



Social Enterprise Leadership Forum: Measuring and Creating Excellence in Schools



Featured Speakers

Ray Fisman

Lambert Family Professor of Social Enterprise
Director
Social Enterprise Program
Columbia Business School

Joel Klein

Chancellor
New York City Department of Education

Caroline Hoxby

Scott and Donya Bommer Professor of Economics
Stanford University

Jonah Rockoff

Sidney Taurel Associate Professor of Business
Columbia Business School

Douglas Staiger

John French Professor in Economics
Dartmouth College

John King

Senior Deputy Commissioner for P-12 Education
New York State Department of Education

James Liebman

Simon H. Rifkind Professor of Law
Columbia Law School
Former Chief Accountability Officer
New York City Department of Education

Emary Aronson

Managing Director of Education
Robin Hood Foundation

Written by Jennifer Itzenson

Photography by Leslye Smith

Design by Carolyn Tharp



Introduction

Despite billions of dollars spent in recent decades with the best of intentions, American students continue to underperform their peers in other developed nations. If a country's economic future is built on the knowledge and expertise of its youth, the United States is a nation at risk.

Yet an enormous amount of energy has coalesced over the past few years around the challenges of improving American education. This has led to new models of governance such as charter schools, entrepreneurial successes like Teach for America, new applications of technology in classrooms and the home, and even the redesign of entire school systems. We're still grappling with whether and how these many innovations have translated into better student outcomes, and also whether proven successes can be adapted or put in place on a larger scale.

The purpose of the inaugural Social Enterprise Leadership Forum was to take stock of some recent successes in education reform, and to reflect on lessons for the future. The forum, held on May 6, 2010, brought together leaders from the New York education community to engage in conversation with leading thinkers on education reform from both academia and practice.

The discussions presented a range of views on what will be required to improve education in New York and beyond. How should teachers be evaluated, and how can this information be used to improve classroom instruction? What lies behind the success of charter schools in New York, and how may their innovations be applied in public schools here and around the country? Should underperforming teachers be reformed or removed?

The discussions heard throughout the day—summarized in the pages that follow—helped to shed considerable light on these and other questions that are central to current debates about school reform. At least as important, as disseminating information on innovative models and ideas, the day finished with new questions for reformers and researchers on what is needed to ensure that students across the country have access to the educational opportunities they deserve.



Keynote Presentation: Accountability: What's It Really All About?

In 1983, a presidential commission issued the landmark report "A Nation at Risk: The Imperative For Educational Reform." The report, which warned the public education system was at risk of failure despite an increase in spending, declared, "If an unfriendly foreign power had attempted to impose on America the mediocre educational performance that exists today, we might well have viewed it as an act of war." The report galvanized a movement, both within school administrations and communities, to improve public education and ensure that the nation's children would be able to compete in a global economy.

Almost three decades have passed since "A Nation at Risk," and in that time, the United States has more than doubled its spending in real dollars on public education, said Joel Klein, chancellor of New York City's Department of Education, in his keynote presentation. Yet the system continues to languish. "If anything, the results are moving in the wrong direction," