with insufficient public input, a Pew research study concluded that the closures happened too quickly and without community input, generating distrust (Samuels 2011). Moreover, while Rhee said that the students in the schools slated for closure were struggling, when their schools were closed, the students went, on average, to schools with even lower test scores and lower odds of making "adequate yearly progress" (AYP), as required under federal No Child Left Behind guidelines (Levy 2011b).

The projected cost savings that Rhee had cited likewise failed to materialize. An audit of the school closures made in 2008 found that "the actual cost to the city of shuttering the 23 schools was steep: \$39.5 million," quadruple the initial reported cost of \$9.7 million (Gartner 2012a). Unanticipated costs included additional moving expenses beyond those initially accounted for, 2010 demolitions of two elementary schools, patrolling of closed schools, transportation of students to further-away schools, and, heaviest of all, "capital asset 'impairment losses' of \$21.8 million, or the near-total loss of the value of eight buildings when they were no longer used as schools."

And while the current chancellor, Kaya Henderson, and Mayor Vincent Gray have announced plans to close dozens more schools, it will be hard to do so because, among other problems, young children will have no safe (or reasonably fast) way to get to their new schools (Scott 2012).²⁴ Many of the schools closed in the first round were close to others

that provided a viable alternative, but that is no longer the case; with duplicative neighborhood schools now closed, schools are increasingly farther from students' homes. The district provides school buses only for special education students whose Individualized Education Programs (IEPs) require it; other students beyond walking distance take public transportation, paying half the normal fare. Most schools are not near subway stations, and bus routes, designed principally to bring workers downtown, are not set up to serve students. Transportation is expensive for families with multiple children or with young children whose parents need to ride with them. The combination of distance and cost impedes attendance, particularly for secondary students, who have farther to travel, and teachers report that students sometimes fail to attend for lack of bus fare. For those already disengaged, the added barriers may tip the scales toward dropping out altogether. The mayor and chancellor should assess these considerations as they target additional schools for closure.

Charter schools. Another aspect of the reform platform is increased access to charter schools. As documented above, the proportion of DCPS students served in charter schools was on the rise before Rhee arrived and continued to increase during her time in office. However, while Rhee and Mayor Fenty had envisioned using their authority over DCPS to reopen the closed public schools as charter schools, they did not do so. Rhee did turn over operations of two high schools to New York-based charter operator Friends of Bedford: Dunbar High School in 2008 and Coolidge High School in 2010. She also transferred management of Stanton Elementary School in Ward 8 to Philadelphia-based charter organizer Scholar Academies in 2010 (Turque 2010b).

When the overhaul of Stanton was announced in June 2010, *Washington Post* education writer Bill Turque noted that Stanton, along with five other schools to be “reconstituted” for failing to make “Adequate Yearly Progress” under No Child Left Behind, “ha[d already] undergone waves of federally mandated improvement and restructuring... [yet four of the six] had not met annual progress benchmarks for at least seven years” (Turque 2010a).

None of the three schools turned over to charter operators improved their performance under new charter management, and, in fact, both high schools have since been reconstituted again for very poor performance. Coolidge High School “took [a] significant tumble” in test scores under the first year of Bedford management, with math scores on the city’s Comprehensive Assessment System (CAS) down 17 points and reading scores down 15 points from 2010 to 2011 (Lerner 2011). At Stanton Elementary, reading scores dipped from 13 percent proficient into the single digits after Scholar Academies’ takeover, and “[o]nly 9 percent of the 135 Stanton students tested were proficient or better in math, essentially unchanged from 2010” (Lerner 2011).

In December 2010, the *Washington Post* reported that Dunbar Senior High School “will get new administrators and extra security this week to quell violence and disorder, two years after the school was placed under a private management team by former chancellor Michelle A. Rhee” (Turque 2010b).²⁵ LaTanya Cherry, president of the Dunbar parent-teacher-student association (PTSA), was among those who called for the resignation of Bedford chief executive George Leonard, whose autocratic approach and disdain for input was noted in the *Post* article.

At Bedford Academy [in New York City], he said, discipline included the automatic suspension of any male student who cursed or disrespected a female. Leonard’s approach was also less than parent-friendly. According to a New York Times profile, he once told an audience: “Just stay out of my way and let me create the scholar, because you’re usually the problem. I’ll see you at graduation.” (Turque 2010b)

Central office growth. In addition to pledging to improve teacher and principal effectiveness through replacement, Rhee came into office vowing to pare down what she called the excessive size and inefficiency of DCPS central administration (DeBonis 2007). While there are indications that the central office runs better in some respects, such as increased responsiveness to parents and teachers, it has grown even larger than when Rhee arrived, and more expensive. Central office total staff, which had increased 17 percent, from 535 full-time employees (FTEs) in 2003 to 626 in 2007, rose by another 20 percent, to 775, by the time Rhee left in 2010 (Levy 2012b).²⁶ And the number of central office employees making \$100,000 or more, which had grown from just 10 in 2003 to 35 in 2007, shot up to 99 by the time she left (Levy 2012b).²⁷

New York City: Merit pay failed while school reshuffling produced mixed results

In New York City, a large-scale merit pay experiment proved a failure, and the combination of closing “failing” schools and opening new ones produced mixed results, with charters and new small schools potential, but controversial, bright spots.

Test-based teacher evaluations. Unlike Washington, D.C., New York City has yet to formally employ student test scores as a basis for firing individual teachers, though it has incorporated them into decisions about tenure, which has been granted less frequently in the past few years (Fleisher 2012a). However, test scores have been used in other decision-making processes, including both merit-based teacher bonuses and targeting schools for closure.

From 2007 to 2010, New York City spent \$50 million on awards to teachers whose high-needs elementary, middle, and high schools made strong progress in raising student test scores. In 2011, however, the city announced that it was ending the program after a RAND study, commissioned by the city’s education department, confirmed “mounting evidence that all those bonuses weren’t having much of an effect” (Sparks 2011). The Schoolwide Performance Bonus Program (SPBP), a joint initiative of the New York City Department of Education and the city’s teachers’ union, explored the use of bonus pay as a means to “motivate educators to change their practices to ones better able to improve student achievement” (Marsh et al. 2011, xix). Although it avoided some of the pitfalls associated with merit bonuses for individual teachers, the results disappointed officials who had backed the program. Specifically, SPBP did not improve student achievement in any grade level, had no effects on school progress report scores, had mixed results in creating the circumstances that foster success, and did not produce the intended effects on teachers’ reported attitudes and behaviors (Marsh et al. 2011, xxvii).

In their discussion of potential explanations for the lack of impact, the researchers rejected the notion that the program had too little time to produce effects, but they noted bumps in the implementation process that might have limited the program’s impact. They also cite broad research findings indicating that “there could be a flaw in the underlying theory of action of SPBP and school-based pay-for-performance programs more generally. As some have argued in the past, motivation alone does not improve schools” (Marsh et al. 2011, xxviii). Finally, already high levels of pressure to improve student test scores from other accountability systems likely made this one less potent.

School closures and new small schools. Student test scores have also been a measure used to target a school for closure due to poor performance. During the period explored in this report, schools were labeled “persistently low-achieving” if they posted low student test scores and/or graduation rates. The exact criteria were not always clear, however, as test score or graduation rate increases in some schools slated for closure were much higher than those of the city’s schools

overall (Cramer 2011). Some schools fought off closure only to be targeted again the following year. Under the state's No Child Left Behind waiver plan, the number of targeted schools in New York City has nearly doubled, with 123 "priority schools" now slated for closure by 2015 unless they dramatically improve their low graduation rates and test scores (Cromidas and Decker 2012).

Between 2002 and early 2012, the New York City Department of Education closed 140 schools, "all of which served the students with the highest needs" (Ansari 2012).²⁸ About 15 percent of those were large, comprehensive public high schools that were broken up into smaller, themed schools. Indeed, between 2003 and 2005, Bloomberg created 100 "new, small and academically rigorous" secondary schools across the city's five boroughs (Goldstein 2010). (Bloomberg and Klein also oversaw the establishment of roughly 100 new charter schools, as discussed in more detail later in this section.) The small schools did not have space for all of the students from the closed high schools; many students went, instead, to other, large comprehensive high schools.

The closures combined with the creation of new, smaller schools created decidedly mixed results as measured by high school student outcomes. A 2005 *New York Times* analysis found the graduation and college readiness rates of students in the smaller high schools "were no better and, in some cases, worse than those of larger schools" (Santos 2005). Recent research found more positive impacts from newer small schools, but mixed outcomes for the city's students and schools overall, and raised questions about whether the small schools can be replicated citywide. A report by three New York University professors found that newer small schools show much greater gains in achievement than their first-generation counterparts (Schwartz, Stiefel and Wiswall 2011). And a 2012 MDRC study funded by the Bill & Melinda Gates Foundation, which compared students who won entrance through a lottery with those who did not, found strong positive outcomes for students who moved to the new, smaller schools (Bloom and Unterman 2012). These included a 10 percentage-point increase in the share of ninth-grade students on track to graduate (from 48.5 percent to 58.5 percent) and a 7.6 percentage-point increase in college readiness in English, though no increase in math, as reflected in Regents exam scores (Bloom and Unterman 2012, Table 1, 4). The MDRC report also found that the new smaller schools were able to provide students with a unique set of resources.

[M]ost were founded with community partners who offer students relevant learning opportunities inside and outside the classroom and provide school faculty with additional staffing support and resources during start-up. For example, a law firm provides internships and embeds real-world examples in the curriculum for the Urban Assembly School for Law & Justice; and the Urban Assembly New York Harbor School is growing one of the largest oyster beds in the region.²⁹ (Bloom and Unterman 2012, 2)

Extra staff and other resources also enabled the schools to focus on the intensive personal teacher-student interactions that proponents of small schools see as likely contributors to those gains.

A less rigorous study conducted for the *New York Daily News* compared test scores for students in five poverty categories in new versus older schools, rather than comparing those who did and did not win entrance in the small school lotteries. The *Daily News* findings contradicted those of MDRC: "Of the 154 schools, nearly 60 percent were found to have lower [reading test] passing rates than the average school in the respective poverty group" (Huffington Post 2012).

These discrepancies in student outcomes across two reports highlight the potential to manipulate state and city test scores, and the need to interpret them cautiously. Because any test scores used to grade teachers or schools bring with

them incentives to raise scores by teaching to the test, narrowing the curriculum, etc., they should always be assessed skeptically. NAEP scores, which are unavailable for comparisons between types of schools, such as new and older schools, or charters versus traditional public schools, are the only truly reliable basis for comparison.³⁰ New York City NAEP scores, and changes in them over this period, are discussed in “The reforms have not improved student outcomes” section later in this report.

Let’s assume for a moment that the better outcomes in the smaller high schools reported in the MDRC study are reliable, and that the comparison employs equivalent student populations. The benefits reported in the MDRC report would need to be assessed against the impact of the changes on the well-being of the majority of students, who are not enrolled in the smaller schools. In particular, it is important to include any impact on students who moved to other, perhaps not better, schools. As of 2007–2008, the new small schools served roughly one in five New York City high school students; 239,000 of the 297,000 total attended other schools. As one analysis of the school closure and opening policies notes, “[m]ost of the students who would have attended the closed high schools were not admitted to the small schools that replaced them. Instead, most of these students were enrolled in other large comprehensive high schools, which consequently became academically overwhelmed, making them additional targets for closure” (Urban Youth Collaborative 2011, 9).

A 2009 report by The Center for New York City Affairs assessed the broader, district-wide impacts of the small schools and found that students who were not enrolled suffered real harm.

An analysis of 34 large high schools in Brooklyn, Manhattan and the Bronx (defined as those with more than 1,400 students in 2007–08) found that 26 saw their enrollments jump significantly as other high schools were closed. Enrollment increases ranged from 150 to more than 1,100 students. Of these 26 schools, 19 saw their attendance decline and 15 saw their graduation rates decline between the fall of 2002 and the spring of 2007. Fourteen saw both attendance and graduation rates decline.³¹ (Hemphill and Nauer 2009, 35)

Critics of school closures contend not only that they are detrimental to students but that the mechanisms employed to determine closures are unfair because they fail to account for inequities in school resources. Historically, a major source of inequity has been the lack of qualified, experienced teachers in schools serving low-income students. A recent report by the Schott Foundation for Public Education found that Community School Districts in New York City with higher poverty rates had fewer experienced and highly educated teachers and less stable teaching staffs. As is true in low-income districts across the country, this pattern results in the least qualified and experienced teachers serving students who have the greatest needs. The report describes a pattern of “education redlining,” with resources, staffing, and opportunities varying by race, income and neighborhood, and noted, too, that “[t]he teachers who are expected to perform miracles are paid the least” (Holzman 2012, 11).

A 2011 report by the Urban Youth Collaborative assessing student outcomes among students in the high schools that had completed phase-out since 2000 noted that “[t]he students who attended the 21 closed high schools, almost all of whom were Black and Latino, had significantly higher needs and were much more academically under-prepared than the students across the city’s school system,” as measured by eligibility for free lunch, English language learner status, being over-age for grade (generally synonymous with having been held back in grade at least once), and being below

grade level in reading and math (Urban Youth Collaborative 2011, 6). The evidence also showed that the closures were seemingly intentional.

Instead of intervening aggressively to help the lowest performing schools improve, the DOE has consistently neglected to provide the comprehensive guidance and supports that struggling schools need. Reports from the NY State Education Department (SED) on 17 schools identified by the state as Persistently Low Achieving (PLA) found that at least 14 of the schools were not provided the assistance from the DOE necessary to raise student achievement. Furthermore, SED reviews of the 11 schools currently implementing the federal transformation model found that the DOE had largely not met their commitment to guide and support the school transformation plan. (Urban Youth Collaborative 2011, 7)

Unlike the process in some other cities, school closures in New York City are gradual, so students and teachers remain in a school marked for closure for several years before it is closed. News reports over the past few years describing the reactions of staff and students in schools targeted for closure depicted absorption of the highest-risk students from other schools that had recently closed and rapid loss of staff and other resources. As early as 2007, in Canarsie High School in south Brooklyn, targeted for closure, “teachers said that as other large high schools in southeast Brooklyn closed, Canarsie became a dumping ground for struggling students. Administrators said the school became overcrowded, with not enough resources” (Medina 2007).

Recently, Legacy High School for Integrated Studies in Manhattan gradually lost staff, electives, and afterschool programs as it shut down. The father of a 19-year-old special education student described the damage done to his daughter when the school’s only social worker was excessed (Powell 2012b). As the reporter noted, “So far, officials have closed 140 schools, which they routinely describe as failing, and replaced them with smaller schools and charters, which they routinely describe as making ‘historic gains.’ Perhaps this is so. But for tens of thousands of children who live in the purgatory of schools marked for closing, boasts of an education revolution bring little comfort.”

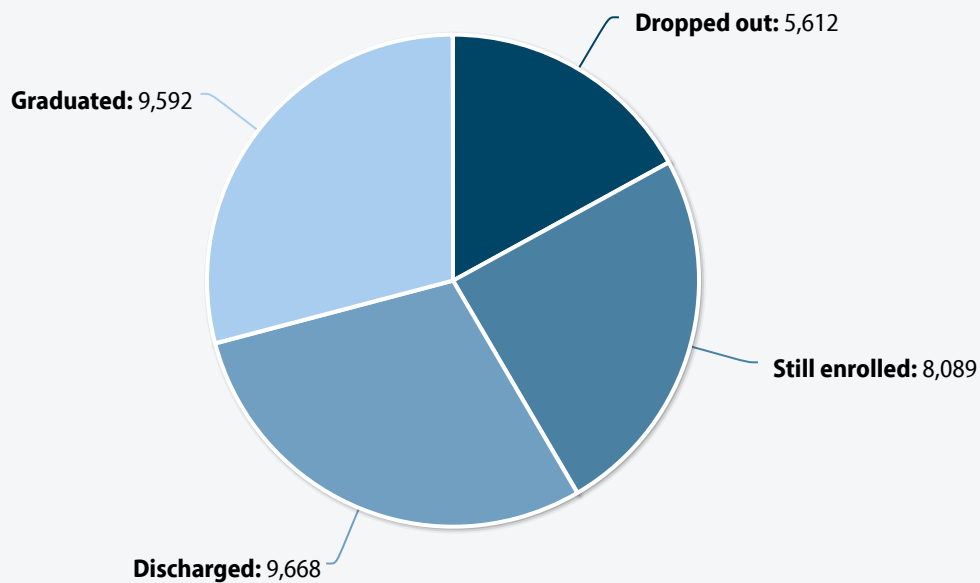
Among those caught in the middle were likely the students in the 21 high schools assessed by the Urban Youth Collaborative, where discharge and dropout rates “skyrocketed” between the year a school was slated for closure and the year it closed: from 25 percent to 70 percent at Taft High School, and from 33 percent to 55 percent at Morris High School. Across all of the schools, of 32,961 students who attended during those years, 5,612 dropped out, 8,089 were still enrolled when the school closed, 9,668 were discharged, and 9,592 graduated (Urban Youth Collaborative 2011, 6). See **Figure C**.

As illustrated in the top panel of **Figure D**, annual reports by New York City’s Independent Budget Office (IBO) comparing high schools slated to be closed with all high schools illustrate the challenges faced by the former, including substantially higher percentages of poor students (77.1 percent versus 67.2 percent), black (40.4 percent versus 31.6 percent) and Hispanic (55.4 percent versus 39.2 percent) students, and over-age students (10.3 percent versus 6.1 percent) (IBO 2012b). One exception is English language learners (ELL), who comprised nearly the same proportion in each type of school—13.7 percent in schools to be closed, versus 12.2 percent overall.

“But for tens of thousands of children who live in the purgatory of schools marked for closing, boasts of an education revolution bring little comfort.”

FIGURE C

Outcomes for 32,961 students at 21 NYC high schools closed since 2000



Note: This chart focuses on 21 high schools that had been phased out and closed and had reported full student outcome data between 2000, when the New York Department of Education announced the first school closings, and June 2009, when the Urban Youth Collaborative launched its study. The outcomes show what happened to the 32,961 students during the four years their school was being phased out.

Source: Authors' analysis of Urban Youth Collaborative (2011, 6)

ECONOMIC POLICY INSTITUTE

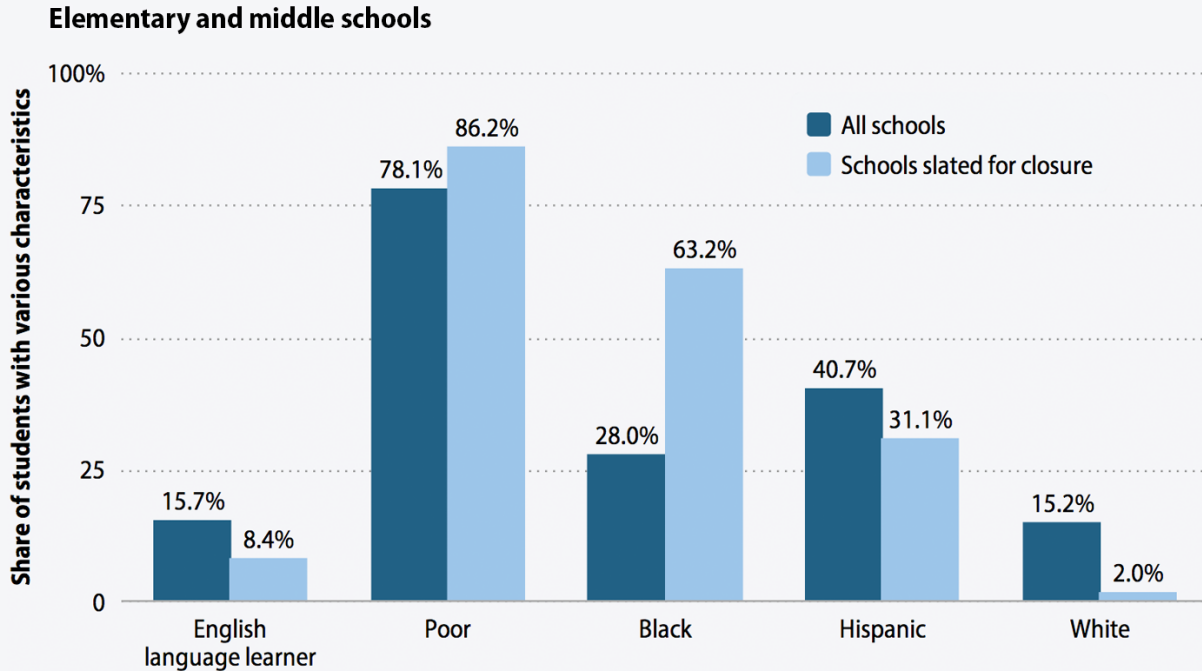
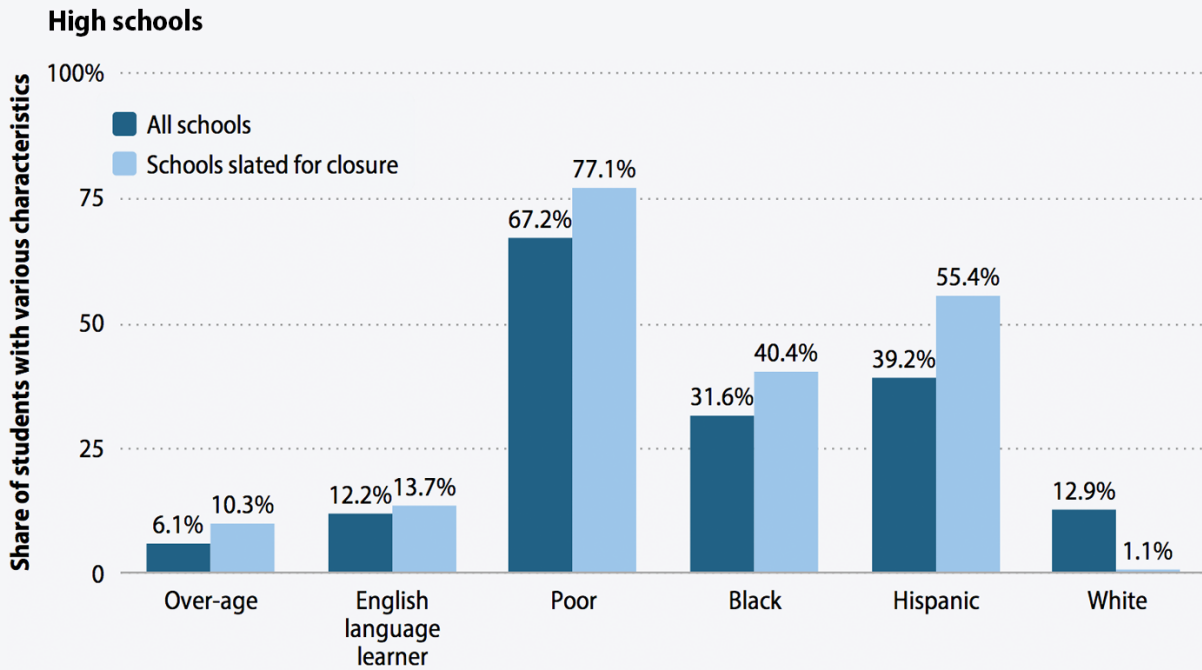
School closures did not only affect high school students. There is substantial evidence that New York City schools targeted for closure—both primary and secondary—were “set up” to fail, and that the closure process itself hurt students enrolled in the schools. As the bottom panel of **Figure D** illustrates, while elementary and middle schools slated for closure had fewer ELL and Hispanic students than the city average—8.4 percent versus 15.7 percent, and 31.1 percent versus 40.7 percent, respectively—they had higher percentages of poor students (86.2 percent versus 78.1 percent), over twice as many black students—63.2 percent versus 28.0 percent, and virtually no white students, just 2.0 percent compared with the city average of 15.2 percent.

Like Rhee’s successor Henderson, Bloomberg has promised more school closures before he leaves office in 2013 (Epstein 2012). However, he experienced a setback in June 2012, when an arbitrator ruled against his plan to close 24 more schools under the federal turnaround process, replacing half the staff and reopening the schools with new programs and personnel (Baker 2012).

Charters and co-location. As stated earlier, Bloomberg and Klein oversaw the establishment of roughly 100 new charter schools during their joint control of the city’s schools, often employing the controversial practice of “co-locating” new charter schools in the same building as existing public schools.³² At an April 2012 city council hearing, Education Committee Chair Robert Jackson testified that co-locations were disruptive to learning: “Problems can arise such as

FIGURE D

Challenges confronting New York City schools, all schools vs. schools slated for closure



Note: Over-age refers to students who are older than the expected age for that grade; normally this indicates that the student has been retained in grade one or more times. Poor students are those who are eligible for free or reduced-price meals.

Source: Independent Budget Office (2012b)

overcrowding, unsafe hallways, inadequate resources, friction over shared space, and a climate of mistrust and conflict” (Cromidas 2012).

Observers also have raised concerns about inequities in resources for public schools sharing buildings with charter schools. For example, critics claimed that students and teachers in Eva Moskowitz’s Success Academy Charter Schools receive more resources, space, and favorable treatment than public school students and teachers sharing the building (Powell 2012a). The merits of such claims are difficult to assess, given limited access to detailed data, and the greater share of expensive students—very low-income students, English language learners, and those with other special needs, in particular—in regular district schools than in charter schools. However, a 2010 report by New York City’s Independent Budget Office, which employed sophisticated methods to estimate per-pupil funding for students in three different types of New York City schools (traditional public schools, charter schools located in public school buildings, and charters located independently), found that charters co-located in public school buildings had the highest per-pupil funding (Domanico and Smith 2011). It estimated the “value of the savings from co-locating for these charter schools on facility, utility, and school safety costs” at \$2,712 per pupil.

Moreover as another report, by two Rutgers University professors, found, the added funds were not distributed equitably; there was no evidence that schools receiving facility “subsidies” had higher student needs, and thus costs (Baker and Ferris 2011). But the real disparities appear to be far greater. The 2011 report discovered that large private donations greatly expanded the public-charter school funding gap. While funding, like other aspects of charter schools, varies tremendously, the researchers found that a small cluster of charter schools had funding substantially higher than the district average and that two—KIPP Academy and Opportunity Charter—spent over \$25,000 per pupil. These levels of spending were enabled by large grants from foundations such as the Bill & Melinda Gates and Walton Family foundations.³³ When comparisons were made between charter and nearby public schools that serve substantially more low-income and special-needs students, the differences were stark. Per-pupil spending at the public A. Philip Randolph school, \$14,123, was comparable to nearby Harlem Link charter, \$13,105, but the former serves a student body in which 81 percent qualify for free lunch and 12 percent are English language learners versus 70 percent qualifying for free lunch and just 1 percent ELL at Harlem Link. At Samuel Stern public school, where 86 percent of students qualify for free lunch and 19 percent are ELL, per-pupil spending is \$12,476. At nearby Harlem Day charter school, with 62 percent of students qualifying for free lunch and no ELL students, per-pupil spending was \$19,632 (Baker and Ferris 2011, 24).

In May 2011, the United Federation of Teachers (UFT), the NAACP, and other organizations sued to stop school closings and the co-location of schools. The New York City Parents Union also filed a lawsuit, asserting that the NYC education department “continues to push public school parents into a privately managed school system where many charters perform worse than the public schools while failing to serve their fair share of students with special needs and English Language Learners” (New York City Parents Union 2011). While the overall data on charter school enrollment show that charter school students are more likely to be low-income and minority than their public school peers, regular public schools serve the lowest-income students and the most-expensive students (English language learners and those with special needs):

- Charter and public schools serve virtually the same share of students eligible for free or reduced-price lunch (75.6 percent at charter schools versus 75 percent in district schools in 2010–2011), but district schools serve slightly

more of the poorest students (67.6 percent versus 65.2 percent free lunch), and fewer reduced-price lunch students (7.4 percent versus 10.4 percent) than do the city's charter schools. Compared with district schools in the same community school district (CSD), fewer than a third of charter schools have an equal-or-higher percentage of subsidized-lunch-eligible students (Merriman et al. 2012, 25).

- Citywide, charter schools serve a lower proportion of special-needs students (students with IEPs), 12.9 percent versus 14.4 percent, with much larger gaps in the K-8 years in which the vast majority of charter schools operate. This gap also holds regionally: 23 percent of charter schools are at or below the 10th percentile with respect to students with IEPs in the same community school district (Merriman et al. 2012, 26–27).
- The charter-public gap is much larger for English language learners; the city's regular public schools serve nearly three times as many ELLs—15 percent of the student body—as do the charter schools, where ELLs make up just 5.8 percent of the student body (Merriman et al. 2012, 28).
- While they serve more African American students than comparable district schools, charter schools serve far fewer Hispanic students than district schools in their CSD. Citywide, charter schools are 30.9 percent Hispanic, whereas district schools are over 40 percent Hispanic, and 40 percent of charter schools are at or below the 10th percentile in the same CSD in enrollment of Hispanic students (Merriman et al. 2012, 29).

Chicago: Neighborhood school replacement increased instability but not performance

Firing hiring, and rewarding teachers. During Duncan's eight-year tenure, Chicago Public Schools laid off 1,300 teachers, citing budget shortages. CPS never provided teachers with the hearing required for due process, and in October, 2010, a group of teachers, including 749 who had tenure, won a discrimination suit against CPS. In a complaint supported by University of Chicago research, the teachers alleged that they were being replaced with less experienced, younger, and whiter teachers (de la Torre et al. 2012). U.S. District Judge David H. Coar held that CPS had violated the rights of over 1,000 tenured Chicago teachers and ordered the district to work with the union to recall them. He noted that most of the laid-off teachers had not received unsatisfactory reviews (Ahmed 2010).

This loss of teachers has been exacerbated by a high rate of teacher turnover that is especially problematic in schools serving low-income minority students. According to a 2009 University of Chicago report on teacher instability in Chicago, “these [seemingly average] one-year stability rates hide a sobering statistic—within five years, the typical CPS school loses over half of its teachers. Many schools turn over half of their teaching staff every three years” (Allensworth, Ponisciak, and Mazzeo 2009, 1). In 100 mostly low-income minority schools, turnover was even higher, with the schools losing one-fourth or more of their teachers annually. Because comparable data are unavailable for prior years or for other urban districts, it is impossible to determine whether this represents an increase in the turnover rate during Duncan's tenure, or whether the rate is higher or lower than in other large city school systems.

At the same time, recognizing the sharp inequities in the distribution of qualified, effective teachers across lower- and higher-income schools, Duncan employed a multifaceted strategy to improve the quality of the CPS teacher pool. Changes included giving schools more capacity to hire earlier, in an effort to make them more competitive with their suburban counterparts in attracting better teachers, and assessing which teacher-preparation programs were producing the best teachers, so that CPS could recruit from them. An Arizona State University report also notes Duncan's launch of a marketing campaign to better brand CPS schools and new pipeline programs to attract strong new teachers, including

the establishment of job fairs to bring together principals and prospective teachers, and offers of higher starting salaries for those new teachers. These efforts appear to have paid off; while the state of Illinois overall had modest but significant improvements in teacher qualifications, driven by improvements in the lowest-performing schools, Chicago's gains were much greater. According to the report, "On average, teacher qualifications in Chicago schools were more than one standard deviation lower than they were in schools in all other locales," but the sharp increase in scores among Chicago's schools "narrowed the gap in average teacher qualifications between Chicago schools and schools in non-urban localities in Illinois by approximately 30% during the six-year period [from 2001–2006]" (DeAngelis, White, and Presley 2010, 10–11).

So, unlike other policies implemented during Duncan's tenure, these changes disproportionately benefited the city's highest-risk and neediest students.

In terms of teacher qualifications, Chicago schools showed noteworthy average improvements from 2001 to 2006 in all but its most economically-advantaged schools, none of which fell into the top quartile statewide. Moreover, the improvements were greatest in its highest poverty and highest minority schools, which resulted in a more equitable distribution of teacher qualifications across schools within Chicago itself. (DeAngelis, White, and Presley 2010, 13)

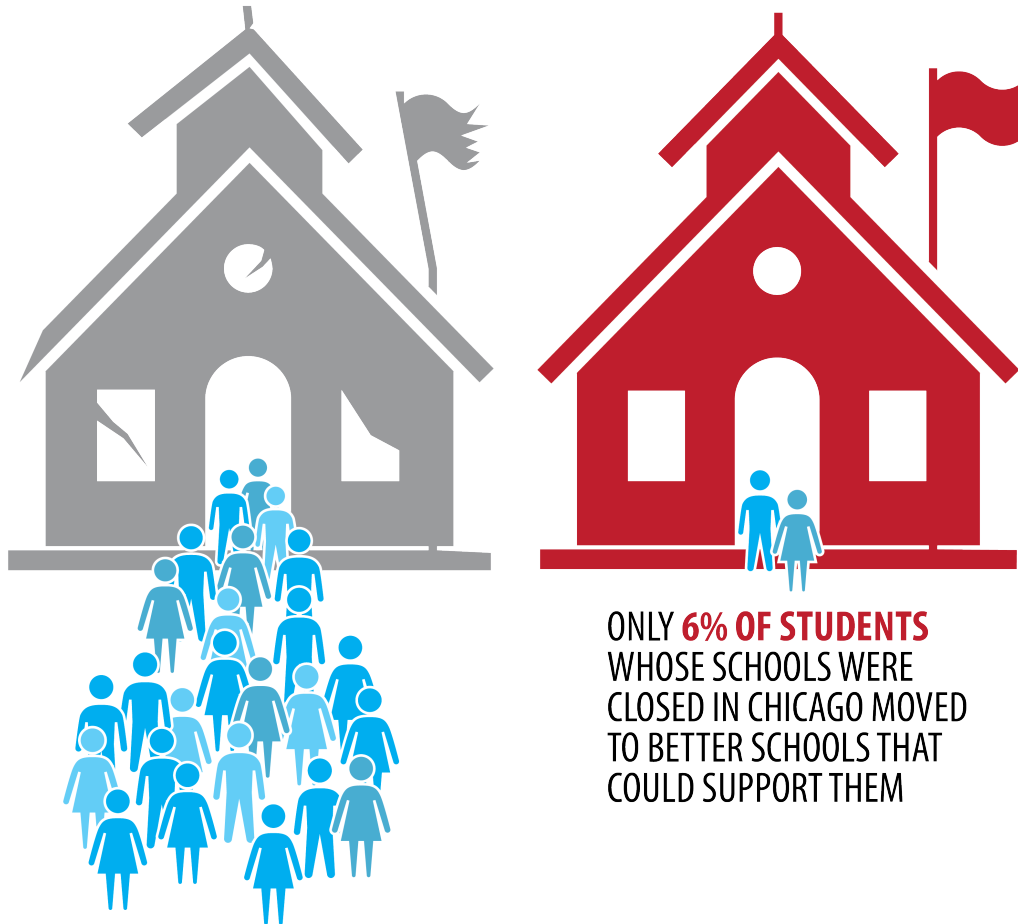
This improvement was the result of both a 70 percent decline in the percentage of teachers lacking full certification and a one-third standard deviation improvement in the average teacher ACT score. The researchers attributed these improvements to a combination of national pressure under No Child Left Behind to place more highly qualified teachers in low-income schools and classrooms, state-level efforts to make teaching in schools and on subjects with shortages more attractive through alternative certification, higher standards for the state licensure exam, and Duncan's Human Capital Initiative. The latter, launched in 2001, consisted of three components: recruitment (particularly credited with fostering improvement), district human resource operations, and talent management.³⁴

A smaller teacher-improvement initiative championed by Duncan proved less successful. Though Chicago's federally funded merit-pay program was designed quite differently from that of New York City, it, too, had disappointing results. The Chicago Teacher Advancement Project (Chicago TAP) placed less emphasis on bonuses and awarded smaller ones, and it included a strong career ladder and teacher improvement component. Nevertheless, after two years, Chicago TAP showed no effect on student test scores or teacher retention, the main objectives of the program (Glazerman and Seifullah 2010). A final report on the program after four years also found no impact on student achievement, though there was some increase in school-level teacher retention (Glazerman and Seifullah 2012).

School closures, selective magnets, and turnarounds. In 2004, Chicago Public Schools' CEO Arne Duncan and Mayor Richard Daley announced the Renaissance 2010 reform program, which had at its core a focus on closing "failing" schools and replacing them with better schools, including charter and selective-enrollment magnet schools.³⁵ (Later in his tenure, Duncan continued to close schools but also shifted to a turnaround strategy for some schools, keeping students in 11 troubled schools but changing the entire staff, including the principal and all teachers (Hill 2010)).³⁶ While Duncan closed schools that CPS deemed "failing," more recently, CPS has cited underenrollment as a reason to close more schools in 2013–2014 (Caref et al. 2012). Research shows that these test-score-based closures exacerbated instability and alienated parents, without improving overall student or school performance.

SENDING STUDENTS FROM ONE 'LOW-PERFORMING' SCHOOL TO ANOTHER DID NOT IMPROVE ACHIEVEMENT

Reforms closed “failing” schools in DC, NYC, and Chicago, promising better student outcomes. But students stagnated or lost ground, as new schools were no better, and moves also meant instability and longer (and sometimes more dangerous) commutes.



ONLY 6% OF STUDENTS
WHOSE SCHOOLS WERE
CLOSED IN CHICAGO MOVED
TO BETTER SCHOOLS THAT
COULD SUPPORT THEM

Source: Marisa de la Torre and Julia Gwynne, *When Schools Close: Effects on Displaced Students in Chicago Public Schools*, Consortium on Chicago School Research, 2009.

From 2001 to 2008, CPS under Duncan’s leadership closed dozens of “failing” elementary and high schools and replaced them with (nonunionized) charter, contract, performance, and military schools (Dillon 2009 and Lipman and Gutstein 2011). A 2009 report from the Civic Committee of the Commercial Club of Chicago, a founder and leading supporter of Renaissance 2010, concluded that the initiative had spurred little progress toward its stated goals: “Most of Chicago’s students drop out or fail. The vast majority of Chicago’s elementary and high schools do not prepare their students for success in college and beyond” (Civic Committee of the Commercial Club of Chicago 2009, 2). Researchers from the Consortium for Chicago School Research (CCSR) at the University of Chicago confirmed this lack of progress. Their 2011 study found that most of the students who transferred to different schools as a result of the closures did not improve their performance (Lupescu et al. 2011).

An earlier CCSR study found that school closings under the Renaissance 2010 initiative had almost no effect on student achievement because most elementary school students who were displaced by closing schools transferred from one low-performing school to another. While there were very few high-performing schools to receive students, the “small number of students who transferred to academically strong receiving schools and found supportive teachers made significant gains” (de la Torre and Gwynne 2009).³⁷ The 2011 CCSR researchers also attributed some elementary school problems to the closing and opening of many schools under Duncan (Hood and Ahmed-Ullah 2011). Students experiencing multiple closures and their schools were particularly negatively affected: “CPS transferred some students in the Mid-south [area of the city] to as many as four schools in three years as the district closed one school after another. Receiving schools were also destabilized by the influx of dislocated students” (Lipman and Gutstein 2011, 20).

Additionally, how, why, and where schools were closed and reopened has exacerbated existing gaps in student achievement. As in New York City, there have been allegations that Chicago schools closed for low performance were not provided with the resources needed to succeed, and thus “set up for failure” (Lipman and Person 2007, 6). A Chicago Teachers Union report, for example, noted that nine of the 10 CPS high schools with the lowest Educational Planning and Assessment System (EPAS) scores were targeted for drastic action, despite the fact that “the test is given in September of students’ freshman year, [so] high schools have practically no impact on how students fare on the [test]” (Caref et al. 2012, 11). (EPAS is a three-part system CPS employs to assess students’ college readiness. It uses ACT test scores, which are a widely accepted benchmark.)

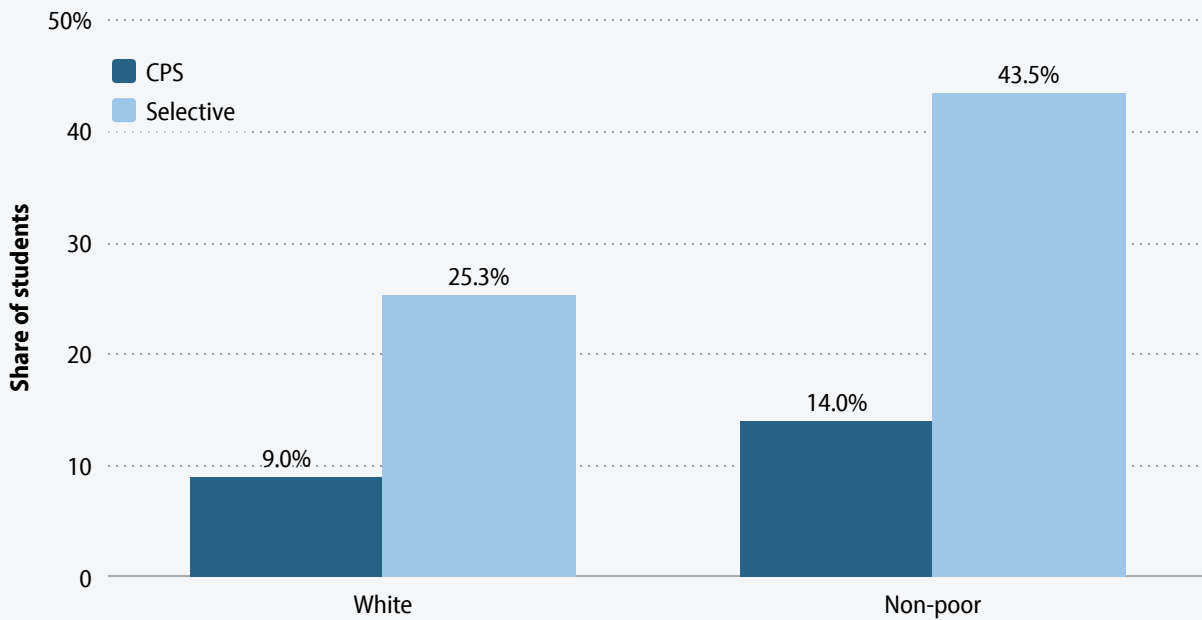
The report also pointed out that the schools targeted for drastic action were intensely poor (in all but one, over 90 percent of students qualify for subsidized school meals) and overwhelmingly black (in all but one, the student body is over 95 percent African American). Yet no CPS policy seeks to ameliorate these root causes of low achievement, the report noted (Caref et al. 2012, 11). Indeed, The CCSR report warned that high schools were being “saddled” with incoming freshman classes “woefully unprepared” for the rigors of college preparatory work (Luppescu et al. 2011). Instead of improving neighborhood schools, CPS has replaced them with charter schools, a key focus of Renaissance 2010, which raised \$47 million through 2009 for these new schools.

This pattern of school closings (correlated with race- and income-based achievement gaps) was accompanied by declining parental and public input in education policy decision-making. While the right balance between parent control and input from other decision-makers is admittedly difficult to determine, the replacement of an elected school board with one appointed by the mayor seems to have swung the pendulum too far toward outside input. As a 2011 report produced by the Collaborative for Equity and Justice in Education at the University of Illinois at Chicago asserted,

Chicago’s mayor-appointed board is comprised of elite decision-makers who are neither representative of the student population of CPS nor directly accountable to the public. Board structures and processes severely limit public input in decisions. The Board is composed primarily of corporate executives, while the district is 92 percent students of color and 86 percent low-income students whose communities have no role in school district decisions. This is problematic because perspectives and knowledge of parents, educators, and students are essential to good educational decision-making. It is evident that community members feel that the Board is unresponsive to their input and concerns. Parents have felt it necessary to take extreme measures to be heard, including candlelight vigils, marches, campouts in front of Board headquarters, a hunger strike, and a recent 43-day occupation of a school field house to get a long-needed school library. (Lipman and Gutstein 2011, 3)

FIGURE E

Chicago Public Schools’ high school enrollment, white and non-poor students, overall vs. in selective-enrollment magnet schools, 2010–2011



Note: Non-poor students are those who are not eligible for free or reduced-price lunch.

Source: Authors’ analysis of Lipman and Gutstein (2011)

ECONOMIC POLICY INSTITUTE

Most school closures were in gentrifying black and Hispanic neighborhoods, as demonstrated by the sharp increase in housing prices in these neighborhoods (Lipman and Gutstein 2011, 21). At the same time, many of the new selective-enrollment magnet schools were opened in neighborhoods that were either gentrifying or affluent. Because the new schools did not admit all students from ones that were closed, however, the combination of closures and new school openings left many students without a nearby option. Chicago’s eight selective-enrollment high schools are very limited in their reach; only one in 10 CPS high school students attends them. Furthermore, they do not represent the CPS student body. As **Figure E** shows, the CPS high school student body as a whole was 9.0 percent white and 14.0 percent non-low-income in 2010–2011, while these selective high schools were 25.3 percent white and 43.5 percent non-low-income (Lipman and Gutstein 2011, 13).

These data also demonstrate that, for the small number of low-income students who attend them, benefits of selective-enrollment high schools include not only the enhanced curriculum and resources they provide, and the stronger teachers they likely attract, but also a more integrated student body. Numerous studies have shown that this advantage, which is very rare within CPS and most urban districts, benefits minority and low-income students in particular (Kahlenberg 2012; Rothstein and Santow 2012). The pertinent questions, then, include whether the selective schools serve a sufficiently large number of students to make a real difference, and whether their addition has had negative impacts on the district as a whole that counter the benefits they confer to the students who attend them.

As schools are closed, more students have to travel long distances to get to their new schools. Among charters that opened under Renaissance 2010, the share of elementary school students who commute six miles or more increased from 9 percent in 2004 to 13 percent in 2008; over the same period, the share of high-schoolers commuting that distance increased from 10 percent to 15 percent (Lipman and Gutstein 2011, 20). This change is partly to blame for increasing youth violence and gang activity in Chicago neighborhoods, since closures disrupt “safe routes” and other antiviolence strategies (Steinberg, Allensworth, and Johnson 2011). Indeed, community organizers and parents protesting school closings warned the city that these shifts could lead to greater gang violence (Ryan 2012). In September 2009, a spike in violence at Fenger High School after students were transferred there from a school that was turned into a selective-enrollment military academy led to the highly publicized death of a student.

Aware of the disruptions caused by the school closings, Duncan added a second strategy in 2006 to improve educational opportunities for low-income students—comprehensive school turnarounds. He began to work with the principal of one of the city’s high-performing selective-enrollment schools, Jones College Prep, to establish a network of low-performing schools targeted for special assistance, specifically replacement of the entire staff followed by intensive training (Hill 2010). This new network built on the work of the independent Academy for Urban School Leadership, established in 2001. In addition to extensive training for new staff, Chicago’s Office of School Turnaround has provided “ongoing management of the school . . . [and] ongoing and extensive technical assistance and coaching on the school level.” Initial results include improved attendance in some schools, improved student behavior, and an increasing number of high school students on track to graduate.³⁸

Charter schools. Charter schools decreased segregation for some students, but, as in New York City, co-location proved disruptive, and systemic costs were high.

While the overall composition of charter and traditional public school students is similar with respect to race and ethnicity, poverty, and school performance, black and Hispanic students experience greater racial diversity and thus less segregation in charter schools than in traditional public schools. As discussed earlier, research shows that integrated schools confer substantial benefits, especially for minority and low-income students. However, these benefits are also conveyed by nonselective magnet schools. And magnets do not require the difficult application process, including parent contracts and often long essays, that many charter schools use; as such, they do not “select out” the highest-needs students (Caref et al. 2012). As discussed later in the section on student outcomes, magnets far outperform their charter school counterparts among comparable student bodies.

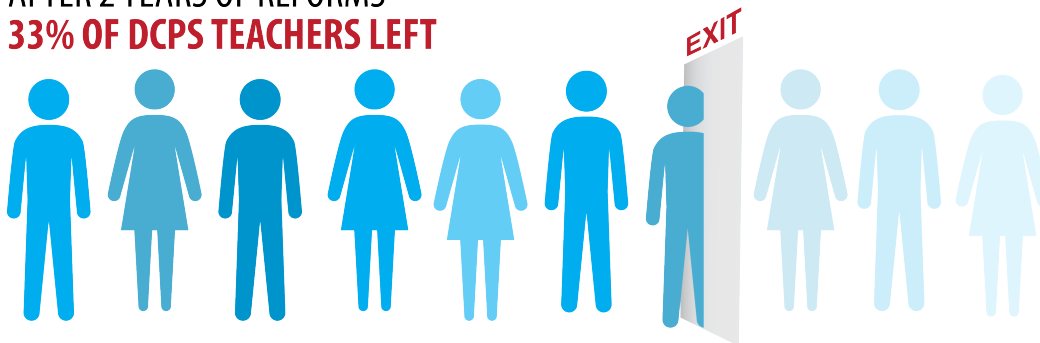
Moreover, as in New York City, preferential treatment for some charter schools has negative impacts on other schools and on the entire system. For example, because charter schools are allowed to decide where to operate, CPS forces neighborhood schools to share facilities with a charter and sometimes a series of charters as charters move from one space to another in search of the best option (Caref et al. 2012). Also, the city has paid to rehabilitate buildings deemed unusable for students in neighborhood schools.

Charters also benefit from sweetheart facilities deals, often after a traditional public school has been closed and the building turned over to a charter operator. CPS routinely leases buildings to charter operators for the bargain-basement market rate of \$1 per year. At least thirty-three charter campuses currently benefit from \$1 leases. (Caref et al. 2012, 23)

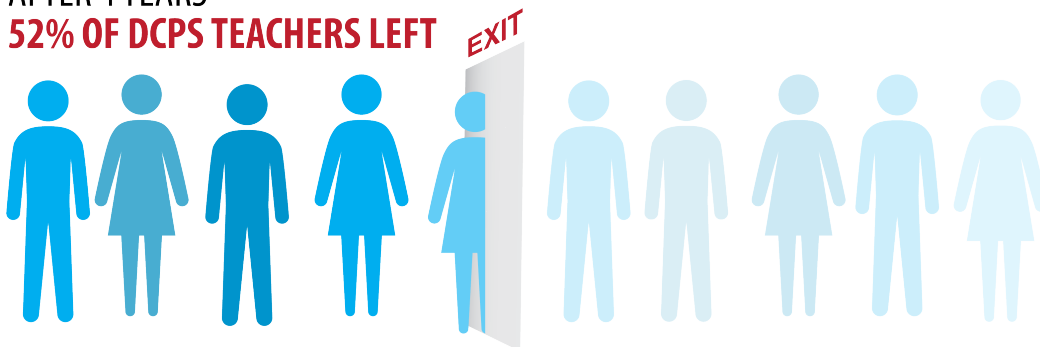
USING TEST SCORES TO EVALUATE EDUCATORS AND SCHOOLS LED TO TEACHER CHURN AND INEXPERIENCE

Reforms in DC, NYC, and Chicago that used student test scores to evaluate, reward, and fire teachers and to target schools for closure delivered increased turnover and fewer experienced, qualified teachers, but no improvement in student achievement.

AFTER 2 YEARS OF REFORMS 33% OF DCPS TEACHERS LEFT



AFTER 4 YEARS 52% OF DCPS TEACHERS LEFT



Source: DCPS teacher retention data provided by D.C. budget consultant Mary Levy in 2012. Levy compared turnover in 2001–2007, before DCPS’s IMPACT evaluation program began (in 2009) with turnover in 2008–2012.

The bottom line: Highlighted reforms in the three cities produced upheaval, overshadowing smaller promising programs

Students, parents, and teachers in all three cities experienced substantial churn as their neighborhood schools were closed, experienced teachers were fired, new and often less experienced teachers replaced them, and students were forced to take longer, sometimes dangerous routes to school.

Indeed, in Washington, D.C., teacher turnover was higher under Rhee and IMPACT than ever before, but apparently not due largely to the elimination of bad teachers, but rather to an increase in voluntary exits by experienced teachers as well as to their “excessing” and replacement by novices.

The cumulative evidence suggests that districts seeking to substitute novices for experienced teachers, and to close schools in order to save money or attain better outcomes, should be skeptical regarding promised gains and aware of the substantial harms that can result.

School closures based on test-score-based labels of “failure” can decrease, rather than improve, odds of student and school success. All three districts failed to protect their most at-risk students from the damage that can occur due to school closures; those students experienced more upheaval and less support in the process and lost ground as a result. For example, in both Washington, D.C., and Chicago, students ended up in schools that, on average, had lower test scores than those they left. In all three cities, closures affected black and low-income students disproportionately, with New York City’s neediest students suffering a gradual loss of staff and other resources as their schools were closed.

In all three cities, reformers worked to bring more middle-class students into their districts. However, the degree to which schools grow more integrated, which would benefit minority and lower-income students, especially, remains to be seen. Currently, the new students are largely clustered in a small group of charter and/or selective-enrollment schools. Moreover, test score changes occurring amid shifting demographics must be interpreted with an even more skeptical eye than usual. As discussed later, the small gains in NAEP scores in Washington, D.C., appear to have resulted largely from new, higher-scoring students entering the system, rather than actual improvements occurring among existing students of any racial or socioeconomic group.³⁹

Charters, a core component of the reform agenda, have both positive and negative aspects, but they do not solve the broad set of problems confronting low-income students, and to the extent that some charters are doing substantially better, discovering how and why they are different is important.⁴⁰ Like many of the small and magnet school benefits however, some of the charter benefits may not be replicable. For example, using funding to extend school days and provide enriching extracurricular activities may not be feasible in larger numbers or based on public school budgets. Parent contracts and commitments cannot necessarily be rendered from most parents, nor are all parents capable of providing the consistent support that benefits some charter school students disproportionately. None of this is to disparage the real gains made by some charter school students, but rather, to caution that such benefits are unlikely to be available to the broader student body and, particularly, to the minority of students with the greatest needs, very few of whom are currently served by charters.

Of course the changes focused on here—test-score-based decisions on teachers and schools, closures of “failing” schools, and expanded access to charters—were not the only ones made. But they drew most of the attention, leaving other, much more promising policies without the resources to expand. In Chicago, for example, the Renaissance 2010 initiative eclipsed and to some extent impeded the successes of other, often more effective, reforms, according to Barbara Radner, who heads the Center for Urban Education at DePaul University.

There has been some good and some bad in Renaissance 2010, but overall it wasn’t the game changer that people thought it would be. ... In some ways it has been more harmful than good because all the attention, all the funding, all the hope was directed at Ren10 (sic) to the detriment of other effective strategies CPS was developing.” (Banchero, Germuska, and Little 2010)

One of these effective but overlooked strategies grew out of Arne Duncan’s recognition that the city was having a hard time attracting the strong credentialed teachers that disadvantaged students particularly need. So he implemented a set

of strategic recruitment and support policies. Researchers documented their success while Duncan was still CEO, yet these policies have been absent in his federal efforts. Indeed, both Race to the Top and No Child Left Behind waivers emphasize test-based strategies to evaluate those teachers already in schools, with a major goal of firing those deemed “ineffective.” Teacher credentials are often dismissed as meaningless, with value-added scores seen as the proper metric.

In all three cities, the focus on the three main reforms appeared to divert attention from the need to address socioeconomic factors that impede learning. With student eligibility rates for free or reduced-price meals of 67 percent in Washington, D.C., 72 percent in New York City, and 77 percent in Chicago in 2010, child poverty has been a significant factor contributing to low student test scores and graduation rates in these three cities (NAF 2009–2010).⁴¹

As noted earlier, Klein and Rhee take a “no excuses” tone, dismissing this factor, but other reform leaders have acknowledged the need to alleviate some poverty-related obstacles to learning. For example, at the American University forum on education reform in 2012, Duncan noted that early childhood education was a high priority and a smart investment, and Chicago Mayor Rahm Emanuel declared that “there is no greater investment than after-school programs, especially at high school, for their safety and [students’] own personal development – athletic, academic or artistic” (American University 2012). The other panelists, including Mayor Bloomberg, agreed. Yet these cities’ policy agendas largely sidelined evidence-based strategies shown to have potential to narrow income-based opportunity gaps.⁴² These strategies include the afterschool programs that Emanuel noted, which avert learning loss in out-of-school hours; early childhood education that improves school readiness; health clinics that reduce chronic absence; and full-service community schools, which bring together these and other resources to address many opportunity gaps associated with growing up poor. Yet, while Duncan did support community schools as CEO of CPS, his main emphasis was testing children who had entered kindergarten already behind their better-off peers, and using the results of those scores to close schools, among other policies.

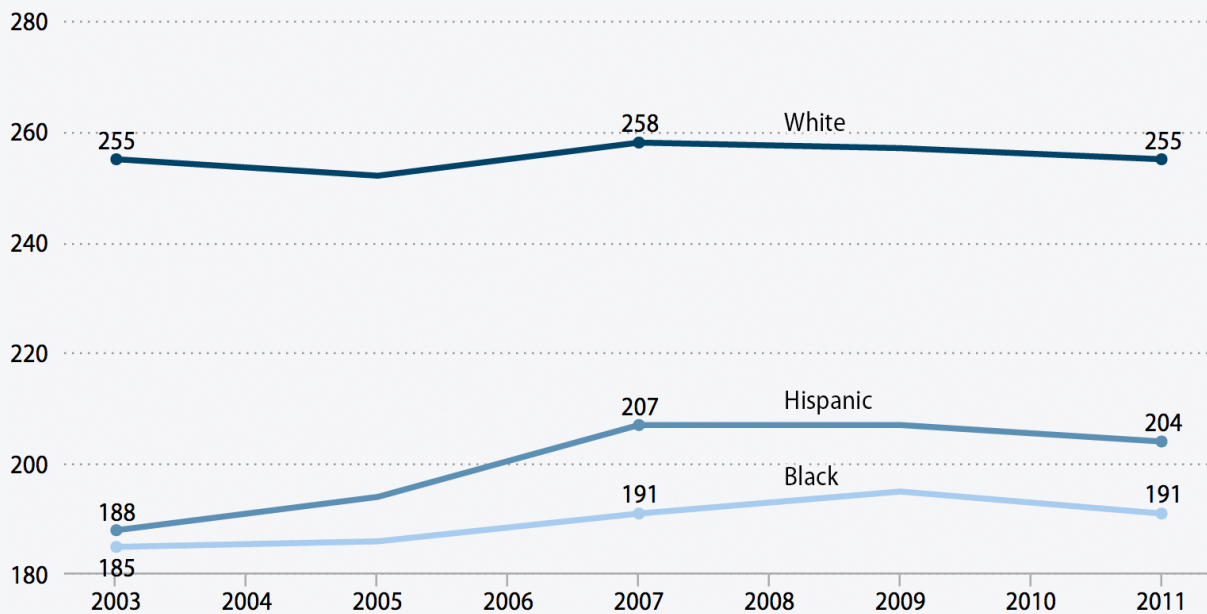
Though evidence pointed to improvements for low-income DCPS third graders who had attended quality pre-K programs, Rhee did not mention that factor, but rather credited IMPACT and other reform components with any gains. Indeed, among her early policy initiatives was one that impeded access to that program, contradicting the intent of the expansion passed just before her arrival to make pre-K available to all students. Bloomberg’s 2012 budget, which called for cutting afterschool and summer programs for more than 47,000 children, would have heralded the fifth straight year that he cut child care and afterschool programs. Added to prior cuts, the budget would have resulted in 90,000 fewer children having access to these programs than in 2009. A strong, organized advocacy campaign worked to avert these cuts by focusing media and public attention to the issue and working with key city council members to restore the cuts, at least temporarily (Ursula Helminski, personal communication 2012).⁴³ While fiscal realities forced budget choices, deep cuts to a relatively small line item suggest that afterschool was not truly seen as a high-priority investment.

The reforms have not improved student outcomes

The reforms have increased instability for students, teachers, and schools, with low-income communities bearing the brunt. As noted, this was somewhat intentional; the reformers believed that schools were so bad that only real disruption could change them enough to fix them, and this strategy inevitably led to churn. The fundamental question, then, is whether this disruption brought about real improvements for the students. If so, then the reforms fulfilled their promise, though perhaps with the need to improve stability. As this section documents, however, the disruption seems to

FIGURE F

District of Columbia Public Schools' 4th-grade reading scores, by race and ethnicity, NAEP, 2003–2011



Note: Graph shows scores for odd-numbered years; lines interpolate points in-between.

Source: Authors' analysis of National Assessment of Educational Progress data (Levy 2012c)

ECONOMIC POLICY INSTITUTE

have offered few gains for the low-income, minority, and other high-needs students that these policies sought to benefit. Indeed, in many respects, predicted outcomes went in the opposite direction: Test scores increased less than in other cities or in periods prior to the reforms, and already large achievement gaps grew.

Washington, D.C.: Race and income-based achievement gaps grew

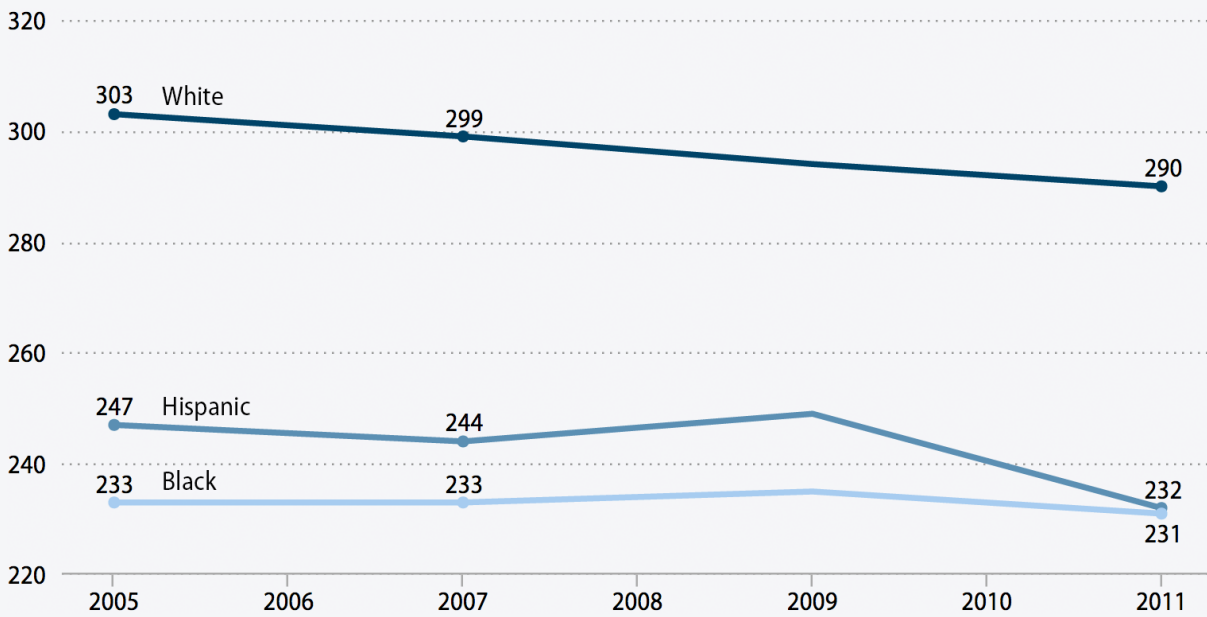
Using DC Comprehensive Assessment System (DC CAS) data, Chancellor Rhee and Mayor Fenty claimed major gains for all subgroups of students during their short control of DCPS from 2007 to 2010 (District of Columbia Public Schools 2010). Reliable data from the National Assessment of Educational Progress, combined with a more careful exploration by subgroup of the test score data they presented, shows that their claims were exaggerated, and that some students most in need of a boost lost ground.

As illustrated in **Figure F**, average fourth-grade NAEP reading scores of white, Hispanic, and black students rose in the four year period before Rhee came to Washington, D.C., but have since leveled off or reversed; since 2007, scores for black students have been stagnant, while those for white and Hispanic students have declined slightly. This combination of changes has slightly narrowed the race-based gap, but, again, only because of stagnant or declining scores, not gains (Levy 2012c).

Before and during Rhee's tenure, eighth-grade reading scores declined among all student subgroups, as seen in **Figure G**. White and Hispanic student scores dropped much more than black student scores, which shrank the black-white

FIGURE G

District of Columbia Public Schools’ 8th-grade reading scores, by race and ethnicity, NAEP, 2005–2011



Note: Graph shows scores for odd-numbered years; lines interpolate points in-between.

Source: DCPS National Assessment of Educational Progress data (Levy 2012c)

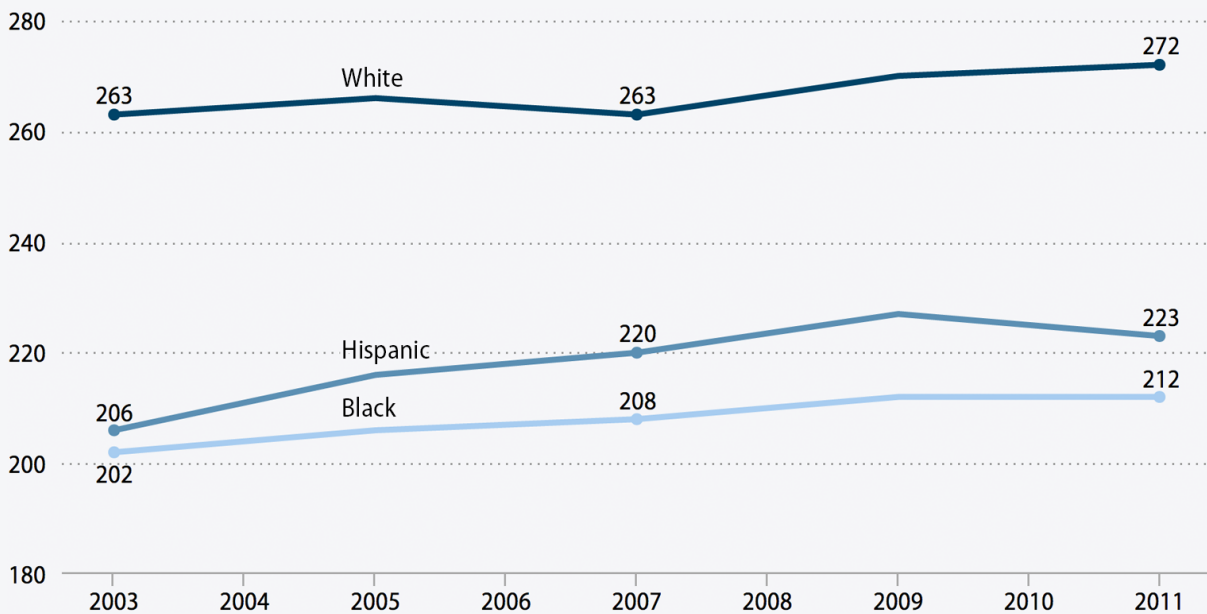
ECONOMIC POLICY INSTITUTE

gap but enlarged the white-Hispanic gap. This contrasts sharply with large, urban districts overall, and with the nation as a whole: At both of those levels, all three subgroups gained ground, with Hispanics gaining the most: six points from 2005 to 2011 both in large, urban districts (from 243 to 249) and nationally (245 to 251), closely followed by black eighth-graders, who gained five points in large, urban districts (240 to 245) and six points nationally (242 to 248), and whites gaining the least: three points in large, urban districts (270 to 273) and across the nation (269 to 272). As a result, gaps narrowed across the board (NCES 2005b and 2011b). (Throughout this report, DCPS NAEP scores are from data from D.C. budget consultant Mary Levy, who includes both regular public school students’ and charter school students’ scores in her calculations. We believe that these data are optimal, because they represent the vast majority of DCPS students. However, they are not perfectly comparable with the NAEP data for large, urban districts, which come from Trial Urban District Assessment reports, which do not always include charter school students’ scores. This does not, however, change any trends or ultimate findings. For example, TUDA data for DCPS confirm DCPS students’ relative lack of gains compared with their peers in other cities, as shown in Figure G; indeed, TUDA data would have shown a larger drop in scores for DCPS students. In the Appendix tables to this report, we present both TUDA and Levy data for DCPS, compared with large, urban district data.)

One bright spot before and through Rhee’s tenure was a small increase in “proficiency”—those demonstrating skills above the basic level—among black eighth-graders, from 8 percent in 2005 to 10 percent in 2011.⁴⁴ Unfortunately, this was accompanied by a much larger drop in proficiency for white students, from 77 percent to 63 percent, and a smaller decrease, from 18 percent to 14 percent, for Hispanic students (Levy 2012c).

FIGURE H

District of Columbia Public Schools' 4th-grade math scores, by race and ethnicity, NAEP, 2003–2011



Note: Graph shows scores for odd-numbered years; lines interpolate points in-between.

Source: DCPS National Assessment of Educational Progress data (Levy 2012c)

ECONOMIC POLICY INSTITUTE

Math scores often rise more easily than reading scores, which proved to be the case in D.C. Public Schools, as demonstrated in **Figure H**. Fourth-grade math scores improved among all student groups from 2003 to 2011; however, black students gains were less than half their white peers' gains, so the math gap among black and white fourth graders in 2011 was 60 points, up from 55 when Rhee took office in 2007, at which point the gap had been narrowing for several years (Levy 2012c).⁴⁵

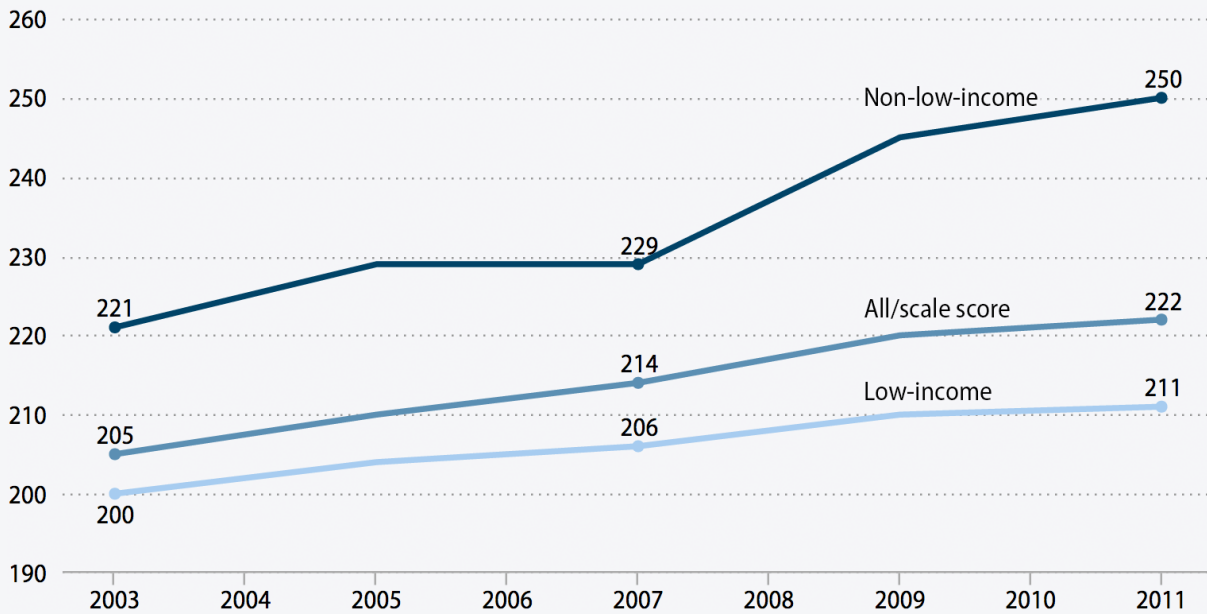
In eighth grade, the gap in 2011 was nearly as large, 59 points, down substantially from 70 points in 2003. However, the gap shrank only because black students lost substantially less ground (two points) than their white peers, whose scores dropped by 13 points.⁴⁶

As in reading, there was an increase in the share of students deemed “proficient” in math. Both white and black fourth-grade students gained real ground though, again, black students made only half the gains of their white counterparts. As a result, the gap widened, from 68 percent in 2007 to 73 percent in 2011. Among eighth-graders, both groups made real gains between 2005 and 2011, though the gap stayed nearly the same (68 percent in 2005 versus 69 percent in 2011), and the percentage of black students who were considered proficient grew only from a miniscule 3 percent to 9 percent, still fewer than one in ten (Levy 2012c).

NAEP data enable researchers to track scores for students who are eligible for free or reduced-price lunch, and thus low-income, and compare those scores with students who are ineligible (not low-income). As shown in **Figure I**, the

FIGURE I

District of Columbia Public Schools' 4th-grade math scores, all and by income group, NAEP, 2003–2011



Note: Low-income students are those who are eligible for free or reduced-price lunch. Graph shows scores for odd-numbered years; lines interpolate points in-between.

Source: DCPS National Assessment of Educational Progress data (Levy 2012c)

ECONOMIC POLICY INSTITUTE

gap in scores between these two groups of students—the income-based achievement gap—increased even more than the gap across racial groups. The gap in NAEP math scores between low-income and non-low-income fourth-graders grew from 21 points (200 for low-income versus 221 non-low-income) in 2003 to 39 points (211 low-income versus 250 non-low-income) in 2011. The gap was virtually unchanged in the four years before Rhee, when both groups saw modest increases, but grew dramatically from 2007 to 2011, when low-income students gained five points, versus their wealthier peers’ dramatic 21 points (Levy 2012c).

The gap between poor and non-poor eighth graders also increased substantially, although this was due to real gains for both groups, with non-low-income students reaping double the increases of their low-income counterparts (Levy 2012c).

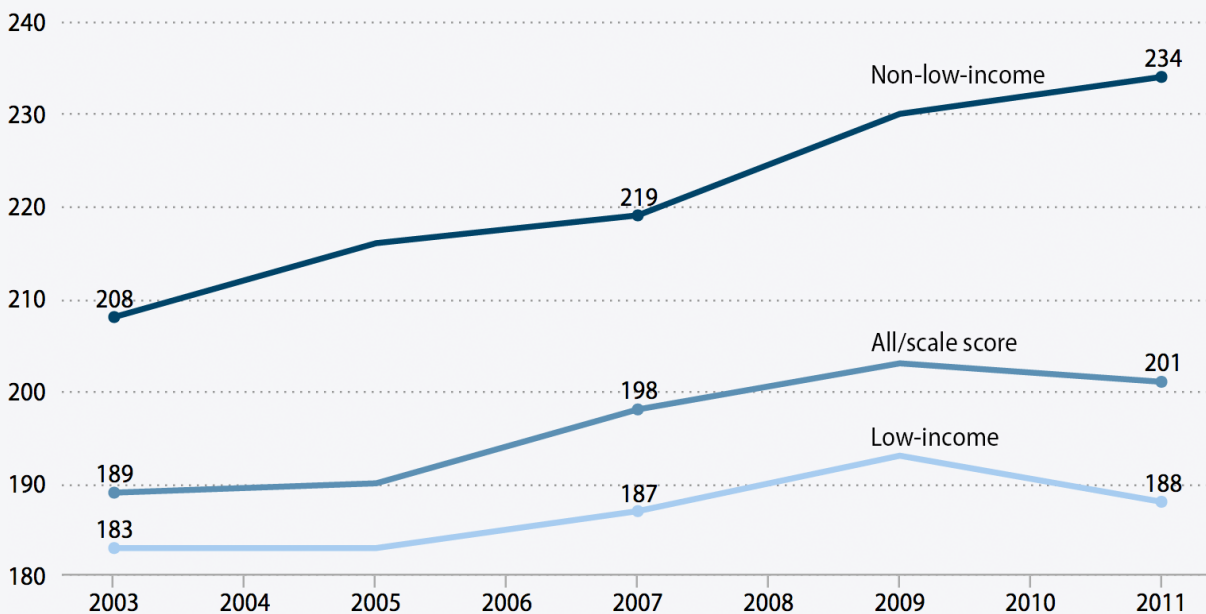
Reading scores were, again, less promising: While non-low-income fourth graders posted real gains between 2007 and 2011, low-income students gained just one point, resulting in a gap, shown in **Figure J**, that grew from 32 to 46 points in just four years.

Low-income eighth-graders also did not improve their reading, while their higher-income peers posted small gains, resulting in a gap that grew more modestly, from 24 to 31 (Levy 2012c).

The average gains by DCPS students in both reading and math from 2007 to 2011, during a period of little or no progress among low-income students, could be explained by disproportionate improvements in high-income students’

FIGURE J

District of Columbia Public Schools' 4th-grade reading scores, all and by income group, NAEP, 2003–2011



Note: Low-income students are those who are eligible for free or reduced-price lunch. Graph shows scores for odd-numbered years; lines interpolate points in-between.

Source: DCPS National Assessment of Educational Progress data (Levy 2012c)

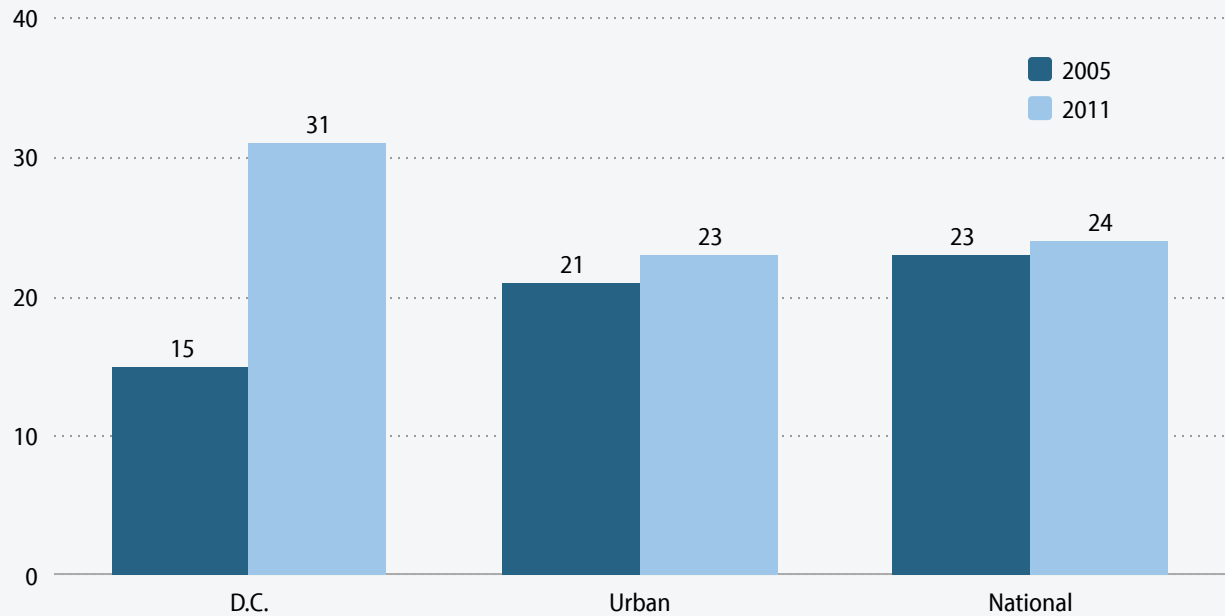
ECONOMIC POLICY INSTITUTE

schools. Another explanation is that, rather than gains among a (minority) group of higher-income students, DCPS had an influx of new higher-scoring, higher-income students, who account for much of these gains. If so, then reforms enacted under Rhee might not even account for the gains seen by those higher-income students. The data clearly affirm this second explanation. In fourth grade, for example, overall math scores increased from 214 in 2007 to 222 in 2011, or 3.7 percent. Among black students, who constitute the vast majority of the student body, however, the increase was just 1.9 percent (from 208 to 212), and similarly small among Hispanics. Gains among white students, who constitute a small minority of the student body, were no bigger (from 263 to 272, or 3.4 percent), so the only explanation is an influx of higher-scoring (white and high-income) students. Fourth-grade reading scores make this even clearer: average (scale) scores rose from 198 to 201, or by 1.5 percent. Among both white and Hispanic students, scores fell by 3 points over the period, and they remained at 191 among black students. If two groups lost ground, and the majority was stagnant, only new students could account for the small overall gain.

While NAEP is the only truly reliable source of year-to-year comparable information, other tests that are not independently reliable but that illustrate differences along geographical lines reveal a similar pattern. Students in schools in wards 7 and 8 (the lowest-income wards in the city) trailed their Ward 3 peers (who live in the wealthiest area) in reading and math pass rates by margins of 41 to 56 percentage points on the 2011 DC CAS exams. As *Washington Post* education reporter Bill Turque wrote, this gap hadn't budged:

FIGURE K

Income-based 8th-grade reading test score gaps, 2005 and 2011, Washington, D.C., vs. other large urban and all districts



Note: The income-based gap is the difference between test scores of low-income students (those eligible for free or reduced-price lunch) and non-low-income students.

Source: National Center for Education Statistics (2005b and 2011b)

ECONOMIC POLICY INSTITUTE

[In] Ward 8 elementary schools ... 28 percent of students read at proficiency level or better, down about 2 percentage points from 2010. That level is almost identical to the pass rate in 2007, when then-Mayor Adrian M. Fenty (D) took control of the city school system. (Turque 2011b)

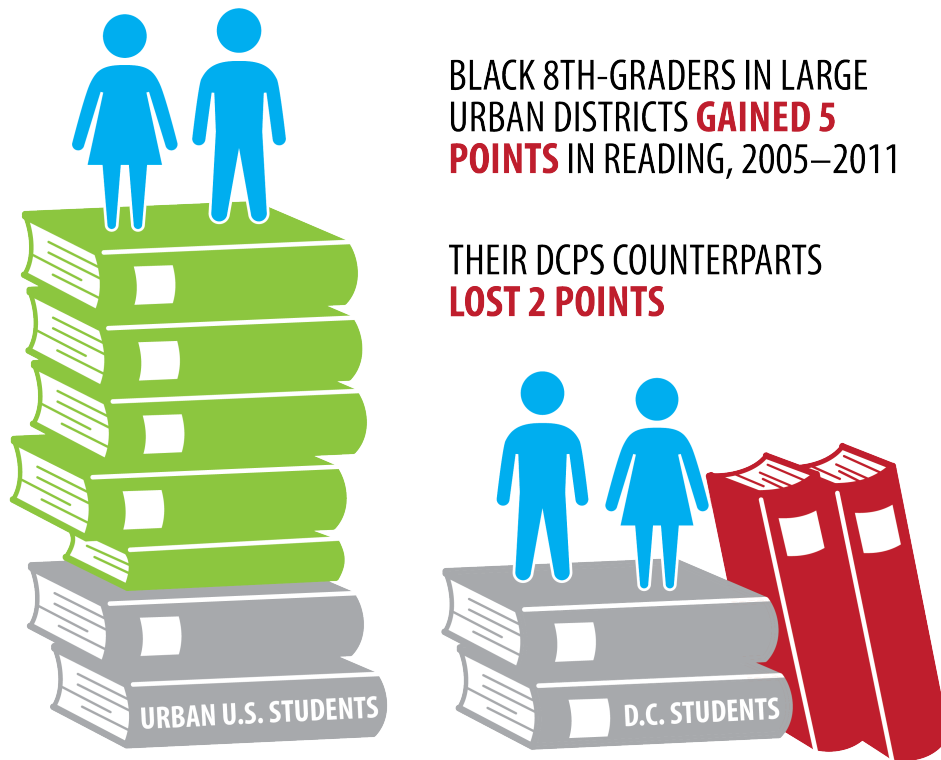
Ward 3 elementary schools had an 84 percent pass rate, for a gap of 56 points.

While the income-based gap in eighth-grade NAEP reading scores increased by just one point at the national level and two points in other large, urban school districts from 2005 to 2011, it more than doubled in Washington, D.C., as illustrated in **Figure K**.

After two years of declines between 2009 and 2011, citywide scores on the DC Comprehensive Assessment System rose very slightly in 2012 for mathematics, but were flat for reading (Chandler and Brown 2012). And while charter schools had slightly higher scores than public schools, their gains were smaller, and their reading proficiency rates fell slightly. In 2012, these small gains left 46 percent of regular public school students and 55 percent of charter school students proficient in math, and just 44 percent and 49 percent, respectively, proficient in reading (Mahaley 2012). The reported scores were not broken down by race or income, which makes it impossible to assess impacts on subgroups of students. Moreover, as discussed earlier, this report treats non-NAEP test scores as a less valid source of information, given incentives to teach to tests and otherwise boost scores due to the high stakes attached to them. One interesting

REFORMS LED TO LOST GROUND, BIGGER GAPS FOR THE STUDENTS THEY WERE SUPPOSED TO SUPPORT

In most large urban districts studied, test score gains among minority students narrowed race-based achievement gaps, and low-income students had gains comparable to their affluent peers. This contrasts with reform cities, where achievement gaps grew as poor and minority students' scores fell further behind those of their peers.



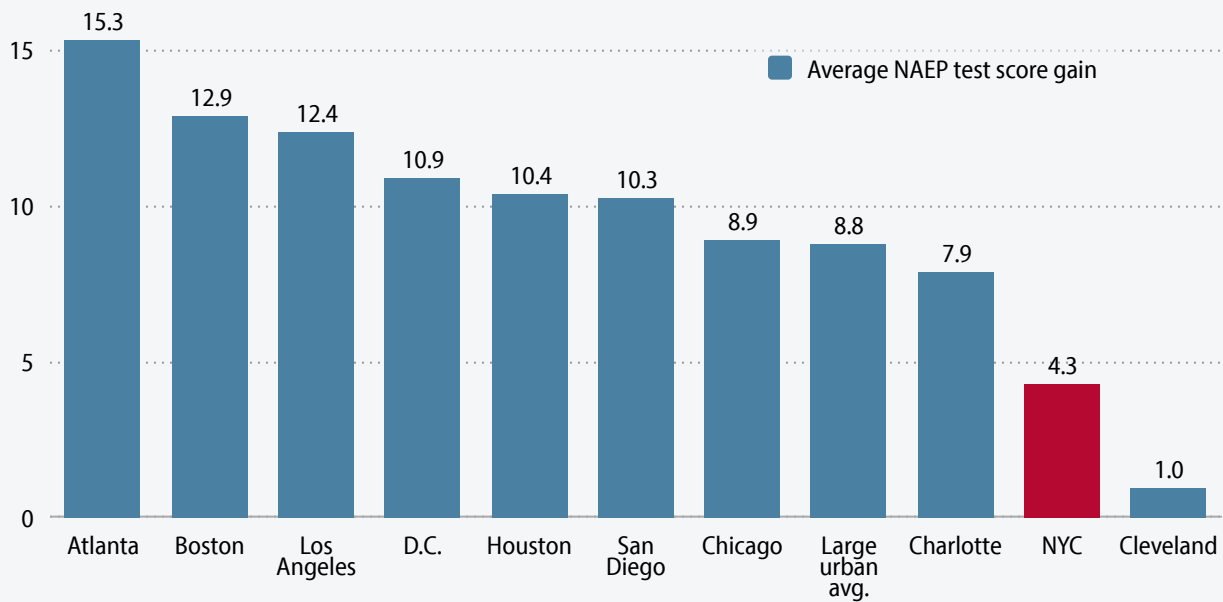
Source: National Center for Education Statistics, *Trial Urban District Reading Assessment*, 2005 and 2011; National Assessment of Educational Progress scores for District of Columbia Public Schools provided by D.C. budget consultant Mary Levy in 2012.

finding that was not subject to manipulation, however, was that, across ethnic and income groups, third-graders who had participated in prekindergarten programs outscored their non-pre-K peers (Mahaley 2012).⁴⁷

High school graduation rates in DCPS were decreasing before Rhee arrived. Between the 2001–2002 and 2007–2008 school years, the NCES-published “averaged freshman graduation rate”—the estimated percentage of ninth-graders who graduated four years later—for DCPS dropped 12.4 percentage points, from 68.4 percent to 56 percent.⁴⁸ This was much lower than both the national average of 75 percent and large, urban district average of 65 percent in 2007–2008 (the latter of which was down four percentage points from 2001–2002). In other words, graduation rates dropped much more sharply in the years preceding Rhee’s arrival in Washington, D.C., than in other large, urban districts. DCPS’s graduation rate rose the first year of Rhee’s tenure (2008–2009) to 62.4 percent, but as more recent comparable data are unavailable, neither the DCPS trend during her tenure nor comparisons to other districts can be assessed.⁴⁹ While these data have likely fluctuated due to changes in how they are calculated, and, as discussed below, graduation rates are

FIGURE L

Average NAEP test score gains from 2003–2011 in 10 large, urban school districts



Note: The bars reflect test score gains from 2003 to 2011 averaged across reading and math for both fourth and eighth grades. The 10 large, urban districts shown are those participating in the NAEP Trial Urban District Assessment (TUDA) in 2003. TUDA uses NAEP scores from a representative sample of students in large, urban districts with high proportions of low-income and minority students. TUDA began in 2002 with six districts and had 21 as of 2011.

Source: Haimson 2012

ECONOMIC POLICY INSTITUTE

notoriously hard to use for valid city-to-city or year-to-year comparisons, the city’s graduation rate seems, at best, to be returning to the low level of a decade ago, with no overall gains over the past 10 years.

New York City: Nine years of market-based reforms failed to improve test scores or narrow achievement gaps

In March, 2012, Mayor Bloomberg repeated his claim that the achievement gap between white/Asian students and black/Latino students in New York City public schools had been halved between 2003 and 2011 (American University 2012). Actually, averaged across state reading and math test scores in the fourth and eighth grades, the achievement gap had stagnated; it was virtually identical in 2011 (25.8 percentage points or 0.73 of standard deviation) to its 2003 level (26.2 percentage points or 0.74 of standard deviation), and not statistically significantly different.⁵⁰ Indeed, Columbia University professor Aaron Pallas, who calculated the actual size of the reduction in the gap (of 1 percent), pointed out, “The careful reader will note that the mayor has thus overstated the cut in the achievement gap by a factor of 50” (Pallas 2012b).⁵¹ Another study that compared NAEP test score gains from 2003 to 2011 averaged across reading and math in fourth and eighth grades found New York City to be second to last, ahead only of Cleveland, among the 10 TUDA districts, as illustrated in **Figure L**, with a gain of only 4.3 points (Haimson 2012). The average large district gained 8.8 points over this period, with Atlanta gaining the most—15.3 points.⁵²

REFORMERS CLAIMED MASSIVE TEST-SCORE GAINS THAT DATA PROVED FALSE

Reformers in DC, NYC, and Chicago reported “success” in large test score gains and shrinking achievement gaps. When the data were recalibrated, broken down by subgroup, and compared with reliable numbers, however, the gains vanished and gaps grew.



NYC MAYOR BLOOMBERG CLAIMED TO HAVE CUT THE RACE-BASED
ACHIEVEMENT GAP **BY 50%** FROM 2003 TO 2011



IN REALITY, THE GAP CLOSED **BY 1%**

Source: Aaron Pallas, “The Emperor’s New Close,” *GothamSchools*, March 7, 2012.

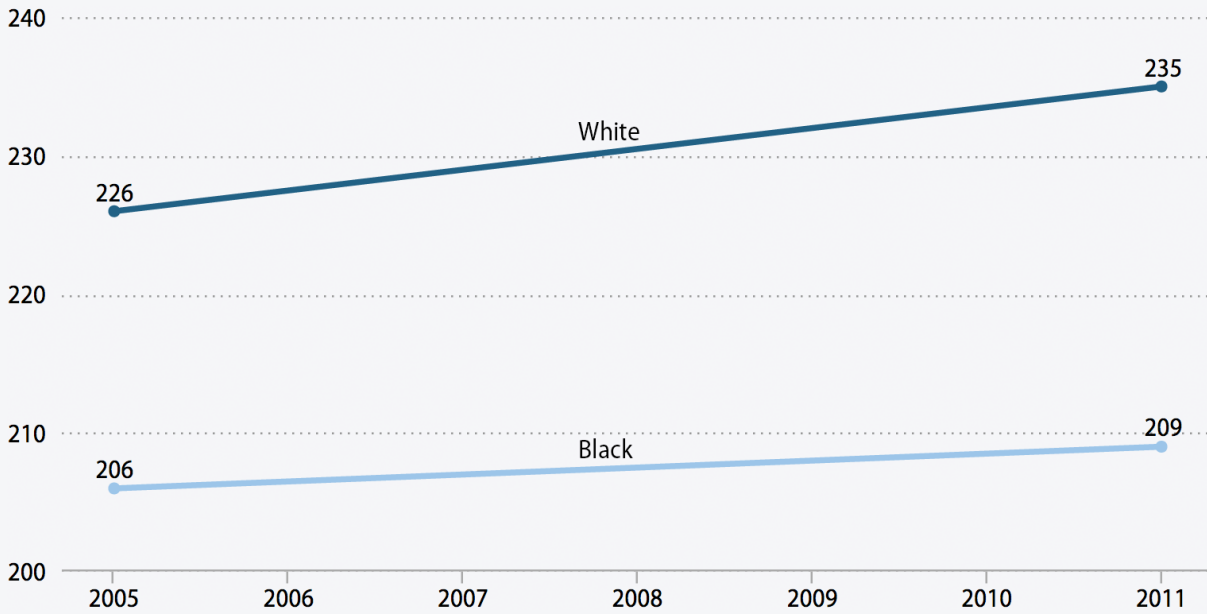
The gap failed to shrink partly because white fourth-graders gained three times as much ground as their black peers between 2005 and 2011, as shown in **Figure M**. Indeed, this gap grew by 6 percentage points during those years, versus a slight decrease in the same gap in large, urban districts overall. The income-based gap for that age grew by even more, also bucking both national and urban trends during the period.

Bloomberg made similarly exaggerated claims regarding the city’s public school students’ “proficiency” on state test scores. When the New York State Department of Education recalibrated the scores, however, the gains vanished, and the proportion of students passing the state reading test fell from 68.8 percent to 42.4 percent, and from “an astonishing 81.8 percent to a disappointing 54 percent in mathematics” (Ravitch 2010). Again, NAEP scores confirm unimpressive gains. Between 2003 and 2011, average NAEP math scores in New York City public schools rose 8 points in fourth grade and 6 points in eighth grade between 2003 and 2011, gains similar to those in the nation as a whole, and half the size, at the eighth-grade level, of gains in other large, urban districts that did not engage in similar reforms (NCES 2011a).⁵³

As **Figure N** illustrates, New York’s eighth-grade students improved at rates similar to those of their urban counterparts in reading; between 2005 and 2011, the city’s white students gained just two points (versus three in large, urban districts on average), while the city’s black students gained seven points (versus five in large, urban districts on average). These gains slightly narrowed the black-white gap in the city, but not enough to counter growing gaps in other subjects and grades.

FIGURE M

New York City 4th-grade reading scores, by race, NAEP, 2005 and 2011



Note: Graph shows scores in 2005 and 2011; lines interpolate in-between years.

Source: National Center for Education Statistics (2005b and 2011b).

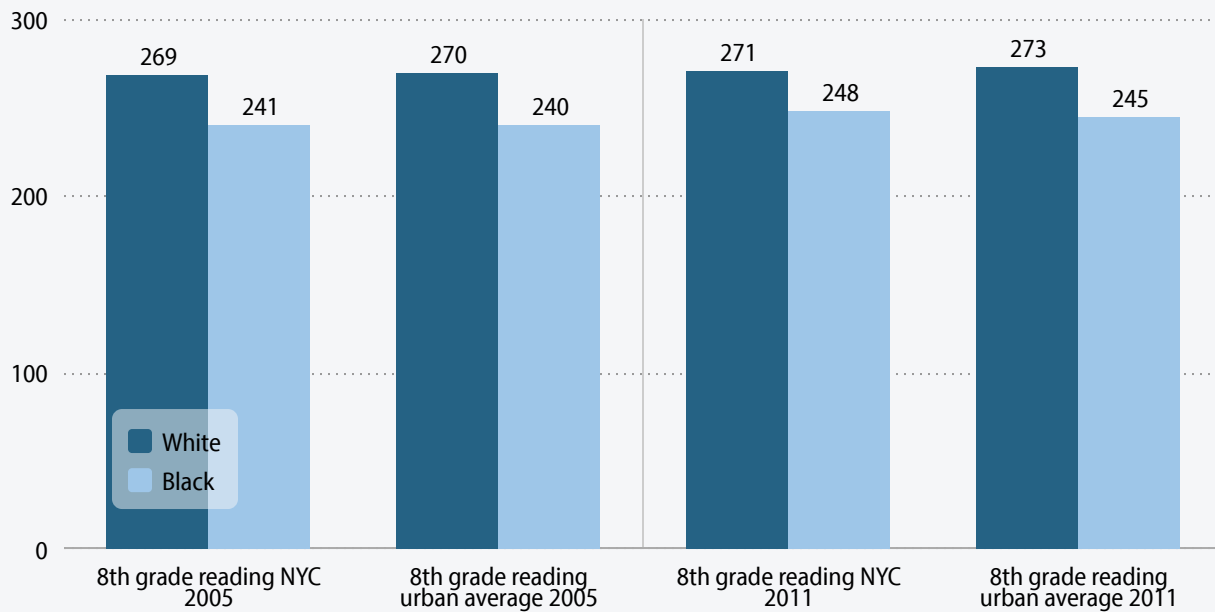
ECONOMIC POLICY INSTITUTE

One definite bright spot in the New York City data come from graduation rates, which increased sharply under Bloomberg and Klein’s leadership, according to data tracked by the city. While just under half (46.5 percent) of the cohort of 2001 (class of 2005) graduated in four years, that share rose to just over 60 percent for the cohort of 2007 (class of 2011) (NYC DOE 2012). The most recent nationally collected data for New York City found similar gains, though the different method of calculating national data produced a lower overall percentage: 56.9 percent of the class of 2008 graduated, an increase of 11 percentage points from 46 percent for the class of 2005 (NCES 2005a and 2010a).

It is unclear, however, what impact, if any, the Bloomberg/Klein reforms had on graduation rates. The National Center for Education Statistics tracks graduation rates for the 100 largest urban districts (some of which, such as Montgomery County, Md., are not cities). The 11 percentage-point average increase in New York City’s high school graduation rate from the classes of 2005 to 2008 is consistent with the 12 percentage-point increase in graduation rates for the 100 largest urban districts in the same period, and it still left New York City well below the 65 percent average class of 2008 graduation rate for those districts (NCES 2010b). It is also important to be aware of critical caveats pertinent generally to claims regarding high school graduation rates. Due to differences in definitions, graduation requirements, and data collection and calculation methods, such data are problematic, especially when used to make comparisons over time or across states and districts. For example, New York includes GED completers in its definition of a high school graduate, whereas the U.S. Department of Education does not, which is likely one factor in the difference between city and federal recorded rates. Researchers have also found that administrative data are often “jumpy” across the years for reasons that they do not fully understand (Roy and Mishel 2008). Other issues include disproportionate “discharges” by race

FIGURE N

New York City 8th-grade reading scores, by race, compared with large, urban district average, 2005 and 2011



Note: Large, urban districts are those participating in the NAEP Trial Urban District Assessment (TUDA), which uses NAEP scores from a representative sample of students in large, urban districts with high proportions of low-income and minority students. TUDA began in 2002 with six districts and had 21 as of 2011.

Source: National Center for Education Statistics (2005b and 2011b)

ECONOMIC POLICY INSTITUTE

that may inflate graduation rates by “pushing at risk students out of school” and not counting them as dropouts (Jennings and Haimson 2009).

There is legitimate disagreement regarding the degree to which charter schools improve their students’ achievement, relative to comparable public schools. Two rigorous studies found substantially better outcomes for students who gained admittance via lottery to city charter schools compared with outcomes for their peers who were not admitted. One, a Stanford University Center for Research on Education Outcomes (CREDO) study of New York City charter schools that employed “composite virtual” students from “competitor traditional public [feeder] schools” found benefits of charter school attendance across almost all student groups, including low-income and minority students, both African American and Hispanic (CREDO 2010, 2). However, two caveats apply to such findings. First, as described earlier, NYC charter schools generally receive more funding than comparable public schools while serving fewer hard-to-teach students. Second, as emphasized in this paper, state standardized test scores, the unit of measurement used in the CREDO study, are easily manipulated. If charter schools focus more on test preparation than their peer neighborhood public schools, this could account for some, even all, of the difference in test scores. Because NAEP scores for New York City are not available by charter versus public school, as they are in Washington, D.C., and since the type of analysis conducted by the Consortium on Chicago School Research (Lupescu et al. 2011) to alleviate some of the potential

biases of Illinois state tests has not been done in New York City, these results must be interpreted through that skeptical lens.

The CREDO researchers themselves cautioned that it is also impossible to infer much from the finding that New York City charter schools have no impact on students receiving special education services. Because the highest-needs students are least likely to enroll in charter schools, and thus make up a relatively small share of the student body, any findings are likely skewed in favor of charter school benefits; i.e., findings of equal impact may actually be masking poorer outcomes for higher-needs students. The CREDO study found a similar lack of impact for English language learners. Finally, students retained in grade did not benefit in reading from being in charter schools, and had poorer outcomes than their public school counterparts in math.

The second of the studies to find benefits for charter school students, an earlier 2009 study funded by the National Center on School Choice at Vanderbilt University, estimated that, “on average, a student who attended a charter school for all of grades kindergarten to eight would close about 86 percent of the achievement gap in math and 66 percent of the ‘Scarsdale-Harlem’ achievement gap in English” (Hoxby, Murarka, and Kang 2009, IV-I). It also reported that students who attended charter high schools were roughly 7 percent more likely to earn a Regents diploma for each year in that school, and that their Regents scores would be about 3 points higher for each year spent in the school (Hoxby, Murarka, and Kang 2009). The impacts on students varied quite a bit across schools, enabling the researchers to assess potential mechanisms driving student gains. Among the policies associated with charter school success were a longer school year (two and a half weeks longer than regular public schools); more time devoted to English during the school day (which is also longer than that of traditional public schools, by roughly 90 minutes); a small rewards/small penalties disciplinary policy; teacher pay based in part on performance or duties; and a mission statement that emphasized academic achievement versus other goals. Charter schools also differed from public schools in offering Saturday classes that may be optional or mandatory, and the vast majority of charter school students attended schools that offer after-school programs. Roughly half of charter schools employed parent contracts that set expectations regarding attendance, on-time school arrival, homework, and other issues.

In his review of the study, however, economist Sean Reardon issued several caveats regarding these findings. These include the misuse of test score data from a prior year that may render moot the purported randomization provided by the lottery; inappropriate extrapolation of impacts that may overstate cumulative benefits; a weak standard for statistical significance that suggests true differences where none may exist; and the omission of ineffective charter schools from the distribution such that the range of effectiveness may be skewed. In general, “the results presented appear to overstate the cumulative effect of attending a charter school” (Reardon 2009, 1).

The 2011 Rutgers report on funding disparities also called some of the reported charter school benefits into question. Disputing the CREDO findings, the researchers asserted that,

controlling for a school’s [English Learner] enrollment, its free-lunch enrollment, its location and its enrollment stability, NYC’s charters did not do better, and arguably did worse, in terms of their student’s (sic) test score outcomes. Contrary to press accounts of NYC charter school success at raising student achievement, our analysis found no statistically significant differences in charter versus non-charter school performance (level, not gain)

for grades 4, 6 and 7, and charter school test performance lower than that of non-charter schools in grade 5.” (Baker and Ferris 2011, 17)

However, their methodology appears to be less rigorous than that employed by the CREDO researchers, with comparisons of scores versus growth a potential weakness. In sum, there is fairly solid evidence that NYC charter schools confer students an advantage in growth and/or proficiency, but this advantage is not as high as has been reported in previous studies. Moreover, because they spend more money per student, while educating fewer hard-to-teach and expensive students, relative to regular NYC public schools, this overall advantage in effectiveness is likely small.

Finally, assuming findings of slightly higher test scores and/or graduation rates among charter high school students are valid, this doesn't necessarily translate into the goal of college readiness emphasized by reformers: Only 10.2 percent of 2010–2011 charter school graduates were deemed “college ready” based on NYC Department of Education progress reports, versus over twice as many, 26.9 percent, among regular district school graduates (Merriman et al. 2012, 21). District students in non-charter schools were also slightly more likely to enroll in college (51.3 percent) compared with their charter school peers (48.8 percent).⁵⁴

As evidence of the importance it placed on college readiness, the city instituted school report cards, grading each school from A to F, with A and B having “the highest percentage of students who took college-level courses, did not need remediation classes upon graduation and went to college within four years of entering high school,” according to a *New York Times* analysis. The gap between that goal and the city's capacity to reach it via the reforms, or even to faithfully report it, however, is evident; even among A schools (a few of which graduated the vast majority of their students), only a small minority of students were “college-ready” (Santos 2011). Also, according to NYC DOE reports of schools' “college and career readiness index,” “college readiness was much higher in what the DOE calls ‘traditional’ schools, those that predate Mayor Bloomberg. In those schools, 30 percent of students graduated college ready, compared with 20 percent in the Bloomberg-era schools” (McAdoo 2012).⁵⁵

Chicago: As achievement gaps grew, African American students fell further behind, and college readiness remained disappointing

Although students made some progress in both reading and math during Arne Duncan's 2001–2009 tenure as CEO of Chicago Public Schools, the race-based achievement gap increased, with white students making slightly more progress than Latino students, and black students falling behind all other groups (Lupescu et al. 2011). The Chicago Consortium on School Research (CCSR), a respected institution that has studied reforms and outcomes in CPS for several decades, employed a complex statistical process to translate the Illinois Standard Achievement Test (ISAT) data into more consistent and reliable measures of changes in student achievement over time.⁵⁶ CCSR found that, in elementary and middle schools, in particular,

test scores for African American students improved at a much slower rate than those of other students. Average scores for African American students improved slightly in math, while improving moderately among other students. There were virtually no improvements in reading scores among African American students, while white and Asian students showed modest improvements and Latino students showed slight improvements. (Lupescu et al. 2011, 7)

CCSR's methods enabled researchers to gauge the impact of test-taking and stakes attached to it on instructional practices and learning. They noted, with respect to a decline in test scores at the point at which the district switched tests, that

[t]he decline in scores in 2006 seems to be an artifact of the ways in which students were prepared for the test. ... [this decline] was driven by schools with many low-achieving students—schools that were at risk of accountability sanctions based on students' performance on the tests. These schools had strong incentives to gear instruction specifically towards the content of the high-stakes test, and the types of questions asked on the tests. When the district switched to a different test, students' performance on the tests dropped. (Luppescu et al. 2011, 24)

This confirms not only that test results tied to high stakes are unreliable measures, but that students in struggling schools are likely losing out on real instruction as test preparation increasingly becomes the norm.

The researchers also noted that apparent test score gains were due to non-equivalent scoring on the ISAT rather than real change, and that CPS student improvement was comparable to that of students across the state, a reality confirmed by NAEP scores. Chicago trailed its large, urban district counterparts by about 10 points in fourth-grade math and eight points in eighth-grade math in 2003, gaps that narrowed very slightly to nine and seven points, respectively, in 2009. In reading, however, the gaps grew a bit larger over that period, from six points to eight points in fourth grade and from one point to three points in eighth grade (Luppescu et al. 2011, 30). **Figures O** and **P** show that improvements in Chicago Public Schools' fourth- and eighth-grade reading scores from 2003 to 2009 were not keeping pace with gains in large, urban districts over that period, as measured by NAEP scores, and that this trend increased the gap in performance between CPS and its large, urban school district counterparts.

A 2011 report by the Collaborative for Equity and Justice noted this consistent gap and lack of academic improvement: "The only place Chicago significantly leads the Large Cities is in the percent of students Below-Basic in math (grade 4 and 8) and reading grade 4" (Lipman and Gutstein 2011, 14).

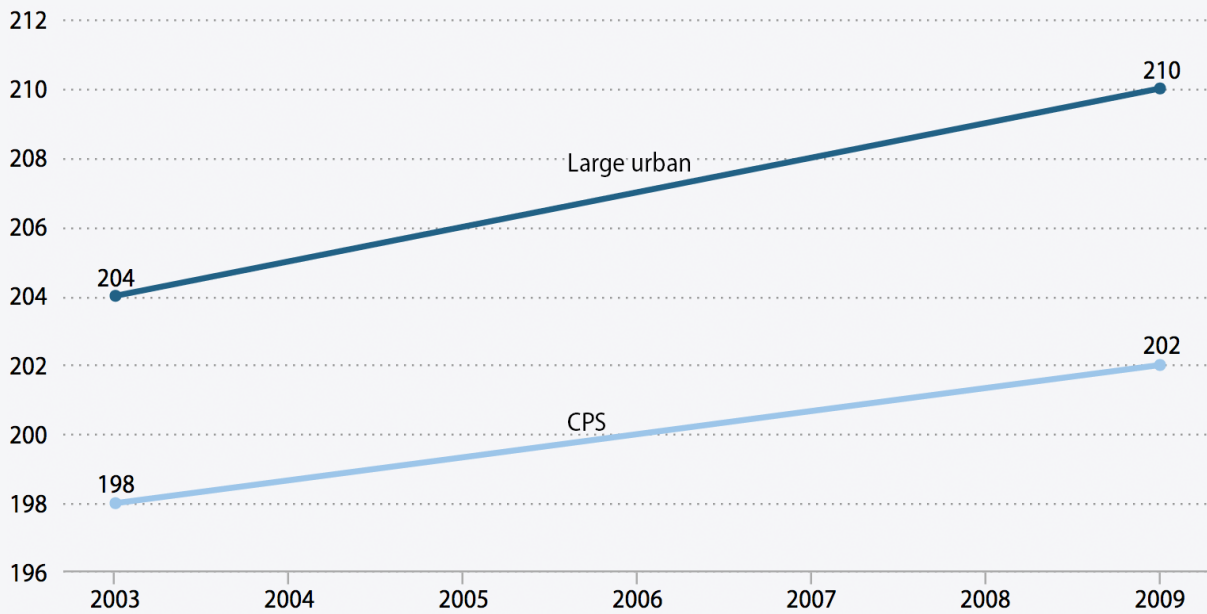
Indeed, a 2009 report by the Civic Committee of The Commercial Club of Chicago noted that the city's schools had made little progress since 2003, refuting President Obama's 2008 assertion that elementary school test scores during Duncan's seven-year tenure had risen 29 percentage points, from 38 percent of students meeting the state standards in 2001 to 67 percent in 2008, and similar claims by Duncan.

As recently as January 2009, CPS distributed brochures showing that 8th grade reading scores improved from 55% of students meeting/exceeding standards in 2004 – to 76% in 2008. And 8th grade math scores improved from 33% in 2004 to 70% in 2008. But these huge increases reflect changes in the [state] tests and testing procedures – not real student improvement. The reality is that most of Chicago's students are still left far behind. Real student performance appears to have gone up a little in Chicago elementary schools during the past few years – and even those gains then dissipate in high school. (Civic Committee of the Commercial Club of Chicago 2009, 2)

Adjusting for changes in tests and procedures, the Commercial Club study found that pass rates—the percent of elementary and middle school students at or above grade level—grew by only about 8 points between 2001 and 2008

FIGURE O

Chicago Public Schools' 4th-grade reading scores compared with large, urban districts, NAEP, 2003 and 2009



Note: Large, urban districts are those participating in the NAEP Trial Urban District Assessment (TUDA), which uses NAEP scores from a representative sample of students in large, urban districts with high proportions of low-income and minority students. Graph shows scores in 2003 and 2011; lines interpolate in-between years.

Source: Luppescu et al. (2011), Figure 13

ECONOMIC POLICY INSTITUTE

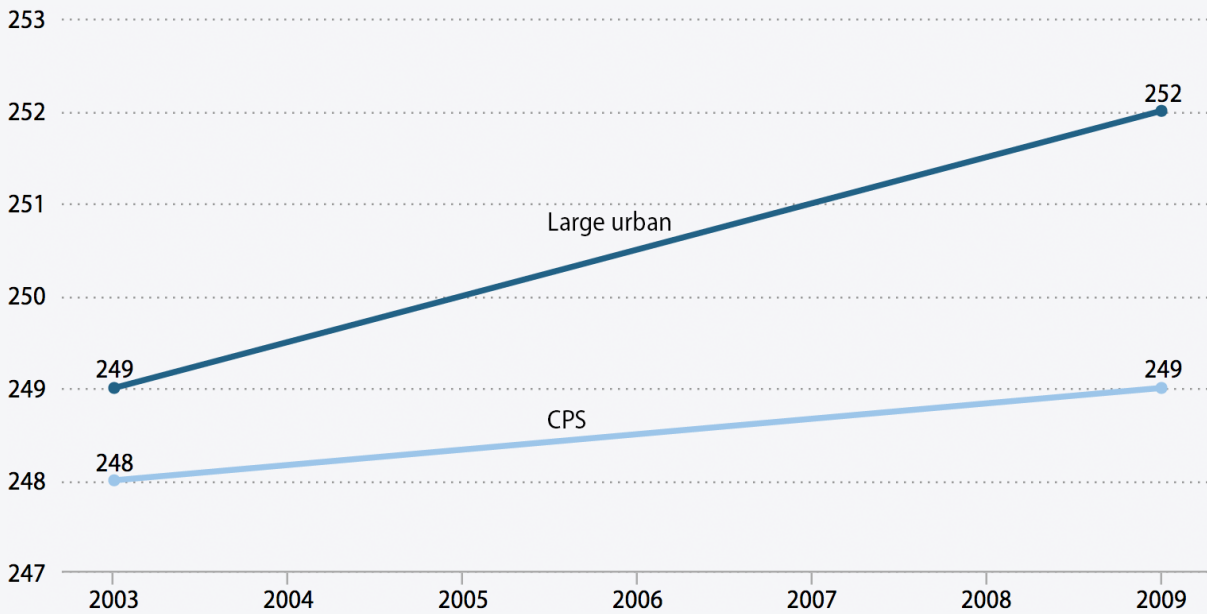
(Civic Committee of the Commercial Club of Chicago, 10). In high schools, the increase was almost imperceptible—just a point and a half.

And to the extent that low-income and minority students made any gains, their concentration in lower-level skills raises concern that those students remain unprepared to compete with better-off peers for schools, college, and jobs. According to the Collaborative for Equity and Justice report, “most of the gains for these students were in lower-level computation skills, rather than conceptual understanding, higher-order thinking, and problem-solving skills that would have shown up in gains at the ‘Advanced’ level. That is, to the extent there are gains, low-income students of color are likely acquiring very basic skills” (Lipman and Gutstein 2011, 17). The Consortium on Chicago School Research noted, too, Chicago’s break with more positive national data.

The increase in the gap in reading and math scores between white and African American elementary grade students in Chicago was quite different from national trends. On the national NAEP exam, fourth grade racial gaps closed substantially over the course of [the three eras studied for this report] in both reading and math, while eighth grade gaps were not consistently up or down. (Luppescu et al. 2011, 33)

FIGURE P

Chicago Public Schools' 8th-grade reading scores compared with large, urban districts, NAEP, 2003 and 2009



Note: Large, urban districts are those participating in the NAEP Trial Urban District Assessment (TUDA), which uses NAEP scores from a representative sample of students in large, urban districts with high proportions of low-income and minority students. Graph shows scores in 2003 and 2009; lines interpolate in-between years.

Source: Luppescu et al. (2011, Figure 13)

ECONOMIC POLICY INSTITUTE

Finally, different types of schools fared differently, with schools that served largely African American students and those that started out with the weakest reading and math scores gaining the least in both subjects during Duncan’s tenure. This contrasted with the prior eras of (and different types of) reforms in Chicago, in which lower-performing schools made more progress relative to their non-low-income counterparts (Luppescu et al. 2011, 36–39).⁵⁷

As **Figure Q** indicates, to the extent the Duncan-era reforms produced any gains, they accrued to Asian or white students; Latino students essentially stagnated and black students lost ground.

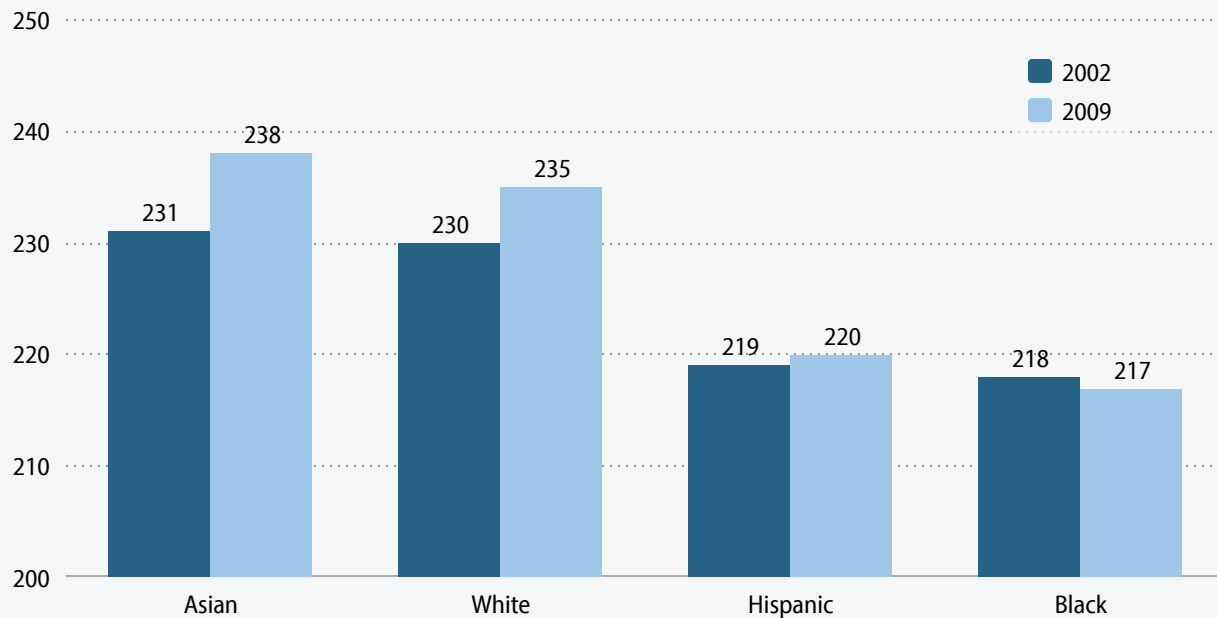
Despite enrolling easier-to-serve students, Chicago’s charter schools seem to perform no better than comparable city public schools. While multiple studies of the charters found somewhat different outcomes, with some finding small gains for specific subgroups and others finding small losses, none showed dramatic improvements in student performance.

Chicago charter school outcomes are mixed, overall showing roughly comparable performance to neighborhood schools. On average, Chicago charter high schools served fewer English language learners and low-income

“The increase in the gap in reading and math scores between white and African American elementary grade students in Chicago was quite different from national trends.”

FIGURE Q

Chicago Public Schools' elementary/middle-school reading scores (ISAT), by race and ethnicity, 2002 and 2009



Note: This figure represents changes to Illinois Standard Achievement Test (ISAT) data by the Consortium for Chicago School Research (CCSR), which employed complex statistical analyses to adjust the raw ISAT data for changes in text content, scoring, and other inconsistencies, in order to enable viable year-to-year comparisons.

Source: Luppescio et al. (2011), Figure 17

ECONOMIC POLICY INSTITUTE

and special education students, and on average, Chicago charter schools replaced more than half of their staff between 2008 and 2010. (Lipman and Gutstein 2011, 3)

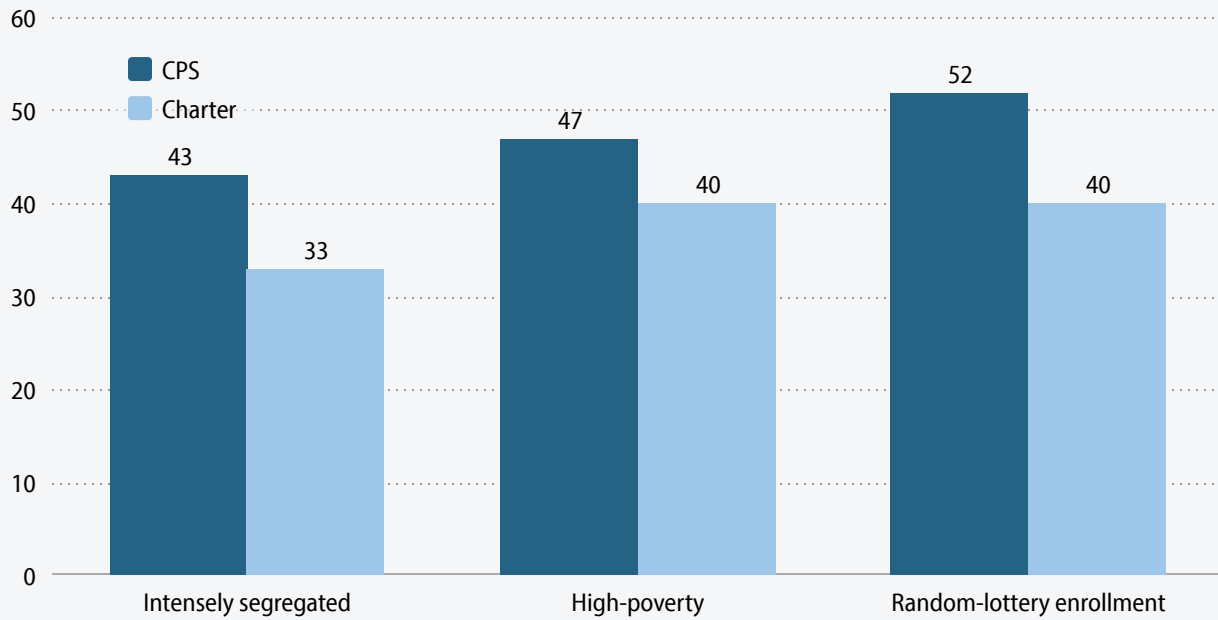
Moreover, to the extent that some charter school students have done better, the gains, as in the city's regular public schools, have largely accrued to nonminority students.⁵⁸

A RAND study published in 2009 noted a caveat to these findings: Charter schools operating for three years or less were the poorest performing charter schools, and students lost more ground their first year in the charter, potentially making small gains (as measured by test scores), depending on subject or year, in subsequent years. Overall, however, the lack of distinction between charters and traditional public schools held (Booker et al. 2009).

Moreover, while the evidence shows that charter schools have no overall impact on district-wide student performance, some studies suggest that they serve fewer high-needs students (Brown and Gutstein 2009, Caref et al. 2012). The CREDO study that found benefits for low-income students in Chicago charter schools relative to comparable neighborhood schools also found disadvantages in math for Hispanic students and in both reading and math for African American and Hispanic students (Brown and Gutstein 2009; CREDO 2009b). **Figure R** compares how well Chicago's traditional public schools and charter schools serve high-needs students, as measured by the "value-added" performance score that CPS uses to determine whether a school is "failing" and should be closed. The value-added score measures

FIGURE R

Chicago school percentile ranking on reading ‘value-added’ score, traditional public vs. charter, by school subgroup, 2012



Note: In high-poverty schools, the share of students eligible for free or reduced-price lunch is higher than the district average of 85 percent. The "value-added" score measures reading test score growth but with an attempt to control for demographic factors and prior achievement. Rankings show in which percentile school subgroups fall in the distribution of scores across all CPS elementary schools.

Source: Caref et al. (2012, 21)

ECONOMIC POLICY INSTITUTE

reading or math test score growth and attempts to control for demographic factors and prior achievement. Schools are ranked in order of their scores. The lower the school falls in the percentile ranking, the higher its risk of closure. By looking at the value-added percentile ranking of the various subgroups of schools most likely slated for closure, the figure depicts how well public schools slated for closure due to poor performance, as defined by the district, compare with the charter schools likely to take their students (Caref et al. 2012).⁵⁹ As Figure R illustrates, among the subgroups of schools that are intensely segregated, have higher-than-district-average subsidized lunch rates, and are non-selective magnet schools, public schools outperform charter schools.

As in New York City, Chicago’s high school graduation rates are a bright spot. The Chicago Public Schools’ graduation rate increased from 53 percent in the 2002–2003 school year to 64 percent in 2007–2008 (NCES 2005a and 2010b). A comprehensive analysis accounting for each subgroup’s initial rate (at the start of the era), discharges from schools, and other factors that can impede interpreting graduation rates reliably found that rates rose for all subgroups of students during the last decade. African American males gained the least, but still a substantial 15 percentage points between 1991 and 2004, compared with 20 points for Latino males and 25 points for white males (Luppescu et al. 2011, 56). The rise began during the 1990s, but increased dramatically during the past decade. Interestingly, however, the increase was only marginally greater in the new schools opened under Duncan than in schools that had existed for 20

years or more. Moreover, “[d]ropout rates at schools on probation—essentially all low-income students of color—have increased” (Lipman and Gutstein 2011, 18).

Charter schools in Chicago also have higher high school graduation rates than regular public schools, though the discrepancy may reflect a number of factors.⁶⁰ A 2009 RAND study, which was funded by a consortium of foundations, found that Chicago high school students in charter schools were 7 percent more likely to graduate and 11 percent more likely to enroll in college than high school students in traditional public schools. Charter students also had higher ACT scores in math and reading. However, the study’s authors cautioned that these findings applied only to students in charter schools serving both middle and high school students.⁶¹ Moreover, black students transferred into charter schools with significantly higher math and reading scores than black students remaining in traditional public schools, making it difficult to determine how much the schools contributed to their higher scores (Booker et al. 2009).

After he realized that some charter schools were graduating students from high school at higher rates than were regular CPS district schools, Duncan worked with Greg Darnieder, head of the newly-created CPS Office of College and Careers, to establish the Department of Postsecondary Education and Student Development in 2003.⁶² This effort incorporated several new initiatives to improve the transition from high school to college and career and increase the number of low-income students enrolling in postsecondary education. These included “[h]olding principals accountable for getting each student to fill out financial aid forms and apply to colleges, and giving high schools college coaches to help students choose and follow through” (Karp 2012a). CPS developed a new website, which provides information and guidance to students regarding high school choices and graduation requirements; provided career and college choices, including career and technical information; and offered to pay for college and scholarship opportunities. The program is credited with substantially increasing scholarships awarded to CPS graduates, from \$147 million in 2011 to nearly \$276 million in 2012 (Karp 2012a). Unfortunately, the future of this promising initiative is in question; budget cuts have sharply reduced the resources allotted to the department.⁶³

Other positive trends during Duncan’s tenure included the 11 percent increase (from 58 percent to 69 percent) in the share of students entering CPS high schools who take the ACT exam by their junior year, and the increase in average ACT scores. While researchers found that an increase in the number of high-achieving students accounts for some of the improvement, African American and Latino students’ scores also rose, albeit at a slower pace, leaving a larger gap. ACT scores increased the most in selective-enrollment schools, second most in charter schools, and least in vocational and general education schools, the latter of which represents by far the largest segment of CPS high schools (Luppescu et al. 2011, 45).

Reports on “college readiness” among CPS students showcase the complexity of that subject. Duncan promoted college readiness as a key goal for low-income and minority students. But, likely due partly to his own efforts to boost students’ access to higher-level courses, many took such courses without foundational skills, and the results look dismal. The Consortium on Chicago School Research report notes that, “[while the growth is encouraging], scores remain below district goals for college and career readiness. ... At the current rate of increase, it would take another 17 years before the average ACT score in Chicago reached [the district average goal established by Duncan of] 20” (Luppescu et al. 2011, 44).

The report cited several pieces of evidence that lack of college readiness remains a major concern across CPS, particularly among black and Latino students. For example, the authors noted the gap between college preparatory high school course completion rates of white/Asian students and black/Latino students. They also found that while “many more students [took] college preparatory math and science courses [during this period], ... there were no accompanying improvements in college-going outcomes.” While the percentage of students who took and passed more than one Advanced Placement (AP) class rose to 15 percent, only about one in three of those students passed the test needed for college credit (Luppescu et al. 2011, 62). And while AP course pass rates increased among all groups, students’ incoming achievement gaps were reflected in race-based differences in those pass rates, which in 2008–2009 ranged from just over 10 percent among African Americans to 16 percent among Latinos, 30 percent among white students, and over 50 percent among Asians (Luppescu et al. 2011, Figure 38).

The Commercial Club of Chicago report also cited low college readiness rates as a problem.

The ACT tests results show the percentages of 11th graders who meet “college readiness” benchmarks (as established by ACT) in math and science are tiny: 16% in math, and 9% in science; most of these are in Chicago’s few “selective enrollment” schools. When one looks at the non-selective enrollment high schools—those which serve the neighborhoods of Chicago—the percentage of 11th graders ready for college are even lower [than indicated by overall numbers]: 6.4% in math, and 2.3% in science. In many high schools, not one 11th grader is on track to succeed in a college-level math or science course. (Civic Committee of the Commercial Club of Chicago 2009, 1)

In summary, while there have been small gains in Chicago Public Schools’ math scores, virtually all of those gains accrued to higher-income, nonminority students. Among low-income students, test preparation also replaced more in-depth instruction and enrichment that is critical for college, career, and life readiness. There were virtually no improvements in reading, and African American students lost a bit of ground. High school graduation rates increased, though less so for African American and male students than others, and no more in the new schools opened under Duncan than in older ones. Finally, while ACT test-taking and scores increased, college- and career-readiness goals remain a distant goal. Some of Duncan’s efforts to improve the odds for at-risk students in reaching college seem promising, but will require less grandstanding and more patience than current Chicago and national reformers are demonstrating.

It is possible, for example, that these low college-readiness numbers in early years of a reform effort reflect a necessary hurdle. Because Duncan was aware that minority students’ failure to take Advanced Placement courses hurt their chances of enrolling in college, he promoted AP class-taking, knowing that less-prepared students might not, at least initially, fare well, but that this was a necessary first step in remedying the gap. As discussed earlier, he also created a new department to address some of the other barriers to low-income students’ successful completion of high school and their smooth transition to college and careers. The lesson here is for policymakers to be patient: Improving such a longstanding and complex problem as inadequate college readiness among groups of students who have lacked not only higher-order, but foundational academic skills, requires multiple strategies unfolding across many years. Some indicators may even trend downward before they begin to improve.

The reforms have not saved money

In the aftermath of the recession, states and school districts continue to struggle with fiscal crises that force difficult choices regarding class size, the range of courses offered, teacher quality and experience, administrative capacity, and various support services. In the midst of this, delivering quality education without harming at-risk students is a growing challenge.

Cleveland, Detroit, and Philadelphia are among many cities asserting that they will save money by adopting a market-oriented set of education reforms like those assessed in this paper.

While the reformers who led the three cities examined here did not claim such savings, their experiences are nonetheless instructive. This section explores the budgetary trends in the three cities during the reform period and compares them to budget trends across large urban districts generally—an assessment complicated by the fact that budgets in all three cities rose substantially over much of the period of the reforms but then declined due to the recession and its aftermath.

An analysis of budget changes is relevant for additional reasons. First, school closures are often presented as cost-cutting measures, but as this section shows, they do not inevitably save money, a reality that the many other districts planning to close schools should take into account.

Second, while the level of spending on education is not the main predictor of student outcomes, it is reasonable to expect that districts that increase spending more than others will reap larger-than-average gains for their students. If not, policymakers would want to question the rationale for the spending, and to redirect it to more efficient uses, especially in this time of fiscal constraints.

Third, new sources and types of spending—and conditions for their use—have attended these policy changes. As the following analyses illustrate, the influx of foundation and corporate dollars into schools in all three cities raises critical questions regarding how democratic is decision-making in these cities and how replicable are the reforms for other districts considering similar policies.

Washington, D.C.: Increased spending without public and parental input produced few benefits

In Washington, D.C., spending increased substantially, but through mechanisms that depressed public and parental decision-making, and to unsustainable levels. Nonetheless, as shown earlier, benefits to students were hard to detect.

During the 2010 fiscal year, District of Columbia Public Schools spent \$18,667 per student.⁶⁴ This was second only to New York City, which spent \$19,597 per student. In comparison, the national per-pupil average for the 2010 fiscal year was \$10,615 (Dixon 2012). According to the most recent data available across the 100 largest public elementary and secondary school districts, average per-pupil spending was \$12,571 for the 2008–2009 school year (NCES 2010a).

DCPS per-student spending from all funding sources rose by 27 percent in real dollars, or \$3,744, between 2007 and 2011. This increase occurred at a time when most other school districts were cutting budgets in response to the recession. Indeed, in 2011 the Center on Budget and Policy Priorities reported that, of 46 states for which state education data could be compared, “37 [provided] less funding per student to local school districts than the year before, 30

[provided] less than they did four years ago, 17 states [had] cut per-student funding by more than 10 percent from pre-recession levels, [and four states had] reduced per-student funding by more than 20 percent” (Oliff and Leachman 2011).

Some of the increase was enabled by foundation grants (\$20 million), increases in special congressional earmarks (\$27 million), and federal entitlement grants (\$6 million).⁶⁵ In 2010, the Walton, Robertson, Laura and John Arnold, and Broad foundations pledged an additional \$64.5 million over three years on the condition that the money be used to support salaries for excess teachers not hired by other principals and for performance bonuses (DC Public Education Fund 2010). Additional foundation supporters include the Bill & Melinda Gates, Carnegie, Dell, and CityBridge foundations. The foundation money was slated to run out in 2012; it would need to be replaced if the merit pay it supported was to continue. When the 2010 contract with the Washington Teachers’ Union authorizing the IMPACT merit pay program was approved, Rhee said that “savings and efficiencies achieved elsewhere in the schools system—such as reductions in private special education tuition and transportation—could be used to help sustain the contract” (Turque 2012b). But that did not happen; instead, the costs have been passed down to individual schools, and 2012–2013 proposed budget allocations included larger class sizes and cuts to positions such as special education coordinators.

Spending on research-backed, effective programs, especially for low-income students, can provide a substantial benefit. However, the literature review above points to a lack of such evidentiary backing for these reform policies’ effectiveness. And DCPS students’ meager gains, despite large spending increases, indicate that these policies are no more effective in best-scenario real-life settings.

News media reports suggested that spending did not significantly improve outcomes in part because it did not go to the right places. First, as discussed earlier, even as the district increased overall and per-pupil spending each year, Rhee cited budget crises as the rationale for firing experienced teachers.

In a radio program in September 2011, budget expert Mary Levy decried what she characterized as an increasing lack of accountability or transparency in the budget, and the apparent failure of the new dollars to go to schools, and classrooms:

[The DCPS budget] had a \$77 million increase this year and yet they cut spending in the local schools. We do not know where that extra money is going, but they moved a number of functions out of central office accounts and charged the schools for them, while cutting the schools’ per-pupil funding. So there’s a tremendous amount of money some place, but I haven’t been able to find out where. (Levy 2011a)

Among the costs for services shifted from the central office to individual schools over this period, were funding for special needs programs, social workers, high school ROTC instructors, and school psychologists, Levy said.⁶⁶

Secondary school class sizes also increased, while librarians, special education coordinators, and school supplies and materials were cut (Barry 2012). At Wilson High School, the Advanced Placement biology teacher died and was replaced by a substitute who had never taught at the AP level.

In 2012, *Washington Post* education reporter Bill Turque criticized this lack of budget transparency in his report on the fallout for schools as private foundation support for teacher merit pay ended: “Where the rest of the \$64 million has

gone is not clear. I've asked for clarification from DCPS. A District study at the time of the contract's approval estimated that without continued private help, the annual cost of supporting the 2010 contract [with merit-based bonuses built in] was about \$30 million" (Turque 2012b).

The foundations that helped initiate IMPACT-based merit bonuses stipulated in a letter that they could withdraw their money if Rhee no longer led DCPS (DC Public Education Fund 2010).⁶⁷ These conditions for private support raise critical questions regarding who is making, and should make, key policy decision about DCPS schools, teachers, and students, as well as the role of the taxpaying public in choosing the leader for the school system that they fund. Perhaps more troubling, the letter links DCPS's continued receipt of the funding through the three years to requirements that it increase both student test scores and teacher retention (in essence, immediately).

Given the funding commitments were based on a belief that [these test-based methods for evaluating and making decisions about teachers] will have a significant impact on the outcomes of your reform efforts, DCPS will be required to report on key metrics and outcomes, such as student achievement growth and teacher retention. In the case that the anticipated outcomes are not being realized, the third party funders reserve the right to reconsider their support. (DC Public Education Fund 2010)

Moreover, closing underenrolled schools did not produce the budget savings Rhee had predicted. As noted above, Rhee moved to close schools mainly as a savings strategy, but the initial estimate of under \$10 million to close 23 schools in 2008 turned out to be wildly inaccurate. The final price tag was nearly \$40 million, or quadruple the estimate (Gartner 2012a). Also as noted, some of the cost savings reported in the DCPS budget were really shifts of costs from DCPS to families; families' transportation costs, for example, some of which are not covered by the district, rise as students commute longer to schools that are further away.

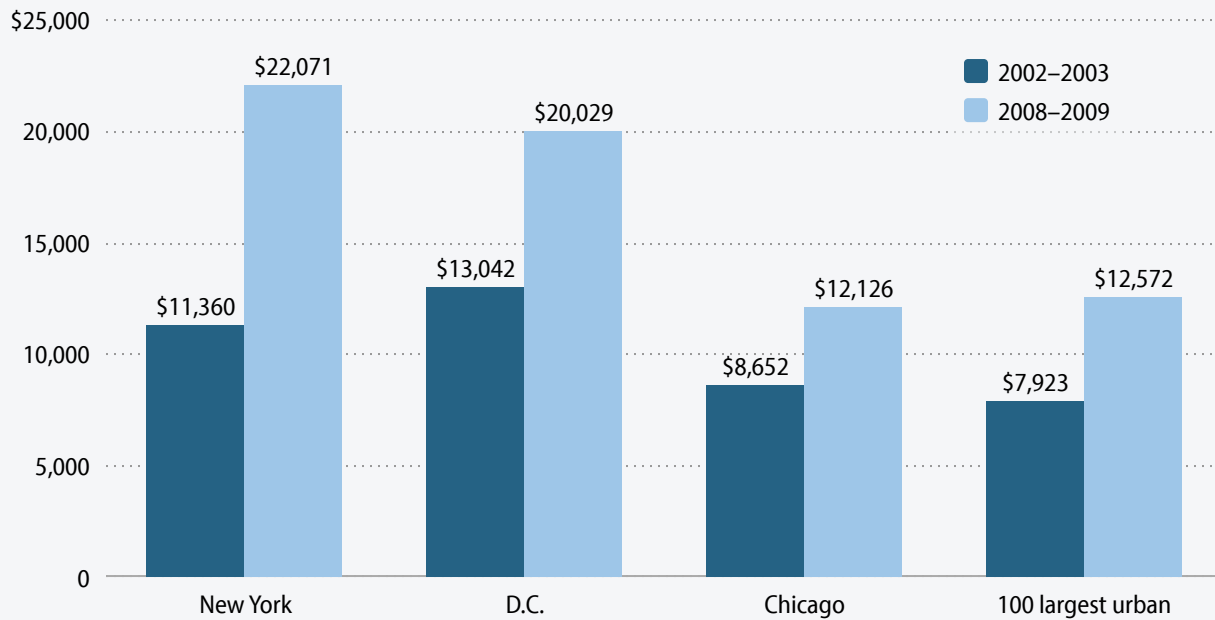
New York City: Reforms hinged on likely unsustainable spending increases

New York City has long been among the U.S. school districts with the highest per-pupil spending; it spent \$19,597 per pupil in 2010, the most recent year for which data are available (U.S. Census Bureau 2012, 104). During Bloomberg and Klein's control of the schools, spending increased at a far greater rate in New York City than it did, on average, across the 100 largest U.S. school districts tracked by the National Center for Education Statistics. In real dollar terms, New York City's per-pupil education spending nearly doubled between the start of Bloomberg and Klein's control of the district (2002–2003) and the 2008–2009 school year. The city spent \$11,361 per-child in real dollars in 2002–2003 and over \$22,000 by 2008–2009. As **Figure S** illustrates, this increase far surpassed that of the other reform cities and the 100 largest U.S. school districts overall (NCES 2010a).⁶⁸ On average, including New York City (and the other two reform cities studied here), per-pupil spending in large school districts rose from \$7,923 in 2001 to \$12,572, a 59 percent increase.

As demonstrated earlier, the added spending has not translated into improved student outcomes. New York City has, on average, achieved less than other large, urban districts whose students' outcomes are reported, while outspending and out-reforming them. Nonetheless, in April, 2012, Klein and Rhee announced that their StudentsFirst New York chapter would spend \$10 million a year for the next five years to sustain the reform model put in place by Mayor Bloomberg after he leaves office following the 2013 mayoral election (Fleisher 2012b).

FIGURE 5

Increase in per-pupil spending, New York City vs. other reform cities and 100 largest urban districts, from 2002–2003 to 2008–2009



Source: National Center for Education Statistics (2005a and 2010a)

ECONOMIC POLICY INSTITUTE

New York City has also benefited far more in absolute dollars than any other urban district from private donations to its schools and education programs (Saltman 2010, 9).⁶⁹ Under Mayor Bloomberg, the city created a Fund for Public Schools, which is described as “dedicated to improving NYC’s public schools by attracting private investment in school reform and encouraging greater involvement by all New Yorkers in the education of our children” (Fund for Public Schools 2012). Four foundations, including the Bill & Melinda Gates Foundation, and two corporations have each given \$1 million or more, The Broad Foundation donated between half a million and a million dollars, another 19 for-profit and nonprofit groups have given between \$100,000 and \$499,999 each, and dozens more have provided smaller sums (Fund for Public Schools 2010). While reforms hew closely to the policy priorities of the larger donors, such as the Gates and Broad foundations, the Independent Budget Office’s analysis illustrates how little total outside donations account for in the system’s overall budget (IBO 2012a). Donors’ impact is thus hugely disproportionate, relative to that of the taxpaying public.

Chicago: Spending increases mirrored those of other large cities, but donors’ policy influence outstripped their funding

In the 2001–2002 school year, when Duncan began his tenure, per-pupil spending in Chicago was \$8,652 in real dollars. In the 2008–2009 school year, the last year of Duncan’s tenure, per-pupil spending in Chicago was \$12,126, an increase of 40 percent from 2001–2002. Unlike the other two districts studied, Chicago’s per-student spending toward

the end of this period was more closely aligned with that of the average large U.S. school district. Likewise, Chicago's spending increase over this seven-year period was similar to that of the average large U.S. school district.

Nonetheless, the sharp increase in a relatively short period of time, during years when districts had much more money than they do now—and the financial and political influence of corporate donors to Renaissance 2010 and related policy reforms—raises the questions of whether the reforms are replicable in other cities, and of whether they are sustainable.

In Chicago, as in the other two cities studied, corporate philanthropists promoted many of the changes implemented, even though their monetary contributions were negligible relative to the city's school budget overall.⁷⁰ As noted earlier, the Civic Committee of the Commercial Club of Chicago helped to spearhead the Renaissance 2010 portfolio of reforms: opening of new charter schools, school closures based on labels of “failure,” and the replacement of some closed schools with new, selective magnet schools. And as in New York City, a small group of corporate philanthropists, from both the for-profit and nonprofit sectors, created a board to shape school policy, in Chicago's case, the Chicago Public Education Fund.⁷¹ Heavily weighted with financial and banking sector executives, the fund describes its market-based approach as follows:

The Chicago Public Education Fund applies principles of venture capital to make systemic improvements in education, with a focus on talent and leadership in our schools. Our approach is further defining a burgeoning field of private sector investment in public education and is being studied by America's top business schools as a model of innovation in philanthropy. More importantly, the venture capital model enables The Fund to stay focused, and get results fast. (Chicago Public Education Fund 2012c)

The fund's asserted intention of “get[ting] results fast” raises serious concerns: This goal appears to conflict with the complex realities of improving outcomes in large, very low-income urban school districts, where longstanding challenges to student success take time and patience to address. Demands for quick fixes can do much more harm than good.

Moreover, as noted earlier, the increase in spending went disproportionately to reform-oriented projects. Much of the capital spending went to building new facilities and rehabilitating old ones for charters and selective schools. While CPS spending on teacher and administrator salaries grew by a modest 10.6 percent between fiscal years 2004 and 2012, annual spending for debt service increased by 52 percent, capital projects spending increased by 84 percent, and spending for charter schools increased by 624 percent (Caref et al. 2012).

Additionally, CPS has recently shifted from a policy of closing schools on the basis of failure to right-sizing the district (and saving money) by shuttering schools labeled underutilized. However, “CPS has not reaped the benefits of closing schools in the past. Of the 76 schools closed since 2001, the vast majority of them—nearly 80 percent—now house a new school”⁷² (Karp 2012b).

A report by the Chicago Teachers' Union raised doubts that schools labeled underutilized really are (Caref et al. 2012). According to the authors, the school district reported a dramatic increase in excess capacity (from 65,000 surplus spots for students in 2011 to more than 100,000 in 2012). In the same year that it closed neighborhood schools it said were not needed, the district announced plans to “open another 60 charter schools over the next several years” (Caref et al. 2012, 13). CPS has the largest class sizes of any school districts in Illinois, so seats counted as excess in CPS wouldn't be considered so in many other school districts. As well, CPS ignores the different ways in which various classrooms

might appropriately be used, which further inflates “underutilization” numbers by failing to build in space for necessary, non-traditional classroom and school use:

It doesn't matter if the room actually has 18 pre-K students or if it's a self-contained special ed room with five autistic students—those rooms have 12 and 25 empty seats, respectively. There's no accounting for whether the school has a gym or lunchroom or playground, or whether a classroom converted to a science lab can fit the same number of bodies as a classroom full of desks. (Caref et al. 2012, 13)

With an initial assessment of underutilization that is so difficult to assess, calculating cost-savings, or their lack, is necessarily challenging and, also, subject to manipulation on both sides.

The bottom line: More spending, bolstered by private donations, depressed democratic input without helping students

Education spending and student outcomes are not closely correlated for a number of reasons. That said, policymakers are reasonable to expect that substantial budget increases, especially in tight fiscal times, will improve outcomes for the students they target. As detailed in this report, the data from these cities fail to demonstrate such improvements. The low-income and minority students who were supposed to benefit the most from market-oriented reforms gained the least and, in some cases, lost ground. It is not even clear that the added money was directed to at-risk students. Rather, funds went to larger central offices at the expense of classrooms (as in Washington, D.C.) or to the small share of relatively better-off students who attend charter schools, small schools, or selective magnet schools with more resources than their neighborhood public school counterparts (as in New York City and Chicago). And even if the reforms had been more effective, they could not easily be replicated in other districts, because they hinge in part on financial resources that are no longer available.

Tight fiscal times also demand that schools focus on programs demonstrated to improve student achievement and narrow gaps. While smart investments can, and do, improve outcomes, one reason that education spending often fails to deliver improvements is the failure of policymakers to follow the evidence. Low-income and minority students (and thus their schools) need more support than their peers from wealthier families because they are less likely to have the family supports and resources that foster learning. For example, studies show that low-income children reap the largest benefits from quality early childhood education programs. Similarly, low-income students in the early grades seem to benefit more than their wealthier peers from smaller class sizes. Other evidence-based investments include health care (such as clinics in schools); nutritional supports such as those from Women, Infants, and Children (WIC) and the Supplemental Nutrition Assistance Program (SNAP), and subsidized school meal programs. Afterschool and summer enrichment help narrow income-based achievement gaps and boost low-income students' performance. As illustrated above, however, low-income students often lacked access to these critical supports; many were taught by less experienced teachers in larger classrooms, had less engaging classroom experiences, and saw other supports reduced due to budget cuts.

Most states are unlikely to substantially increase their education budgets right now, but they can shift existing resources from less effective initiatives to those above, which would produce a stronger return on investment. Even if they did make the investment, however, they would be building on a shakier foundation. Due to the recession and to a focus on reforms, in many districts, class sizes have increased, the teacher pool is less experienced, and pre-K and extracurricular enrichment have been cut.

Finally, the budget increases to fund the reforms in all three cities were driven, though not largely enabled, by private donors' contributions, often with substantial policy strings attached. Such funding raises important questions. First, are programs and policies based on decisions by philanthropists and foundations scalable and sustainable, or will districts face even greater ups and downs than the recession has brought? The example of Washington, D.C., suggests that this can become a real problem, with donors promoting programs that they subsequently drop, and that are then sustained by public dollars siphoned from other priorities. (Many districts may not be able to attract outside funding to begin with.) Second, does the imperative to attract corporate backers stifle the public voice? The explicit goals of the privately managed reform funds in all three cities suggest that there is real cause for concern. When school districts are couched as "portfolios" for investment, is there space for the focus on civic engagement, social and emotional development, and other critical aspects of education that society deems important? Growing teacher and parent backlash in New York, Chicago, Philadelphia, and other cities highlight the need to reconsider whose voices should be front-and-center when major changes are made.

Both academic and real-world evidence strongly counter the Chicago Fund's goal, and the goal of venture capitalists generally, of "get[ting] results fast." As set out above, substantial, sustained improvements in high-poverty schools and the students they serve require that those who understand education best are given the time, space, and resources to address the root causes of problems comprehensively. It is not clear that increased private investment allows, let alone promotes, such priorities.

Conclusion

Turning around a troubled school is a complex and difficult process, as Anthony Bryk and his colleagues discovered in their in-depth study of Chicago reforms. Turning around many troubled schools in large, high-poverty districts is even more complicated. Success is reflected not only in higher student test scores. It is reflected in teachers who are motivated, supported, and work as a team; in a curriculum that is engaging, comprehensive, and hands-on; in parents who are meaningfully involved in supporting their children's achievement; and in schools that serve as integral parts of their communities. Determining success thus requires measures that go far beyond test scores; achieving success requires strategies that reflect those goals and measures.

While education leaders sometimes acknowledge these realities, current "reform" agendas at the federal level and among high-profile advocacy groups do not reflect that recognition. Instead, education policy has been dominated in recent years by a narrow set of market-oriented reforms, and success or failure correspondingly measured by a small set of quantifiable metrics that tell only a slice of the story, often in a distorted manner. In Chicago, New York City, and Washington, D.C., where the reforms played out most comprehensively, substantial gaps between the rhetoric and reality of the market-oriented reform agenda are apparent.

Reformers did not deliver on their promises. Reformers asserted that their policy changes would improve academic achievement, especially among low-income and minority students, and thus narrow achievement gaps. As the report documents, the promised outcomes failed to materialize. Test scores increased no more than in other large, urban districts that did not implement the same reforms, and, in most cases, increased less in the three cities' districts. For example, New York City had the second-smallest average test score gains among 10 large, low-income urban districts studied.

Reforms were framed as specifically targeting low-income and minority students. But these students made less progress than similar students in other urban settings or than their counterparts in years before reforms were implemented. For example, in Chicago, low-income, black, and Hispanic students gained virtually no ground—and in a few cases lost ground—while their peers in other cities did increase their test scores.

To the extent that gains were made in the three cities, they tended to accrue to those students who were better off to begin with. Thus, despite real gains in a few areas, gaps still widened, leaving at-risk students even less able to compete with their higher-scoring, more affluent peers. In Washington, D.C., for example, race-based achievement gaps that were already the country's largest grew still bigger under Rhee.

Reformers made false claims. Reformers in all three cities proclaimed success, reporting higher test scores and smaller achievement gaps. New York City Mayor Michael Bloomberg claimed to have cut race-based achievement gaps in half in less than a decade. As Chicago schools CEO, Arne Duncan said that Chicago's students had risen from far below-average to far above-average in their rates of "proficiency" in a similarly short period of time (President Obama has repeated these assertions). In both New York City and Chicago, purported gains evaporated when proficiency standards were recalibrated to reflect changes in the design of tests. And in Washington, D.C., initially positive test score trends reversed when an investigation into widespread cheating began, and asserted gains disappeared when reliable data were employed and subgroup differences assessed.

Four fundamental flaws in the reform agenda led to the unfulfilled promises. First, the promises are unrealistic. No set of policies, no matter how effective, could raise test scores, especially for students confronting the multiple barriers posed by poverty, or narrow achievement gaps to the extent predicted and claimed. High-quality early childhood education programs, widely seen as one of the most cost-effective educational interventions available, can reduce the odds of receiving special education services by 10 percentage points, and increase the odds of graduating from high school by an estimated 11 percentage points (Lynch 2007).⁷³ The most effective summer enrichment programs have been found to reduce the odds of dropping out of high school by as much as 20 percent among very engaged participants (Heckman and Sanger 2013).⁷⁴ There is no basis to believe that other reforms can do more to close achievement gaps than quality early childhood and summer enrichment programs, and the vast majority will do less.

Second, many aspects of the reform agenda are counterproductive; their practical impacts thwart some of their own objectives. As discussed in this report, increasing the science, technology, and engineering components of STEM education to produce more engineers and computer programmers is difficult when raising reading and math scores assume such high priority, and thus crowd out other subjects. The same is true of the higher-order critical thinking and creativity required to forge productive workers and good citizens; attaching high stakes to tests that assess basic skills all but guarantees that more complex learning falls by the wayside. Indeed, the Chicago Consortium on School Research documents the replacement of real instruction with test preparation in low-income CPS schools in particular. Finally, as scholars warned, and as the outcomes above illustrate, reliance on student test scores to evaluate teachers and schools strongly dissuades the very teachers sought for struggling schools from working in them. The dearth of experienced teachers in District of Columbia Public Schools, and the relative inexperience and lack of qualifications of teachers in low-income and minority schools in districts across the country, point to the need to rethink both what constitutes teacher quality and how to attain it.

Indeed, as eight New York State Teachers of the Year wrote in their letter to the state Board of Regents about the board's decision to evaluate teachers based on student standardized test scores:

These changes, while politically popular, will neither improve schools nor increase student learning; rather, they will cause tangible harm to students and teachers alike. ... We could quote the research of educational experts like Diane Ravitch, Richard Rothstein and Jonathon Kozol as to why poverty and parental support affect test scores significantly more than any curricular changes a school can provide. We could refer to myriad child psychologists who have documented the deleterious effects of high-stakes testing on our nation's youth. We could call upon assessment experts who insist that standardized tests were not developed to evaluate teacher effectiveness. And we could examine the last decade of educational results that followed No Child Left Behind: rampant gaming of the system to provide the appearance of growth, narrowing of the curriculum, excessive teaching to the test and virtually no change in the achievement gap. All of the above would lead the reasonable person to be skeptical about using standardized tests as the engine for school reform. Worse yet, we fear that the competition generated by this approach will reduce the collaboration necessary for true school improvement (Strauss 2011).

Third, the reforms have caused real harm to some students. Again, the highest-risk students have suffered the most from the upheavals reforms have produced. It is the students in under-resourced schools, who have lost literature and poetry to vocabulary drills and seen their curricula stripped of art, music, and physical education to make room for increased test preparation, who are most likely to see their schools shuttered when their test scores do not rise quickly enough. These children face longer and potentially more dangerous commutes to school, the loss of a neighborhood hub, and, sometimes, repeated disruptions of friendships and relationships with teachers.

Fourth, the market-oriented reforms implemented in these cities are no match for the complex, poverty-related problems they seek to solve. This is evident in the continued and, in some cases, expanded gaps between low- and higher-income students and between minority students and their white counterparts. As Bryk and his colleagues also found in their in-depth exploration of 20 years of Chicago school reform efforts, poverty—at the child, family, and community levels—poses a major impediment to student and school success that these narrow reforms fail to address (Bryk et al. 2010). As the findings in this report affirm, African American students have long been concentrated in schools that have made the smallest achievement gains, a reality that is perpetuated by “reforms” that fail to address the root causes of underachievement. According to CCSR research, schools serving communities with little social capital have the hardest time improving: “Generally African Americans live in neighborhoods that have fewer social resources for school improvement,” while “neighborhoods that have more resources bring them to bear on the schools” (Karp 2011). It is hard to imagine a “no excuses” mantra holding up in schools serving what Bryk calls “students living under extraordinary circumstances,” schools in which one in four children has a substantiated case of abuse or neglect, where a typical classroom of 30 has teachers engaging with seven or eight such students every year (Bryk et al. 2010, 176).

“These changes, while politically popular, will neither improve schools nor increase student learning; rather, they will cause tangible harm to students and teachers alike.”

These reforms should not serve as models for other districts. Cleveland, Detroit, Philadelphia, and the many other cities that are implementing or considering market-oriented strategies have no reason to think that they will increase student or school success by relying heavily on student test scores to evaluate teachers, principals, and schools; closing “failing” or underenrolled schools; or treating charter schools as a fix for “dropout factories.” Indeed, in Ohio and Pennsylvania, it is already clear that charter schools perform significantly worse than comparable regular public schools, so students would lose ground, rather than advance, by making that switch. Given the lack of excellent schools to receive students whose neighborhood schools are shuttered, policymakers should not promise unrealistic improvements and should assess very carefully predicted cost savings, which have failed to materialize in other districts.

The reforms may actually be costly. Charter schools tend to be smaller than regular district schools, and thus less cost-efficient by that metric. Findings in these three cities suggest, too, that effective charters, such as those in New York City, may cost more per pupil, and rely on private funding to close that gap. When the compositions of their student bodies are considered, charters are substantially more expensive. Systemic disruptions when charters co-locate in regular public school buildings, and when students awaiting acceptance move from public to charter schools after the school year begins, can further increase the district’s total education budget. Merit-based teacher pay adds to the budget, especially if the assumption is that relatively small bonuses are insufficient to motivate changes in behavior. Moreover, to the extent that philanthropic support has enabled some of these policies, other districts are unlikely to have access to such substantial donor funds. Even in these three cities, some of the policies, such as DCPS salary increases, have already proven hard to sustain, and other initiatives that rely on outside support, such as New York City’s small schools, do not appear likely to increase in scale to serve a substantial share of the city’s students.

Mayoral control, touted as key to reformers’ success in these three cities, is no panacea. When employed wisely, mayoral control of a school district can be used to bridge silos across departments and agencies, and to create a more holistic system of student supports that complements in-school improvements. In these cities, however, it has led to substantial opposition by parents who feel ignored and teachers who feel shut out of the reform process, but little noticeable boost to students or schools.

Reform must be more realistic, patient, and multipronged if it is to achieve real, sustained change. Achievement gaps have their root in opportunity gaps. Only by closing the latter can we begin to shrink the former. The promising “nuggets” in each city illustrate this reality. New York’s small schools offer low-income students opportunities normally available only to those who are much wealthier. Chicago’s college- and career-readiness initiative provides low-income and minority students with good advice about high school courses, scholarship opportunities, and the range of post-secondary options, so that they, like their more affluent peers, can pursue schooling that leads to careers matching their skills. The District of Columbia’s high-quality universal pre-K program employs a holistic, hands-on curriculum designed to boost all aspects of children’s development—physical, mental, emotional, and cognitive—in ways that mirror programs enjoyed by children of professional parents. Effective reform must recognize the huge impact of community factors, leverage the community’s resources, and establish supports to compensate for gaps. Without such supports, gaps in kindergarten readiness, physical and mental health, nutrition, and extracurricular enrichment opportunities will continue to thwart even the most effective reforms that stop at the classroom and school walls.

[There is] compelling evidence that school community factors can act as potent catalysts for the broad-based organizational developments needed to improve student outcomes. The presence of community social capital

BROADER STRATEGIES OFFER MORE PROMISE THAN MARKET-BASED REFORMS

While reformers continue to advance market-oriented policies with few benefits, more holistic strategies with real promise in the three cities have failed to receive the attention or funding needed to have a real impact.



100 new small NYC schools created in 2003–2005 focused on ensuring strong, consistent student-teacher relationships; leveraging community partners for extra staff, coaching, and resources; and providing hands-on learning experiences, such as internships at law firms and seeding oyster beds. These schools reportedly increased the share of ninth-grade students on track to graduate and high school students' college readiness.



To attract more high-quality teachers to Chicago Public Schools, then-CEO Arne Duncan identified the strongest teacher-preparation programs and encouraged CPS to hire from them, moved recruitment dates up, established job fairs to boost recruiting ability, and offered new teachers higher starting salaries. This improved teacher quality and reduced inequities across districts.



Duncan worked to improve low-income and minority students' college readiness by increasing their access to AP courses, putting college counselors in low-income high schools to help students choose courses and schools and match their goals with skills, and holding principals accountable for ensuring that students applied for financial aid, which nearly doubled in one year. Budget cuts have since removed counselors from almost all schools.



DCPS expanded its full-day voluntary prekindergarten program to serve 3- and 4-year-olds at all income levels, and adopted a holistic curriculum designed to nurture all domains of children's development. But Michelle Rhee enacted obstacles to enrollment, and although it boosted DCPS third-graders' reading scores, pre-K is not even a component of the agenda on which Rhee's advocacy group, StudentsFirst, grades every state's education system.

facilitates the work of a school community at enhancing professional capacity, forging vital links to the parents, creating a healthy climate for children, and strengthening instruction. Absent these social resources and confronting dense community problems, while it is still possible for schools to improve, as a few did in Chicago, the barriers appear almost insurmountable. (Bryk et al. 2010, 189–190)

About the authors

***Elaine Weiss** has been national coordinator of the Broader, Bolder Approach to Education (BBA) since 2011. Her public policy background has focused on early childhood education, specifically on policy mechanisms to help low-income families and children escape poverty. In her role with BBA, Weiss has coordinated and led events at the Economic Policy Institute and on Capitol Hill, developed research and advocacy materials, launched a social media presence for BBA, and worked with numerous allied organizations to increase the visibility of comprehensive strategies in education policy. Weiss has a J.D. from Harvard Law School and a Ph.D. in public policy from the Trachtenberg School of Public Policy and Public Administration at The George Washington University.*

***Don Long** has been a consultant for BBA since November 2011. In that capacity, he has conducted research, drafted and edited documents, and identified much of the evidence for this report. Prior to working as an independent education consultant in Virginia, Long was director of the State Collaborative on Assessment and Student Standards (SCASS) and program manager at Pearson Educational Measurement in Austin, Texas. Long has a Master of Public Affairs degree from the LBJ School of Public Policy at the University of Texas, Austin.*

*The **Broader Bolder Approach to Education** is a national campaign that acknowledges the impact of social and economic disadvantage on schools and students and proposes evidence-based policies to improve schools and remedy conditions that limit many children's readiness to learn. It was launched in 2008 by the Economic Policy Institute, but is guided by outside co-chairs and an independent Advisory Council that shape policies distinct from those of EPI.*

Appendix tables

TABLE A.1

4th-grade NAEP reading scores in three cities compared with large urban and national averages, by race and income, 2005 and 2011

	2005		2011		2005		2011	
	White	Black	White	Black	Non-low-income	Low-income*	Non-low-income	Low-income*
<i>Washington, D.C.**</i>	252 (252)	187 (186)	255 (255)	191 (191)	215 (216)	183 (183)	234 (234)	188 (188)
<i>New York</i>	226	206	235	209	230	210	242	214
<i>Chicago</i>	225	190	229	197	222	194	230	200
<i>Large urban average</i>	228	196	232	202	226	198	232	204
<i>National average</i>	228	199	230	205	230	203	234	207

* Students categorized as low-income are those eligible for free or reduced-price lunch under the National School Lunch Program.

** The scores in parentheses are from DC budget consultant Mary Levy's NAEP data, which, unlike TUDA NAEP data, always include charter schools.

Note: Large, urban districts are those participating in the NAEP Trial Urban District Assessment (TUDA), which uses NAEP scores from a representative sample of students in large, urban districts with high proportions of low-income and minority students. TUDA began in 2002 with six districts and had 21 as of 2011.

Source: National Center for Education Statistics (2005b, 2011b) and Levy (2012c)

TABLE A.2

4th-grade NAEP math scores in three cities compared with large urban and national averages, by race and income, 2005 and 2011

	2005		2011		2005		2011	
	White	Black	White	Black	Non-low-income	Low-income*	Non-low-income	Low-income*
<i>Washington, D.C.**</i>	266 (266)	207 (206)	272 (272)	212 (212)	229 (229)	206 (204)	250 (250)	211 (211)
<i>New York</i>	245	222	248	226	243	228	253	233
<i>Chicago</i>	243	208	246	217	237	212	246	221
<i>Large urban average</i>	247	217	251	222	246	221	250	227
<i>National average</i>	246	220	249	224	248	225	252	229

* Students categorized as low-income are those eligible for free or reduced-price lunch under the National School Lunch Program.

** The scores in parentheses are from DC budget consultant Mary Levy's NAEP data, which, unlike TUDA NAEP data, always include charter schools.

Note: Large, urban districts are those participating in the NAEP Trial Urban District Assessment (TUDA), which uses NAEP scores from a representative sample of students in large, urban districts with high proportions of low-income and minority students. TUDA began in 2002 with six districts and had 21 as of 2011.

Source: National Center for Education Statistics (2005b, 2011b) and Levy (2012c)

TABLE A.3

8th-grade NAEP reading scores in three cities compared with large urban and national averages, by race and income, 2005 and 2011

	2005		2011		2005		2011	
	White	Black	White	Black	Non-low-income	Low-income*	Non-low-income	Low-income*
<i>Washington, D.C.**</i>	301 (303)	235 (233)	290 (290)	231 (231)	249 (250)	234 (232)	259 (259)	228 (228)
<i>New York</i>	269	241	271	248	266	249	267	253
<i>Chicago</i>	270	240	271	245	264	246	273	249
<i>Large urban average</i>	270	240	273	245	264	243	271	248
<i>National average</i>	269	242	274	249	270	247	275	251

* Students categorized as low income are those eligible for free or reduced-price lunch under the National School Lunch Program.

** The scores in parentheses are from DC budget consultant Mary Levy's NAEP data, which, unlike TUDA NAEP data, always include charter schools.

Note: Large, urban districts are those participating in the NAEP Trial Urban District Assessment (TUDA), which uses NAEP scores from a representative sample of students in large, urban districts with high proportions of low-income and minority students. TUDA began in 2002 with six districts and had 21 as of 2011.

Source: National Center for Education Statistics (2005b, 2011b) and Levy (2012c)

TABLE A.4

8th-grade NAEP math scores in three cities compared with large urban and national averages, by race and income, 2005 and 2011

	2005		2011		2005		2011	
	White	Black	White	Black	Non-low-income	Low-income*	Non-low-income	Low-income*
<i>Washington, D.C.**</i>	317 (318)	241 (239)	322 (322)	249 (249)	261 (261)	241 (239)	278 (278)	246 (246)
<i>New York</i>	286	257	292	262	286	264	288	270
<i>Chicago</i>	281	245	296	260	275	254	289	267
<i>Large urban average</i>	288	250	295	261	282	256	293	266
<i>National average</i>	288	254	293	262	288	261	295	269

* Students categorized as low income are those eligible for free or reduced-price lunch under the National School Lunch Program.

** The scores in parentheses are from DC budget consultant Mary Levy's NAEP data, which, unlike TUDA NAEP data, always include charter schools.

Note: Large, urban districts are those participating in the NAEP Trial Urban District Assessment (TUDA), which uses NAEP scores from a representative sample of students in large, urban districts with high proportions of low-income and minority students. TUDA began in 2002 with six districts and had 21 as of 2011.

Source: National Center for Education Statistics (2005b, 2011b) and Levy (2012c)

Endnotes

1. Both Cleveland and Detroit leaders make similar claims. Cleveland’s plan asserts that it can address its substantial budget deficit without asking the state for more money, despite declining state support in recent years, because schools will garner more money through improved efficiency: “The district will also identify ways to authorize the transfer of some locally generated tax revenues to charter-operated Transformation Schools that are sponsored by or have agreements with the district. This shift in funding to schools will incentivize schools to maintain excellence and grow enrollment; more pupils means more dollars. Schools will have to attract and retain students, which will require an intense focus on customer service, school safety, enrichment offerings and academic performance” (Jackson 2012, 10). Michigan Governor Rick Snyder’s website states that, “Across Michigan, eight emergency financial managers are working to rescue cities and school districts from the brink of financial ruin—and they’re succeeding. One of those great success stories is happening in the Detroit Public School System, where emergency financial manager Roy Roberts has instituted reforms that have improved the quality of education for students in the district while also saving money, balancing the budget, and sharing success with teachers and other employees” (Brownfield 2012).
2. Indeed, Mayor Bloomberg claimed that policy changes under his and Klein’s leadership had already narrowed by half New York City’s race-based achievement gap.
3. VAM also have many other limitations, most prominent among them the difficulty of deciding how to evaluate teachers in non-tested subjects and how to assess teachers in their first year, when no prior year scores are available to compare to the new ones (VAM has been found to be especially unstable when only a single years’ worth of scores are employed, so some systems may employ three years, rendering many new teachers ineligible).
4. As the seven most recent New York State teachers of the Year and the 1993 Teacher of the Year wrote to the state Board of Regents, “It is with sadness, pain and frustration that we write this letter. We are deeply concerned about recent changes to the State Education Department’s Annual Professional Performance Review system. These changes, while politically popular, will neither improve schools nor increase student learning; rather, they will cause tangible harm to students and teachers alike. . . . To illustrate the challenges of the new APPR system, we offer these stories from our schools: . . . Trudy is a veteran teacher. She volunteered to teach a class of at-risk learners because she has the skills to do so. Her passing rate on the Regents exam will be significantly lower than her peers teaching the stronger students. Under the new APPR, what motivation will teachers have to take on the most challenging students?” (Peneston et al. 2011)
5. As the Fair Test blog authors wrote, “Widespread cheating is an inevitable consequence of overuses of high-stakes testing, as predicted by renowned social scientist Donald Campbell. In 1976 he wrote in what is now called “Campbell’s Law, “*The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor. . . when test scores become the goal of the teaching process, they both lose their value as indicators of educational status and distort the educational process in undesirable ways*” (Wikipedia 2013).
6. In Chicago, the more comprehensive Teacher Advancement Program (TAP) not only provided (smaller) merit-based bonuses, but opportunities for advancement, such as teacher mentors, that came with larger salary increases. Nonetheless, Mathematica’s rigorous evaluation of the program’s first two years of implementation in 10 Chicago schools found no improvement in student outcomes or better teacher retention rates (Glazerman and Seifullah 2010).
7. While “failing” and “underenrolled” schools often overlap, school reformers tend to cite either one or the other as the principal rationale for targeting a school for closure.
8. The specific definition of a “failing” school varies from district to district, but is based heavily on student test scores. The same is true of schools that are labeled “high-performing”—they have above-average, often very high, student test scores. The Broader

Bolder Approach rejects the notion that a school's effectiveness in educating its students can be judged solely, or mainly, on student test scores, which are highly correlated with and contingent on student body income and other socioeconomic factors. Indeed, in the vast majority of cases, "failing" schools serve highly disadvantaged student populations, while "high-achieving" schools tend to serve higher-income and largely nonminority students. For the purposes of this paper, however, we use the terms employed by the reformers, in keeping with other assumptions that reformers make.

9. The states (or representing cities) studied were Arkansas, Arizona, California, Colorado (Denver), the District of Columbia, Florida, Georgia, Illinois (Chicago), Louisiana, Massachusetts, Minnesota, Missouri, New Mexico, North Carolina, Ohio, and Texas.
10. While the charter school movement originally intended that teachers and other public school educators would test innovative approaches as pilots and apply those approaches that produced the most benefit to public schools, the movement has largely shed the real intent of shaping public schools and now poses charter schools as competitors with traditional public schools for students. There are also individual charter schools or smaller groups of charter schools that are established by entrepreneurs, groups of parents, teachers, and others.
11. "With some notable exceptions, Ohio's charter schools are failing to produce higher academic achievement for the studies and families. By most measures, [they] have performed worse than Ohio's public schools ... being twice as likely as comparable schools in large urban districts to be ranked in Academic Emergency and half as likely to be ranked Effective." (Dingerson 2009, 21)
12. Among the states studied, charter schools in Arizona, Florida, Georgia, Minnesota, Ohio, Pennsylvania, and Texas performed "significantly below" their public school counterparts, while charters in Arkansas, Colorado, Illinois, Indiana, Louisiana, Missouri, and New York City performed significantly better. Charter schools in Washington, D.C., performed at roughly the same level as their public school equivalents. Charters in some states performed better in reading, but worse in math. CREDO noted problems with authorization and accountability across most states, and in Pennsylvania, while overall growth was much lower in both subjects in charter schools, the variation across schools was huge.
13. Dingerson writes that, "Because fixed costs such as buses and heating do not go down, the disproportionate reduction in teaching staff has increased class sizes, increased the numbers of teachers teaching outside their fields of expertise and decreased the elective courses available in Ohio's public schools," p.34. In many cases, costs do decline—some bus routes can be eliminated, or fewer classrooms cleaned, or even a school building shut down, but not nearly enough to replace the lost student funding.
14. "[D]istricts with comparatively few highly educated teachers have relatively high turnover rates for inexperienced teachers; districts with comparatively many highly educated teachers have relatively low turnover rates for their inexperienced teachers," (Holzman 2012, 11).
15. Of course, if teacher quality improves with the replacement of old teachers, those disadvantages may be offset or even gone, depending on the degree of improvement relative to the disruption to the teacher pool (and student body) from the loss of teachers. The problem found in these three cities is that teachers fired were often not replaced by better teachers, which was compounded by the voluntary departure of more experienced and qualified teachers, leading to a teaching force that was less experienced, credentialed, and effective.
16. As discussed earlier, parallel agendas are being promoted by, among others, the groups StudentsFirst, Stand for Children, 50Can, Democrats for Education Reform, and the Republican governors of Florida, Michigan, Ohio, Michigan, New Jersey, and Wisconsin.
17. When public schools closed, union teachers lost jobs while charter and contract schools hired nonunion teachers.

18. More information on TUDA is available at <http://nces.ed.gov/nationsreportcard/about/district.asp>.
19. Note that “new teachers” indicate teachers who have just joined the profession, versus “new teacher hires,” who are new hires to the district, but may be either new teachers, or experienced teachers who have moved into the district from a different one.
20. Levy’s data cites no reason for termination in 2008; categories such as “ineffective” or “licensure” are employed beginning in 2009, the year before IMPACT began.
21. Many of the DCPS data calculations in this paper come from Mary Levy, a school finance litigator and researcher, who has studied DCPS data for over 30 years, primarily as the former director of the Public Education Reform Project of the Washington Lawyers’ Committee for Civil Rights and Urban Affairs, and more recently as a budget consultant to the D.C. city council as well as the district’s CFO and charter schools. Unlike in New York City and Chicago, no major academic institution has developed such data or comparable reports, so this paper relies heavily on Levy’s calculations, with all original sources noted. With respect to teacher turnover data, pre-Rhee means 2001–2007, and post-Rhee/mayoral takeover means 2008–2012. Technically, the 2007–2008 year is a transition year, but for this purpose, it is counted in Mary Levy’s data as post-mayoral takeover.
22. Rhee’s successor, Kaya Henderson, has indicated that, while she intends to continue many of the changes Rhee instituted, she will reduce the amount of weight student test scores are accorded in IMPACT ratings (though the new system was anticipated to result in an increased number of teacher eliminations).
23. “For example, key stakeholders, including D.C. Council members and parent groups, told us they were not given the opportunity to provide input on DCPS’s initial proposals regarding school closures and consolidations, the establishment of schools that spanned prekindergarten to grade 8, or the planning and early implementation of the new staffing model that placed art, music, and physical education teachers at schools and which fundamentally changed the way funding is allocated across DCPS. Lack of stakeholder involvement in such key decisions led stakeholders, including the D.C. Council and parents groups, to voice concerns that DCPS was not operating in a transparent manner or obtaining input from stakeholders with experience relevant to the District’s education system.”
24. An Illinois Facility Report, funded by the Walton Foundation, recommended the closure or turnover to charter managers of 38 D.C. public schools, on the basis of Tier III or Tier IV ratings, as determined by DCPS Comprehensive Assessment System test score results.
25. “The complete contract terms were not disclosed, but D.C. school system procurement records show that Friends of Bedford has received about \$1.2 million in payments from the District since September 2009.”
26. These FTEs include staff in central administration, central instructional support, and central noninstructional support.
27. While data are not available to enable comparisons of those in the office prior to Rhee with those who made at least \$100,000 at the time she left, a scan of the latter compiled by Mary Levy shows that some of the lower-paid central office staff making that much were directors in the Office of the Chief Academic Officer and the Office of Data and Accountability. Some teachers, teacher mentors, principals, and curriculum writers made slightly more. At higher pay levels, specialist/manager and analyst/coordinators were paid in the \$120,000–\$140,000 range, with many directors in the mix. At the upper end of the scale, instructional superintendents and high-ranking deputies earned between \$140,000 and \$170,000. Note that at every pay grade, positions varied—the summary here represents an approximation of the range.
28. While Ansari does not specify, other reports indicate that closed schools disproportionately served students with special education needs (self-contained classroom), English language learners (ELL), homeless, and overage students.

29. According to the report (pp. 2–3), small schools of choice (SSCs) “received start-up funding from philanthropic organizations plus technical assistance and policy support from the New York City Department of Education and intermediary organizations that were often experienced in managing the launching of new schools. ... The creation of SSCs was supported by a consortium of funders, led by the Bill & Melinda Gates Foundation, the Carnegie Corporation of New York, and the Open Society Institute.”
30. Beginning in 2009, results for charter schools are excluded from the TUDA results if they are not included in the school district’s Adequate Yearly Progress (AYP) report to the U.S. Department of Education (NCES 2011c).
31. It is not possible to determine whether these declines were within the range that could be expected, given the changing demographics of the receiving schools, or whether they were larger than expected, and partly the negative result of student mobility and upheaval. As noted, however, any overall assessment of the pros and cons of the new small schools must account for impacts on the system overall.
32. The New York State Charter Schools Act of 1998 authorized the establishment of charter schools in New York, with the first twelve operating in New York City in 2000–2001. As of 2009, 78 charter schools were operating in New York City, and 26 more were scheduled to open in 2009 or 2010 (Hoxby, Murarka, and Kang 2009).
33. Such contributions are not easily calculable, so the researchers used a combination of IRS records and estimates of how donations were distributed to estimate the totals they produce.
34. Out of concern that strong local control over principal hiring could lead to some very bad hires, Duncan also centralized this process for some of the worst-performing schools, creating a principal pool from which Local School Councils (LSCs) were forced to choose. The impact, if any, of this change is unclear.
35. Selective-enrollment in Chicago is based entirely on standardized test scores, employing a combination of citywide and neighborhood socioeconomic status “tier” systems. As the CPS website explains, “For Selective Enrollment Elementary and High Schools, the first 30% of available seats at each school are filled strictly according to the student scores—also referred to as the “rank” score—the seats are filled by the top-scoring students. The remaining seats are equally distributed among the four tiers and are filled by the highest-scoring students in each tier” (Chicago Public Schools 2013). Nonselective magnet schools prioritize siblings of existing students and students who live closest to the school, with the remainder of seats filled by lottery, distributed evenly across the four tiers.
36. Between 2001 and 2008, when Duncan was CEO of CPS, the Academy for Urban School Leadership instituted its comprehensive turnaround model in 11 schools. The academy has since expanded its network to 14 more schools, for a total of 25 as of 2012 (AUSL 2012).
37. As in other assertions regarding low-performing or high-performing schools, this CCSR report uses test scores to measure school performance, with the “strongest” receiving schools those in the top quartile on test scores, and the weakest ones those in the bottom quartile.
38. The Academy for Urban School Leadership (AUSL) website’s FAQ (www.ausl-chicago.org/about-faq.html) states that “[s]teady, positive improvements in achievement and attendance results are a hallmark of all 19 of the CPS schools managed by AUSL.”
39. It is possible that new, higher-scoring students also affected scores in New York City and Chicago. Given the size and relative diversity of those two cities, however, such determinations are not possible here.
40. A corollary to this is the need for policymakers to get tougher on closing down charters that are doing worse than their public school counterparts, failing to live up to commitments, or, in some cases, engaging in fraudulent behavior. As Dingerson

documents, such problems are common in Ohio but often do not result in state action to protect parents and students. And as the Fordham Foundation, a major proponent of increased access to charter schools, reported in *Are Bad Schools Immortal?* in 2010, closures of under-performing and otherwise unacceptable charters are rare, notwithstanding charter proponents' and authorizers' promises to revoke failing charters' licenses. (Stuit 2010).

41. New York's 2010 rate is not available, so the 72 percent is for 2009.
42. Such strategies include but are not limited to quality early childhood education, such as pre-K and nurse home-visiting programs; school-based health clinics; breakfast, afterschool, and summer nutrition support programs; parent engagement and education sessions; reengagement programs; and quality afterschool and summer enrichment programs.
43. Ursula Helminski is vice president of external affairs for the Afterschool Alliance. The result of the coordinated campaign was a budget deal that restored \$150 million for child care and after-school programs, including new funds for several initiatives. However, of the \$150 million, only \$41 million will be "baselined." The majority of the funds, \$109 million, are being provided using council "discretionary" funding, which can only be allocated for a one-year period. As a result, these programs likely face a similar last-minute fight for survival again in 2013.
44. While the term "proficiency" is often used to indicate a score cutoff or standard that the majority of students are expected to reach, it was never intended that way by NAEP's governing board.
45. The same pattern is seen with respect to the gap between white and Hispanic students; between 2003 and 2007 the gap narrowed substantially, then widened again through 2011.
46. Among high school students, the white-Hispanic gap also grew slightly, rather than narrowing.
47. The only exception was special education students, among whom those who had participated in Pre-K scored lower in both math and reading than their non-participant special education peers.
48. "Averaged freshman graduation rate (AFGR) is an estimate of the percentage of an entering freshman class graduating in four years. For 2007–08, it equals the total number of high school diploma recipients in 2007–08 divided by the average membership of the 8th-grade class in 2003–04, the 9th-grade class in 2004–05, and the 10th-grade class in 2005–06. Ungraded students were allocated to individual grades proportional to each state's enrollment in those grades." (NCES 2010b, Table A12, FN4)
49. The source for data on local school districts is the U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), *Local Education Agency Universe Survey, Dropout and Completion Public-Use Data File, 2002-03 AND 2007–08, Version 1a*. The source for data on states is the U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "The Averaged Freshman Graduation Rate for Public High Schools From the Common Core of Data, School Years 2001–02 and 2002–03" and U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "Public School Graduates and Dropouts From the Common Core of Data: School Year 2007–08 and 2008–09," Table 1.
50. The author explains why and how he converted what had begun as Standard Deviation differences into percentile change differences: "Given the mayor's penchant for reducing complex phenomena to a single number ... I have summarized the shrinkage in the achievement gap on the NAEP and New York State assessments as the percentage reduction in the gap. (For the technically-minded, this involved calculating group differences in citywide standard-deviation units, weighted by the size of the four racial/ethnic groups, for each grade and subject area, and then averaging those group differences, in both 2011 and 2003. The ratio of the 2011 group difference to the 2003 group difference indicates the extent of the change in the achievement gap over that eight-year period.)

“So here it is: Looking across ELA and math scores on state exams for New York City students in grades three through eight in 2003, the achievement gap separating black and Latino students from white and Asian students was .74 of a standard deviation. In 2011, the achievement gap was .73 of a standard deviation. This represents a 1 percent reduction in the magnitude of the achievement gap. The careful reader will note that the mayor has thus overstated the cut in the achievement gap by a factor of 50.” (Pallas 2012b)

51. Pallas also noted that, when NAEP scores were used to assess the validity of the same claim, the gap grew by 3 percent.

52. Assessments of gains (or losses) in test scores could also take into account the starting point for a city or district as, of course, those with higher starting scores have less room to grow. It is not possible, in this context, to compare starting NAEP scores of every large city, or even a smaller subset of them, and such an analysis also requires more than simple comparison to be valid, since many other factors should be taken into account. For the purposes of this report, the goal is to compare promised, and asserted, gains, against those actually realized, so only these basic comparisons are employed.

53. Moreover, while scores continued to climb between 2009 and 2011 in the nation as a whole, and in the large urban school districts, scores in New York City dropped slightly, though the decrease was not statistically significant.

54. These comparisons should be viewed within the context of the two sets of schools’ student bodies. On average, as described earlier, district schools serve roughly the same proportion of students who qualify for meal subsidies, though more of them are poor (eligible for free lunch), versus low-income (eligible for reduced-price lunch). District schools also serve more special needs students, and likely substantially more of those with very high needs. Charter schools also serve many more African American students, and fewer Hispanic, White, and Asian students than their regular district school counterparts (Merriman et al. 2012).

55. “Developed by the state two years ago, this indicator counts college ready students as those who scored at least 75 on the English Regents and 80 on the math, or got at least 480 on their math and reading SATs” (McAdoo 2012).

56. Luppescu et al. (2011, 11.) explain: “The trends in student achievement displayed in this report frequently do not match the publicly reported statistics. This does not mean that the statistics that are reported publicly are wrong. However, they are often calculated in ways that are not comparable across the years. Decisions about how to produce indicators of student performance change frequently in response to policies at the local, state, and federal levels. Often changes are made in an attempt to produce more accurate indicators, but these changes make the indicators non-comparable to those produced in the past. In this report, we make our own calculations from student-level data, so that student achievement can be compared in a fair way over time. ... There are five general issues that make it difficult to create fair comparisons across time in students’ test scores:

- 1) Changes in tests, standards, scoring, and test administration make scores non-comparable.
- 2) The most commonly used metric—the percent meeting standards—is imprecise and can be misleading.
- 3) The promotion policy instituted in Era 2 concentrates low-scoring students in certain grades and keeps the lowest-scoring students’ scores in district averages for extra years.
- 4) The proportion of CPS students whose test scores were included in the publicly reported statistics has changed over time with various policies.
- 5) The types of students entering Chicago schools have changed over time, and these demographic changes can affect district achievement levels.

This chapter details the methods CCSR researchers used to address each of the five issues outlined above in order to make fair comparisons over time. The complexity of the methodology underscores how difficult it is to gauge improvements in schools and

districts when the statistics that are reported are affected by numerous decisions of policymakers, practitioners, and the makers of assessments.” Note that, notwithstanding the rigorous methodology employed, the report still compares the results against NAEP scores.

57. As the Consortium researchers document, Duncan’s tenure represented a third consecutive wave of (different) CPS reform agendas over the past twenty years. The report not only documents trends between 2002 and 2009, when Duncan was CEO (which it distinguishes as one of “innovation” for the many new schools opened), but compares those to trends between 1990–1995 (which it labels as the era of “decentralization” and 1996–2001 (when “accountability” emerged under Duncan’s predecessor, Paul Vallas).
58. These studies include a 2009 CREDO study limited to CPS students in grades 3-8 (CREDO 2009b); a study conducted by the RAND Corporation in 2008 (Booker et al. 2009); and Brown and Gutstein (2009).
59. The authors acknowledge the serious flaws with this methodology, but reason that, if CPS is going to use it for such serious decisions as school closure, it seems fair to employ it to assess the claim that charters are outperforming comparable public schools serving the highest-needs students.
60. This discrepancy could reflect the unreliability of standardized test scores as a measure of performance, in which case, more weight should be placed on improvement in the graduation rate. However, since high school graduation rates can be manipulated, and since NAEP scores show that CPS students who entered charter schools behind tended to gain less ground than their higher-starting counterparts, putting more stock in improved graduation rates is also risky.
61. The RAND report noted that the existence of charter high schools that spanned primary and secondary grades “creates an ambiguity in interpreting results: Any measured effects might be attributable to charter status or might be attributable to the unconventional grade configuration that eliminates the transition between middle school and HS. Transitions to new schools often have negative effects; charter schools might promote positive outcomes simply by eliminating a transition and adopting 6–12 or K–12 grade configurations. We cannot be sure about the causal *mechanism* producing any effects we measure.” (Booker et al. 2009, 34)
62. Prior to helping Duncan develop CPS’s college- and career-readiness efforts, Darnidier was the executive director of the Steans Family Foundation and also led the Cabrini-Green youth program, CYCLE.
63. “[I]t is unclear what the current CPS administration plans to do to keep the college enrollment numbers rising. The centralized post-secondary office, which included specialists who worked with school teams, has now been whittled down to just a few people. Some network offices have post-secondary specialists, but still [Liz] Monge-Pacheco [the post-secondary coach for the Network for College Success] says people are wondering who will help organize college tours and help individual schools navigate the process. ‘There is a loss of resources this year,’ she says. ‘CPS has put forth no plan on how they will support post-secondary enrollment.’ As part of its effort to close a huge budget gap, the district office told schools that if they want to keep their college coach, they will have to cover the cost completely with discretionary funds. Previously, central office chipped in.” (Karp 2012a)
64. All dollars are adjusted for inflation using the CPI-U-RS and 2008–2009 dollars.
65. Federal grants include the following: Reading is Fundamental (\$24.8 million over 2008–2010), Save the Children Juvenile Delinquency Prevention program (about \$0.5 million over 2008–2010), and Teach for America (\$2.6 million in 2008) (LegiStorm 2012).
66. Levy noted that, although the ROTC mandate is unfunded, schools are forbidden to stop providing the service, so they have no way to recoup that money.

67. “Given that the successful implementation of the tentative agreement is dependent on effective leadership, the third party funders reserve the right to reconsider their support for this initiative if there is a material change in DCPS’ leadership.”
68. One exception is Los Angeles, which increased spending at a rate even greater than that of New York City, though it also started out at a much lower place—roughly \$8,000 per student.
69. Regarding private funding for market-based “portfolio district initiatives,” DePaul University professor Kenneth Saltman (2010) notes, “The amounts available have been significant, by some accounts as high as \$200 million in New York City, half that amount in Chicago, and over \$50 million in New Orleans.”
70. It is possible that these funds fall into discretionary funding. If so, they would constitute a more significant proportion than they do of the budget overall.
71. The Chicago Public Education Fund’s collective investment so far of \$50 million also resembles New York City’s private contributions: The influence of the donors is disproportionate to their share of the school budget (Chicago Public Education Fund 2012b). The fund emphasizes that this “small group of highly engaged and committed individuals... put a lot on the line, and expect a lot in return” and that the board has “Great Power, Great Responsibility” (Chicago Public Education Fund 2012a).
72. Moreover, the article notes, while the purported reason for closing the school is underenrollment in that specific building, the charter and magnet schools that replace them tend to house even fewer students, which makes the building even less cost-effective by that measure. “[Executive director of the 21st Century Fund in Washington, DC Mary] Filardo points out that charter operators often want to keep their student population small. That is true in Chicago. On average, charter elementary schools have about 100 fewer students than neighborhood schools, and 21 of them have fewer than 350 students, according to an analysis of CPS enrollment data.”
73. These statistics are based on studies of students who attended the Chicago Child-Parent Centers, compared with similar children who did not. The odds of special education placement among program participants were reduced from 24.6 percent to 14.4 percent, and the odds of graduating from high school increased from 38.5 percent to 49.7 percent. It is noteworthy that the students remained at very high risk, but the risk levels were significantly reduced; i.e., the program did not help them attain anywhere near the status of their higher-income peers.
74. This is based on studies of LA’s Best, a highly effective summer and afterschool program in Los Angeles. Test score increases from this program and others tend to be substantially lower than metrics like high school graduation and criminal activity, which reflect gains across a range of domains and skillsets, rather than purely cognitive ones.

References

Academy for Urban School Leadership (AUSL). 2012. “Find a School.” www.ausl-chicago.org/schools.html.

Ahmed, Azam. 2010. “Judge Sides with Union in Chicago Teacher Layoffs.” *Chicago Tribune*, October 4. http://articles.chicagotribune.com/2010-10-04/news/ct-met-cps-teacher-lawsuit-1005-20101004_1_teacher-layoffs-union-and-district-officials-tenured-teachers

Alexander, Karl L., Doris R. Entwisle, and Linda Steffel Olson. 2007. “Lasting Consequences of the Summer Learning Gap.” *American Sociological Review*. Vol. 72, 167–180.

Allensworth, Elaine, Stephen Ponisciak, and Christopher Mazzeo. 2009. *The Schools Teachers Leave: Teacher Mobility in Chicago Public Schools*. Consortium on Chicago School Research at the University of Chicago Urban Education Institute. http://ccsr.uchicago.edu/sites/default/files/publications/CCSR_Teacher_Mobility.pdf

- American University. 2012. "Education Now: Cities at the Forefront of Reform" [American University forum]. March 2. <http://www.c-spanarchives.org/program/EducationReform26> [video transcript]
- Ansari, Zakiyah. 2012. "Save Students by Ending Bloomberg's Destructive School Closings Policy." *Black Star News*, May 15. <http://www.blackstarnews.com/news/135/ARTICLE/8216/2012-05-15.html>
- Ashby, Cornelia. 2009. "District of Columbia Public Schools: Implementation and Sustainability of Reform Efforts Could Benefit From Enhanced Planning." Testimony before the Subcommittee on Oversight of Government Management, the Federal Workforce, and the District of Columbia, Committee on Homeland Security and Governmental Affairs, U.S. Senate. July 23.
- Baker, Al. 2012. "Arbitrator Halts City's Plan to Overhaul 24 Schools." *New York Times*, June 29. <http://www.nytimes.com/schoolbook/2012/06/29/city-loses-arbitration-on-staffing-for-24-turnaround-schools/>
- Baker, Bruce D., and Richard Ferris. 2011. *Adding Up the Spending: Fiscal Disparities and Philanthropy Among New York City Charter Schools*. National Education Policy Center at the University of Colorado, Boulder. <http://nepc.colorado.edu/publication/NYC-charter-disparities>
- Baker, Eva L., Paul E. Barton, Linda Darling-Hammond, Edward Haertel, Helen F. Ladd, Robert L. Linn, Diane Ravitch, Richard Rothstein, Richard J. Shavelson, and Lorrie A. Shepard. 2010. *Problems with the Use of Student Test Scores to Evaluate Teachers*. Economic Policy Institute, Briefing Paper No. 278. <http://www.epi.org/publication/bp278/>
- Banchero, Stephanie, Joe Germuska, and Darnell Little. 2010. "Daley School Plan Fails to Make Grade; Renaissance 2010 Officials Defend Efforts to Upgrade Education for Chicago Students Over Last 6 Years." *Chicago Tribune*, January 17. http://articles.chicagotribune.com/2010-01-17/news/1001160276_1_charter-schools-chicago-reform-urban-education
- Barnes, Gary, Edward Crowe, and Benjamine Schaefer. 2007. *The Cost of Teacher Turnover in Five School Districts: A Pilot Study*. National Commission on Teaching and America's Future.
- Barry, Marion. 2012. "Memorandum to the Council Re: DCPS Proposed Budget FY 2013" [memo circulated among members of the D.C. City Council by Ward 8 representative Marion Barry].
- Barton, Paul E., and Richard J. Coley. 2007. *The Family: America's Smallest School*. Policy Information Report, Princeton: Educational Testing Service.
- Berliner, David C. 2009. *Poverty and Potential: Out-of-School Factors and School Success*. Education and the Public Interest Center at Arizona State University. <http://nepc.colorado.edu/files/PB-Berliner-NON-SCHOOL.pdf>
- Bloom, Howard S., and Rebecca Unterman. 2012. *Sustained Positive Effects on Graduation Rates Produced by New York City's Small Public High Schools of Choice*. MDRC Policy Brief. <http://www.mdrc.org/sustained-positive-effects-graduation-rates-produced-new-york-city%E2%80%99s-small-public-high-schools>
- Board on Testing and Assessment, Division of Behavioral and Social Sciences and Education, National Research Council, National Academy of Sciences (BOTA). 2009. *Letter Report to the U.S. Department of Education on the Race to the Top Fund*. http://www.nap.edu/catalog.php?record_id=12780
- Booker, Kevin, Brian Gill, Ron Zimmer, and Tim R. Sass. 2009. *Achievement and Attainment in Chicago Charter Schools*. RAND Corporation. http://www.rand.org/pubs/technical_reports/TR585-1.html

- Brown, Emma. 2011. "New Rating System Will Put More D.C. Teachers at Risk." *Washington Post*, August 2. http://www.washingtonpost.com/local/education/new-rating-system-will-put-more-dc-teachers-at-risk/2012/08/02/gJQAKNyySX_story.html.
- Brown, Emma. 2012a. "Officials Cut Ribbon at Anacostia High." *Washington Post*, September 15. http://www.washingtonpost.com/blogs/dc-schools-insider/post/officials-cut-ribbon-at-anacostia-high/2012/09/15/645f08a2-feb9-11e1-8adc-499661afe377_blog.html
- Brown, Emma. 2012b. "Parents Struggle With 'Wait-list Shuffle' in D.C. Schools." *Washington Post*, September 9. http://www.washingtonpost.com/local/education/dc-school-parents-struggle-with-wait-list-shuffle/2012/09/09/6b10eb26-f2f1-11e1-adc6-87dfa8eff430_story.html
- Brown, Liz, and Eric Gutstein. 2009. *The Charter Difference: A Comparison of Chicago Charter and Neighborhood High Schools*. Collaborative for Equity and Justice in Education, University of Illinois-Chicago. <http://www.uic.edu/educ/ceje/articles/CharterDifference.pdf>
- Brownfield, Mike. 2012. "An Emergency Financial Manager Success Story." Office of the Governor of Michigan, November 15. <http://www.michigan.gov/snyder/0,4668,7-277-60279-289946-,00.html>
- Bryk, Anthony S., Penny Bender Sebring, Elaine Allensworth, Stuart Luppescu, and John Q. Easton. 2010. *Organizing Schools for Improvement: Lessons from Chicago*. Consortium on Chicago School Research at the University of Chicago Urban Education Institute. Chicago: University of Chicago Press.
- Bushaw, William J., and Shane J. Lopez. 2012. "Public Education in the United States: A Nation Divided." *Kappan Magazine*. Vol. 94, no. 1, 8–25. <http://www.kappanmagazine.org/content/94/1/8.abstract>
- Caref, Carol, Sarah Hains, Kurt Hilgendorf, Pavlyn Jankov, and Kevin Russell. 2012. *The Black and White of Education in Chicago's Public Schools: Class, Charters & Chaos: A Hard Look at Privatization Schemes Masquerading as Education Policy*. Chicago Teachers Union. <http://www.ctunet.com/blog/the-black-and-white-of-education-in-chicagos-public-schools>
- Carroll, Thomas G. 2011. *Policy Brief: The High Cost of Teacher Turnover*. National Commission on Teaching and America's Future. <http://chalkboardproject.org/wp-content/uploads/2011/01/ment-13.pdf>
- Carroll, Thomas G. and Elizabeth Foster. 2010. *Who Will Teach? Experience Matters*. National Commission on Teaching and America's Future.
- Cavanagh, Sean, and Stephen Sawchuk. 2012. "Outside Advocacy Groups Target Local School Board Elections." *Education Week*, May 21. <http://www.edweek.org/ew/articles/2012/05/23/32adv-local.h31.html>
- Center for Research on Education Outcomes (CREDO). 2009a. *Charter School Performance in Ohio*. http://credo.stanford.edu/reports/OH_CHARTER_SCHOOL_REPORT_CREDO_2009.pdf
- Center for Research on Education Outcomes (CREDO). 2009b. *Multiple Choice: Charter School Performance in Sixteen States*. http://credo.stanford.edu/reports/MULTIPLE_CHOICE_CREDO.pdf
- Center for Research on Education Outcomes (CREDO). 2010. *Charter School Performance in New York City*. <http://www.nyccharterschools.org/resources/credo-study-charter-school-performance-new-york-city-2010>
- Chandler, Michael Alison, and Emma Brown. 2012. "D.C. Public and Charter Schools Show Gains in Math Scores." *Washington Post*, July 26. http://articles.washingtonpost.com/2012-07-26/local/35486138_1_charter-schools-proficiency-rates-national-tests

Chicago Public Education Fund. 2012a. "Board of Directors." <http://www.thefundchicago.org/index.php?tray=topic&tid=top13&cid=134>

Chicago Public Education Fund. 2012b. "Investment Fund and Goals." <http://www.thefundchicago.org/index.php?tray=topic&tid=top11&cid=128>)

Chicago Public Education Fund. 2012c. "Who is Making a Difference at the Fund" and "What the Fund is Achieving." Chicago Public Education Fund website. <http://www.thefundchicago.org>

Chicago Public Schools. 2013. "Looking for Your Tier? Need an Explanation of the Process?" CPS Office of Access and Enrollment. http://cpsoae.org/apps/news/show_news.jsp?REC_ID=184188&id=0.

Civic Committee of the Commercial Club of Chicago. 2009. *Still Left Behind: Student Learning in Chicago's Public Schools*. <http://civiccommittee.org/initiatives/education/Still%20Left%20Behind%20v2.pdf>

Cramer, Philissa. 2011. "Grad Rate Gains at Some Set-to-Close Schools Outpace City's." *Gotham Schools*, June 15. <http://gothamschools.org/2011/06/15/grad-rate-gains-at-some-set-to-close-schools-outpace-citys/>

Cromidas, Rachel. 2012. "City Council's Hearing on Co-locations Airs Persistent Concerns." *Gotham Schools*, April 19. <http://gothamschools.org/2012/04/19/city-councils-hearing-on-co-locations-air-persistent-concerns/>

Cromidas, Rachel and Geoff Decker. 2012. "State names 123 city schools to improve or close by 2015." *Gotham Schools*, August 30. <http://gothamschools.org/2012/08/30/state-names-city-schools-to-improve-or-close-by-15/>

DC Public Education Fund. 2010. Letter to Michelle Rhee affirming commitment of funds to support performance-based teacher compensation, March 30. http://www.washingtonpost.com/wp-srv/education/DCPEF_Commitment_Letters_04272010.pdf.

DeAngelis, Karen J., Bradford R. White, and Jennifer B. Presley. 2010. "The Changing Distribution of Teacher Qualifications Across Schools: A Statewide Perspective Post-NCLB." *Education Policy Analysis Archives*. Vol. 18, no. 28. <http://epaa.asu.edu/ojs/article/view/722>

DeBonis, Mike. 2007. "The Office. Plenty of Bosses Have Taken on the DCPS Headquarters and Failed. Will Michelle Rhee Be Any Different?" *Washington City Paper*. November 15. <http://www.washingtoncitypaper.com/articles/34099/the-office/>.

De la Torre, Marisa, and Julia Gwynne. 2009. *When Schools Close: Effects on Displaced Students in Chicago Public Schools*. Consortium on Chicago School Research at the University of Chicago Urban Education Institute. <http://ccsr.uchicago.edu/publications/when-schools-close-effects-displaced-students-chicago-public-schools>

De la Torre, Marisa, Elaine Allensworth, Sanja Jagesic, James Sebastian, and Michael Salmonowicz. 2012. *Turning Around Low-Performing Schools in Chicago*. Consortium on Chicago School Research at the University of Chicago Urban Education Institute. <http://ccsr.uchicago.edu/sites/default/files/publications/12CCSRTurnAround-3.pdf>

Dillon, Sam. 2009. "Report Questions Duncan's Policy of Closing Failing Schools." *New York Times*, October 29. <http://www.nytimes.com/2009/10/29/education/29schools.html>

Dingerson, Leigh. 2008. *Reclaiming the Education Charter: Ohio's Experiment with Charter Schooling*. Education Voters Institute, Forum for Education and Democracy.

District of Columbia Public Schools. 2010. *DCPS on the Move: Three-Year Growth and Trends*. <http://dcps.dc.gov/DCPS/Files/downloads/ABOUT%20DCPS/Press/DCPS-Press-Release-DCPS-On-the-Move.pdf>

- District of Columbia Public Schools. 2011. IMPACT – The DCPS Effectiveness Assessment System for School-Based Personnel. <http://www.dc.gov/DCPS/impact>.
- Dixon, Mark. 2012. *Public Education Finances: 2010*. Governments Division Reports, U.S. Census Bureau. U.S. Department of Commerce, Economics and Statistics Administration.
- Domanico, Ray, and Yolanda Smith. 2011. “Charter Schools Housed in the City’s School Buildings Get More Public Funding per Student than Traditional Public Schools.” *IBO Web Blog*, February 15. <http://ibo.nyc.ny.us/cgi-park/?p=272>
- Engberg, John, Brian Gill, Gema Zamarro, and Ron Zimmer. 2011. “Closing Schools in a Shrinking District: Do Student Outcomes Depend on Which Schools are Closed?” *Journal of Urban Economics*. Vol. 71, no. 2. <http://www.sciencedirect.com/science/article/pii/S009411901100060X>
- Epstein, Marc. 2012. “Putting New York’s School Reform in Perspective.” *Huffington Post*, March 12. http://www.huffingtonpost.com/marc-epstein/new-york-schools_b_1319833.html.
- FairTest. 2012. “How to Stop Cheating on Standardized Tests.” *Washington Post’s The Answer Sheet* (blog), February 17. http://www.washingtonpost.com/blogs/answer-sheet/post/how-to-stop-cheating-on-standardized-tests/2012/02/16/gIQAPF0nIR_blog.html
- Fleisher, Lisa. 2012a. “Fewer City Teachers Get School Tenure.” *Wall Street Journal*, August 17. <http://online.wsj.com/article/SB10000872396390443324404577595670912547892.html>
- Fleisher, Lisa. 2012b. “Mayor’s Work on Schools Gets a Backer.” *Wall Street Journal*, April 4. <http://online.wsj.com/article/SB10001424052702303299604577324193681145570.html>
- Fund for Public Schools. 2010. *Annual Report*. <http://fps2010annualreport.org/main.html#/donors>.
- Fund for Public Schools. 2012. “About Us.” Fund for Public Schools website. <https://www.fundforpublicschools.org/about-us>
- Galinsky, Ellen. 2006. *The Economic Benefits of High-Quality Early Childhood Programs: What Makes the Difference*. The Committee for Economic Development.
- Gartner, Lisa. 2012a. “Audit: Closing Schools Cost Quadruple the Reported Price Tag.” *Washington Examiner*, September 9. <http://washingtonexaminer.com/audit-closing-schools-cost-quadruple-the-reported-price-tag/article/2507444>
- Gartner, Lisa. 2012b. “D.C. to weaken link between test scores, teacher ratings.” *Washington Examiner*, August 2. <http://washingtonexaminer.com/d.c.-to-weaken-link-between-test-scores-teacher-ratings/article/2503893>.
- Glazerman, Steven, and Allison Seifullah. 2010. *An Evaluation of the Teacher Advancement Program (TAP) in Chicago: Year Two Impact Report*. Mathematic Policy Research, Inc. <http://ideas.repec.org/p/mpr/mpres/6649.html>
- Glazerman, Steven, and Allison Seifullah. 2012. *An Evaluation of the Teacher Advancement Program (TAP) in Chicago After Four Years: Final Report*. Mathematica Policy Research, Inc. http://www.mathematica-mpr.com/publications/pdfs/education/TAP_year4_impacts.pdf
- Goldstein, Dana. 2010. “Early Thoughts on Joel Klein, Cathie Black and Education Reform in New York.” *Washington Post*, November 9. http://voices.washingtonpost.com/ezra-klein/2010/11/early_thoughts_on_joel_klein_c.html.
- Haimson, Leonie. 2012. *NYC Achievement Gains Compared to Other Large Cities Since 2003*. Class Size Matters. <http://www.classsizematters.org/nyc-second-to-last-among-cities-in-student-progress-on-the-naeps-since-2003/>

Hanushek, Eric A., John F. Kain, and Steven G. Rivkin. 2004. "Disruption Versus Tiebout Improvement: The Costs and Benefits of Switching Schools." *Journal of Public Economics*. Vol. 88, no. 9–10. <http://www.nber.org/papers/w8479>

Haycock, Kati. 2006. *Teaching Inequality: How Poor and Minority Students are Shortchanged on Teacher Quality*. The Education Trust. <http://www.edtrust.org/sites/edtrust.org/files/publications/files/TQReportJune2006.pdf>

Heckman, Paul and Carla Sanger. 2013. "How Quality Afterschool Programs Help Motivate and Engage More Young People in Learning, Schools, and Life." In *Expanding Minds and Opportunities: The Power of Afterschool and Summer Learning for Student Success*. Terry K. Peterson, Executive Editor. Collaborative Communications Group. Washington, D.C.

Hemphill, Clara, and Kim Nauer. 2009. *The New Marketplace: How Small-School Reforms and School Choice Have Reshaped New York City's High Schools*. The New School for Management and Urban Policy. http://www.newschool.edu/milano/nyc affairs/publications_schools_thenewmarketplace.aspx

Hill, Kerry G. 2010. "Fraynd Focuses on Turning Around Chicago Schools." Learning Connections, University of Wisconsin Madison School of Education. <http://campusconnections.education.wisc.edu/post/PEOPLE-Don-Fraynd.aspx>

Holzman, Michael. 2012. *A Rotting Apple: Education Redlining in New York City*. Schott Foundation for Public Education. <http://www.schottfoundation.org/publications-reports/education-redlining>

Hood, Joel, and Noreen S. Ahmed-Ullah. 2011. "University of Chicago Report Says CPS Reforms Have Failed Many Students." *Chicago Tribune*, September 30. http://articles.chicagotribune.com/2011-09-30/news/ct-met-cps-reforms-0930-20110930_1_terry-mazany-cps-jean-claude-brizard.

Hoxby, Caroline M., Sonali Murarka, and Jenny Kang. 2009. *How New York City's Charter Schools Affect Achievement*. New York City Charter School Achievement Project. http://users.nber.org/~schools/charterschoolseval/how_NYC_charter_schools_affect_achievement_sept2009.pdf

Huffington Post. 2011. "D.C. Public Schools Revises Controversial Teacher Evaluation System Implemented By Rhee." September 6. http://www.huffingtonpost.com/2011/09/06/dc-public-schools-revises_n_951178.html.

Huffington Post. 2012. "Bloomberg's New Schools Failed to Improve Academic Standing of Thousands of New York City Students, Analysis Shows." July 23. http://www.huffingtonpost.com/2012/07/23/bloombergs-new-schools-ha_n_1695536.html

Hunt, James B., and Thomas G. Carroll. 2003. *No Dream Denied: A Pledge to America's Children*. National Commission on Teaching and America's Future. http://nctaf.org/wp-content/uploads/no-dream-denied_full-report.pdf

Independent Budget Office (IBO) of the City of New York. 2012a. *Department of Education Funding, Fiscal Years 1990–2016*. Updated Nov. 2012. http://www.ibo.nyc.ny.us/iboreports/Funding_Trends1990_20048.pdf

Independent Budget Office (IBO) of the City of New York. 2012b. *Statistical Profile of Schools on DOE's 2012 Closure List*. <http://www.ibo.nyc.ny.us/iboreports/2012schoolclosing.pdf>

Isaacs, Julia B. 2012. *Starting School at a Disadvantage: The School Readiness of Poor Children*. Brookings Institution Center on Children and Families. <http://www.brookings.edu/research/papers/2012/03/19-school-disadvantage-isaacs>

Jackson, Frank G. 2012. "Cleveland's Plan for Transforming Schools: Reinventing Public Education in Our City and Serving as a Model of Innovation for the State of Ohio." Office of the Mayor of the City of Cleveland, February 2. <http://www.gcpartnership.com/Talent/Primary-and-Secondary-Education/-/media/Files/Talent/ClevelandPlanFinalExecutiveSummary%20FINAL.ashx>

- Jennings, Jennifer L., and Leonie Haimson. 2009. *High School Discharges Revisited: Trends in New York City's Discharge Rates, 2000-2007*. Columbia University and Class Size Matters. http://www.classsizematters.org/wp-content/uploads/2011/04/High_School_Discharge_Report_FINAL.pdf
- Kahlenberg, Richard D. 2012. "Introduction: Socioeconomic School Integration." In Richard D. Kahlenberg, ed., *The Future of School Integration: Socioeconomic Diversity as an Education Reform Strategy*. New York: The Century Foundation Press.
- Karp, Sarah. 2011. "Consortium Study Says Little Improvement in Elementary Students over Two Decades." *Chicago Catalyst*, September 30. <http://www.catalyst-chicago.org/notebook/2011/09/30/consortium-study-says-little-improvement-in-elementary-students-over-two-decades>
- Karp, Sarah. 2012a. "Jones and Urban Prep Shine as Citywide College Enrollment Rises." *Catalyst Chicago*. August 21. <http://www.catalyst-chicago.org/news/2012/08/21/20362/jones-and-urban-prep-shine-citywide-college-enrollment-rises>
- Karp, Sarah. 2012b. "Minimal Cost Savings for Closing Schools: Analysis." *Catalyst Chicago*. October 31. <http://www.catalyst-chicago.org/notebook/2012/10/31/20573/minimal-cost-savings-closing-schools-analysis>
- Kirshner, Ben, Matthew Gaertner, and Kristen Pozzoboni. 2010. "Tracing Transitions: The Effect of High School Closure on Displaced Students." *Educational Evaluation and Policy Analysis*. Vol. 32, no. 3, 407–429.
- Klein, Joel. 2009. "Urban Schools Need Better Teachers, Not Excuses, to Close the Education Gap." *U.S. News and World Report*, May 4. <http://www.usnews.com/opinion/articles/2009/05/04/urban-schools-need-better-teachers-not-excuses-to-close-the-education-gap>
- Ladd, Helen F. 2011. *Education and Poverty: Confronting the Evidence*. N.C.: Duke Sanford School of Public Policy. <http://sanford.duke.edu/research/papers/SAN11-01.pdf>
- LegiStorm. 2012. Paid subscription database. <http://www.legistorm.com/>
- Lerner, Mark. 2011. "Concerns About the Achievement Gap with This Year's D.C. CAS Results." *Washington Examiner*, August 8. <http://www.examiner.com/article/concerns-about-the-achievement-gap-with-this-year-s-d-c-cas-results>
- Levy, Mary. 2011a. "A Back-to-School Special: Mary Levy Discusses DCPS's de Facto Segregation, Lack of Transparency, High Turnover and More." *FightBack Radio*, September 20. <http://thefightback.org/2011/09/a-back-to-school-special-mary-levy-discusses-dcps-de-facto-segregation-lack-of-transparency-high-turnover-and-more/>
- Levy, Mary. 2011b. "School 'Reform' in DC: Is the Problem Choice or What Compels Families to Choose?" *All Things Education* (blog), December 5. <http://allthingsedu.blogspot.com/2011/12/school-reform-in-dc-is-problem-choice.html>
- Levy, Mary. 2012a. "Audited Enrollment in DC Public Schools and Charter Schools, 1998–2011" [unpublished Excel files provided by D.C. budget consultant Mary Levy, based on annual official audits, excluding tuition students].
- Levy, Mary. 2012b. "DCPS Full-Time Equivalent Central Office Employees, 1981–2012" [unpublished Excel files provided by D.C. budget consultant Mary Levy, derived from D.C. Public Schools annual Schedule As and payroll reports].
- Levy, Mary. 2012c. "DCPS NAEP, 2003–2011" [unpublished Excel files provided by D.C. budget consultant Mary Levy, based on data extracted from National Assessment of Educational Progress].

Levy, Mary. 2012d. “DCPS Principals, 1997–2012” [unpublished Excel files provided by D.C. budget consultant Mary Levy, based on data from D.C. Public Schools annual school directories, supplemented by Washington Post reports on the occasional mid-year change].

Levy, Mary. 2012e. “Hiring Before and After Fall 2009 RIF” [unpublished Excel files provided by D.C. budget consultant Mary Levy, based on D.C. employee lists posted on the D.C. government Freedom of Information Act website].

Levy, Mary. 2012f. “Individual School Turnover” [unpublished Excel files provided by D.C. budget consultant Mary Levy].

Levy, Mary. 2012g. “Teacher Retention, 2001–2012” [unpublished Excel files provided by D.C. budget consultant Mary Levy, assembled from D.C. Public Schools employee lists provided to the D.C. Council during the budget process].

Levy, Mary. 2012h. “Teacher Terminations, 2009–2012” [unpublished Excel files provided by D.C. budget consultant Mary Levy, based on data from Ashby (2009) and *Washington Post* (2011–2012)].

Lipman, Pauline. 2009. *Making Sense of Renaissance 2010 School Policy in Chicago: Race, Class, and the Cultural Politics of Neoliberal Urban Restructuring*. Great Cities Institute Working Paper. http://www.uic.edu/cuppa/gci/publications/workingpaperseries/pdfs/GCP-09-02_Lipman.pdf.

Lipman, Pauline, and Eric Gutstein. 2011. *Should Chicago Have an Elected Representative School Board? A Look at the Evidence*. University of Illinois-Chicago, Collaborative for Equity and Justice in Education. <http://www.uic.edu/educ/ceje/articles/Printed%20school%20board%20report.pdf>.

Lipman, Pauline, and Alecia Person. 2007. *Students as Collateral Damage? A Preliminary Study of Renaissance 2010 School Closings in the Midsouth*. University of Illinois-Chicago, Collaborative for Equity and Justice in Education. <http://www.uic.edu/educ/ceje/articles/midsouth%20initial%20report%201-31-07.pdf>

Loeb, Susanna, Matthew Ronfeldt, and Jim Wyckoff. 2012. *How Teacher Turnover Harms Student Achievement*. Stanford University. <http://blogs.edweek.org/edweek/teacherbeat/TchTrnStAch%20AERJ%20R%26R%20not%20blind.pdf>

Lupescu, Stuart, Elaine M. Allensworth, Paul Moore, Marisa de la Torre, James Murphy, and Sanja Jagesic. 2011. *Trends in Chicago's Schools Across Three Eras of Reform*. Consortium on Chicago School Research at the University of Chicago Urban Education Institute. http://ccsr.uchicago.edu/sites/default/files/publications/Trends_CPS_Full_Report.pdf

Lynch Robert. 2007. *Enriching Children, Enriching the Nation: Public Investment in High-Quality Prekindergarten*. Washington, D.C.: Economic Policy Institute.

Mahaley, Hosanna. 2012. *2012 DC CAS Results*. Office of the State Superintendent of Education. http://osse.dc.gov/sites/default/files/dc/sites/osse/release_content/attachments/DCCAS_2012.pdf

Marsh, Julie A, Matthew G. Springer, Daniel F. McCaffrey, Kun Yuan, Scott Epstein, Julia Koppich, Nidhi Kalra, Catherine DiMartino, and Art (Xiao) Peng. 2011. *A Big Apple for Educators: New York City's Experiment with Schoolwide Performance Bonuses*. RAND Corporation Final Evaluation Report. http://www.rand.org/content/dam/rand/pubs/monographs/2011/RAND_MG1114.pdf

Maxwell, Lesli A. 2012. “Growing Gaps Bring Focus on Poverty’s Role in Schooling.” *Education Week*, March 7. http://www.edweek.org/ew/articles/2012/03/07/23poverty_ep.h31.html

McAdoo, Maisie. 2012. “New HS Progress Reports Spotlight College Readiness.” *New York Teacher*, December 6. <http://www.uft.org/news-stories/new-hs-progress-reports-spotlight-college-readiness>

- Medina, Jennifer. 2007. "At Canarsie High, Now Marked for Closing, Loyalty Prevails." *New York Times*, December 24. <http://www.nytimes.com/2007/12/24/nyregion/24canarsie.html>
- Merriman, James, Christina Brown, Michael Regnier, David Frank, Daniel Hayman, Eric Horowitz, Christian Villenas, and Cara Volpe. 2012. *The State of the NYC Charter School Sector*. New York City Charter School Center. <http://www.nycCharterSchools.org/Data>
- Moore Johnson, Susan, Matthew A. Kraft, and John P. Papay. 2012. *How Context Matters in High-Need Schools: The Effects of Teachers' Working Conditions on Their Professional Satisfaction and Their Students' Achievement*. Project on the Next Generation of Teachers, Harvard Graduate School of Education. http://scholar.harvard.edu/files/mkraft/files/teacher_working_conditions_-_tcr_revision_-_final.pdf
- National Center for Education Statistics (NCES). 2005a. *Characteristics of the 100 Largest Public Elementary and Secondary School Districts in the United States: 2002–03*, Table 17. http://nces.ed.gov/pubs2005/100_largest/tables/table_17.asp
- National Center for Education Statistics (NCES). 2005b. *The Nation's Report Card: Trial Urban District Assessment Reading 2005—Results at Grades 4 and 8*.
- National Center for Education Statistics (NCES). 2010a. *Characteristics of the 100 Largest Public Elementary and Secondary School Districts in the United States: 2008–09*, Tables A-12 and A-14. <http://nces.ed.gov/pubs2011/2011301.pdf>
- National Center for Education Statistics (NCES). 2010b. "NCES Common Core of Data State Dropout and Completion Data File, School Year 2007–08. Characteristics of the 100 Largest School Districts in the United States." http://nces.ed.gov/ccd/pub_100_largest.asp
- National Center for Education Statistics (NCES). 2011a. "The Nation's Report Card: Trial Urban District Assessment Mathematics 2011—Executive Summary." <http://nces.ed.gov/nationsreportcard/pubs/dst2011/2012452.asp>
- National Center for Education Statistics (NCES). 2011b. *The Nation's Report Card: Trial Urban District Assessment Reading 2011—Results at Grades 4 and 8*. <http://nces.ed.gov/nationsreportcard/pdf/dst2011/2012455.pdf>
- National Center for Education Statistics (NCES). 2011c. *The Nation's Report Card: Trial Urban District Snapshot Report 2011—New York City Department of Education Grade 8 Public Schools*. <http://nces.ed.gov/nationsreportcard/pdf/dst2011/2012453XN8.pdf>
- New America Foundation (NAF). 2009–2010. *Federal Education Budget Project* [based on U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 2009–2010]. <http://febp.newamerica.net/k12>
- New York City Department of Education (NYC DOE) Accountability Office. 2012. "Graduation Outcomes Citywide" [data table]. <http://schools.nyc.gov/Accountability/data/GraduationDropoutReports/default.htm>
- New York City Parents Union. 2011. "Parents Union Files Lawsuit over School Closings and Co-locations" [press release]. June 13. <http://www.nycparentsunion.org/archives/173>
- Oloff, Phil, and Michael Leachman. 2011. *New School Year Brings Steep Cuts in State Funding for Schools*. Center on Budget and Policy Priorities. <http://www.cbpp.org/cms/?fa=view&id=3569>
- Pallas, Aaron. 2012a. "Meet the 'Worst' 8th Grade Math Teacher in NYC." *Washington Post's The Answer Sheet* (blog), May 16. http://www.washingtonpost.com/blogs/answer-sheet/post/meet-the-worst-8th-grade-math-teacher-in-nyc/2012/05/15/gIQArmlbSU_blog.html

Pallas, Aaron. 2012b. "The Emperor's New Close." *Gotham Schools*, March 7. <http://gothamschools.org/2012/03/07/the-emperors-new-close/#more-78849>.

Peneston, Jeff, Debra Calvino, Vickie Mike, Rich Ognibene, Marguerite Izzo, Stephen Bongiovi, Elizabeth Day, and Dr. Patricia Jordan. 2011. "Teachers of the Year: APPR Regulations Poison 'Spirit of Collaboration.'" New York State United Teachers, May 23. http://www.nysut.org/cps/rde/xchg/nysut/hs.xsl/nysutunited_16614.htm

Perdido Street School. 2010. "Michelle Rhee Alleges Fired D.C. Teachers Had Engaged In Sexual Misconduct." *Perdido Street School Blog*, January 26. <http://perdidostreetschool.blogspot.com/2010/01/michelle-rhee-alleges-fired-dc-teachers.html>

Pink, Daniel H. 2009. *Drive: The Surprising Truth About What Motivates Us*. New York: Riverhead Books.

Powell, Michael. 2012a. "An Upstairs-Downstairs Divide at a Public High School in East Harlem." *New York Times*, June 25. <http://www.nytimes.com/2012/06/26/nyregion/in-east-harlem-school-building-uneasy-neighbors.html>

Powell, Michael. 2012b. "In Schools Cut by the City Ax, Students Bleed." *New York Times*, April 16. <http://www.nytimes.com/2012/04/17/nyregion/in-schools-cut-by-new-york-citys-ax-students-bleed.html>

Ravitch, Diane. 2010. "The Sound of Bubbles Bursting: Student Gains on State Test Vanished into Thin Air." *The Daily News*, August 1. <http://www.nydailynews.com/opinion/sound-bubbles-bursting-student-gains-state-test-vanished-thin-air-article-1.205311#ixzz0vLyDQqbU>

Reardon, Sean F. 2009. "Review of 'How New York City's Charter Schools Affect Achievement.'" Education and the Public Interest Center & Education Policy Research Unit. <http://epicpolicy.org/thinktank/review-How-New-York-City-Charter>

Reardon, Sean F. 2011. "The Widening Academic Achievement Gap Between the Rich and the Poor: New Evidence and Possible Explanations." In Duncan, Greg J. and Richard J. Murnane, eds., *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*. New York: Russell Sage Foundation.

Rhee, Michelle, and Joel Klein. 2010. "How to Fix Our Schools: A Manifesto by Joel Klein, Michelle Rhee and Other Education Leaders." *Washington Post*, October 10. <http://www.washingtonpost.com/wp-dyn/content/article/2010/10/07/AR2010100705078.html>

Rolnick, Arthur J. and Rob Grunewald. 2007. *The Economics of Early Childhood Development, as Seen by Two Fed Economists*. Federal Reserve Bank of Minneapolis. Fall. http://www.frbf.org/publications/community/investments/0709/economics_early_childhood.pdf.

Ronfeldt, Matthew, Hamilton Lankford, Susanna Loeb, and James Wyckoff. 2011. *How Teacher Turnover Harms Student Achievement*. National Bureau of Economic Research. <http://www.nber.org/papers/w17176.pdf>

Rothstein, Richard. 2004. *Class and Schools: Using Social, Economic, and Educational Reform to Close Black-White Achievement Gaps*. Washington, D.C., and New York: Economic Policy Institute and Columbia University Teachers College.

Rothstein, Richard, Rebecca Jacobsen, and Tamara Wilder. 2008. *Grading Education: Getting Accountability Right*. Washington, D.C., and New York: Economic Policy Institute and Teachers College Press.

Rothstein, Richard, and Mark Santow. 2012. *A Different Kind of Choice: Educational Inequality and the Continuing Significance of Racial Segregation*. Economic Policy Institute Working Paper. <http://www.epi.org/publication/educational-inequality-racial-segregation-significance/>

- Roy, Joydeep, and Lawrence Mishel. 2008. "Using Administrative Data to Estimate Graduation Rates: Challenges, Proposed Solutions and their Pitfalls." *Education Policy Analysis Archives*. Vol. 16, no. 11. <http://epaa.asu.edu/ojs/article/view/36/162>
- Ryan, Howard. 2012. "Occupying Parents Win an Audience, but Chicago Schools Still Privatize." *Labor Notes*, February 12. <http://labornotes.org/blogs/2012/02/occupying-parents-win-audience-chicago-schools-still-privatize>.
- Saltman, Kenneth J. 2010. *Urban School Decentralization and the Growth of "Portfolio Districts."* Education Policy Research Unit, and Education and the Public Interest Center. <http://nepc.colorado.edu/publication/portfolio-districts>
- Samuels, Christina. 2011. "Study Warns of Limited Savings from Closing Schools." *Education Week*, November 1. http://www.edweek.org/ew/articles/2011/11/02/10closings_ep.h31.html.
- Santos, Fernanda. 2005. "Mayor Bloomberg's Promises for Education: An Annotated Scorecard." *New York Times Schoolbook*, January 13. <http://www.schoolbook.org/2012/01/13/mayor-bloombergs-promises-for-education-an-annotated-scorecard/>
- Santos, Fernanda. 2011. "College Readiness is Lacking, City Reports Show." *New York Times*, October 24. <http://www.nytimes.com/2011/10/25/education/25progress.html>
- Sawchuk, Stephen. 2010. "Merit-Pay Model Pushed by Duncan Shows No Achievement Edge." *Education Week*, June 9. http://www.edweek.org/ew/articles/2010/06/01/33tap.h29.html?qs=Merit-Pay_Model_Pushed_by_Duncan_Shows_No_Achievement_Edge
- Scherrer, Jimmy. 2011. "Measuring Teaching Using Value-Added Modeling: The Imperfect Panacea." *NASSP Bulletin*. Vol. 95, no. 2, 122–140.
- School District of Philadelphia. 2012. *A Blueprint for Transforming Philadelphia's Public Schools*. http://thenotebook.org/sites/default/files/BlueprintPublicPresentation_4_22_12.pdf
- Schwartz, Amy Ellen, Leanna Stiefel, and Matthew Wiswall. 2011. *Do Small Schools Improve Performance in Large, Urban Districts? Causal Evidence from New York City*. Institute for Education and Social Policy at New York University, Working Paper #04-11. http://steinhardt.nyu.edu/scmsAdmin/media/users/spa2/Working_Paper__01-12.pdf
- Schwartz, Heather. 2012. "Housing Policy is School Policy: Economically Integrated Housing Promotes Academic Success in Montgomery County, Maryland." In Richard D. Kahlenberg, ed., *The Future of School Integration: Socioeconomic Diversity as an Education Reform Strategy*. New York: Century Foundation Press.
- Scott, Liane. 2012. "School Closings and the Displacement Equation." *Grassroots DC* (blog), February 10. <http://www.grassrootsmediaproject.org/2012/02/school-closings-and-the-displacement-equation/>
- Simon, Mark. 2012. "Is Teacher Churn Undermining Real Education Reform in D.C.?" *Washington Post*, June 15. http://www.washingtonpost.com/opinions/is-teacher-churn-undermining-real-education-reform-in-dc/2012/06/15/gJQAigWcfV_story.html
- Sparks, Sarah D. 2011. "Study Leads to End of New York City Merit-Pay Program." *Education Week*, July 20. http://blogs.edweek.org/edweek/inside-school-research/2011/07/a_new_study_by_the.html
- Springer, Matthew G., Dale Ballou, Laura Hamilton, Vi-Nhuan Le, J.R. Lockwood, Daniel McCaffrey, Matthew Pepper, and Brian Stecher. 2010. *Teacher Pay for Performance: Experimental Evidence from the Project on Incentives in Teaching*. National Center on Performance Incentives at Vanderbilt University. http://www.rand.org/content/dam/rand/pubs/reprints/2010/RAND_RP1416.pdf

Starr, Joshua P. 2012. Testimony before the United States House of Representatives, Committee on Appropriations, Subcommittee on Labor, Health, and Human Services, Education, and Related Agencies. March 27.

Steinberg, Matthew, Elaine Allensworth, and David W. Johnson. 2011. *Student and Teacher Safety in Chicago Public Schools: The Roles of Community Context and School Social Organization*. Consortium on Chicago School Research at the University of Chicago Urban Education Institute, and American Institutes for Research. <http://ccsr.uchicago.edu/publications/student-and-teacher-safety-chicago-public-schools-roles-community-context-and-school>

Strauss, Valerie. 2011. "Teachers of Year Fight Test-Based Assessment." *The Answer Sheet* (Washington Post blog), May 27.

Strauss, Valerie. 2012. "Moco Schools Chief Calls for Three-year Moratorium on Testing." *The Answer Sheet* (Washington Post blog), December 10. <http://www.washingtonpost.com/blogs/answer-sheet/wp/2012/12/10/moco-schools-chief-calls-for-three-year-moratorium-on-standardized-testing/>

StudentsFirst. 2011. "The Name-calling Must Stop" [email to StudentsFirst mailing list, accessed August 15 at <http://groups.yahoo.com/group/nyceducationnews/message/39272>].

StudentsFirst. 2012. "Our Mission." <http://www.studentsfirst.org/pages/our-mission>

StudentsFirst. 2013. "State of Education: State Policy Report Card 2013." http://edref.3cdn.net/9e8505b2c4ad5ec0e8_u6m6ikky8.pdf.

Stuit, David A. 2010. *Are Bad Schools Immortal? The Scarcity of Turnarounds and Shutdowns in Both Charter and District Sectors*. Basis Policy Research and the Thomas B. Fordham Institute. <http://www.edexcellence.net/publications/are-bad-schools-immortal.html>

Turque, Bill. 2008. "Rhee Bypasses Talks, Imposes Dismissal Plan." *Washington Post*, October 3. <http://www.washingtonpost.com/wp-dyn/content/article/2008/10/02/AR2008100201672.html>.

Turque, Bill. 2009. "DC Mayor Accused of Ruse to Fire Unionized Teachers." *Washington Post*, September 18. <http://www.washingtonpost.com/wp-dyn/content/article/2009/09/17/AR2009091702880.html>

Turque, Bill. 2010a. "6 D.C. Schools to be Overhauled After Failing to Meet 'No Child' Goals." *Washington Post*, June 15. <http://www.washingtonpost.com/wp-dyn/content/article/2010/06/14/AR2010061405068.html>

Turque, Bill. 2010b. "D.C.'s Dunbar High School Getting New Administrators, More Security to Quell Violence." *Washington Post*, December 4. <http://www.washingtonpost.com/wp-dyn/content/article/2010/12/04/AR2010120403313.html>

Turque, Bill. 2010c. "Rhee Dismisses 214 D.C. Teachers; Union Vows to Contest Firings." *Washington Post*, July 24. <http://www.washingtonpost.com/wp-dyn/content/article/2010/07/23/AR2010072303093.html>

Turque, Bill. 2011a. "D.C. Educators Rated 'Effective' Can Still Lose Jobs." *Washington Post*, November 2. http://www.washingtonpost.com/local/education/dc-educators-rated-effective-can-still-lose-jobs/2011/10/27/gIQAxysNgM_story.html

Turque, Bill. 2011b. "Huge Achievement Gaps Persist in D.C. Schools." *Washington Post*, August 5. http://www.washingtonpost.com/local/education/huge-achievement-gaps-persist-in-dc-schools/2011/08/05/gIQAB7b2yI_story.html

Turque, Bill. 2012a. “‘Creative... Motivating’ and Fired.” *Washington Post*, March 6. http://www.washingtonpost.com/local/education/creative-motivating-and-fired/2012/02/04/gIQAwwZpvR_story.html

Turque, Bill. 2012b. “DCPS Budget: IMPACT’s Impact.” *Washington Post*, March 6. http://www.washingtonpost.com/blogs/dc-schools-insider/post/dcps-budget-impacts-impact/2012/03/04/gIQA1VVhvR_blog.html

Urban Youth Collaborative. 2011. *No Closer to College: NYC High School Students Call for Real School Transformation, Not School Closings*. <http://www.urbanyouthcollaborative.org/wp-content/uploads/2011/05/No-Closer-to-College-Report.pdf>

U.S. Census Bureau. 2012. *Public Education Finances: 2010*, Table 8, “Per Pupil Amounts for Current Spending of Public Elementary-Secondary School Systems by State: 2009–2010.” <http://www2.census.gov/govs/school/10f33pub.pdf>

Weiss, Elaine. 2011. “The Elusive Value in Value-Added: Technical and Practical Obstacles Contribute to Unreliability and Bias, Complicating the Sorting of Teachers on the Basis of Student Performance.” *The School Administrator*. Vol. 9, no. 68, 18–22. <http://www.aasa.org/SchoolAdministratorArticle.aspx?id=20428>

Wikipedia. 2013. “Campbell’s Law.” http://en.wikipedia.org/wiki/Campbell%27s_law

Winerip, Michael. 2012. “Hardworking Teachers, Sabotaged When Test Scores Slip.” *New York Times*, March 4. <http://www.nytimes.com/2012/03/05/nyregion/in-brooklyn-hard-working-teachers-sabotaged-when-student-test-scores-slip.html>

Yoshikawa, Hirokazu, J. Lawrence Aber, and William R Beardslee. 2012. “The Effects of Poverty on the Mental, Emotional, and Behavioral Health of Children and Youth: Implications for Prevention.” *American Psychologist*. Vol. 67, no. 4, 272–284.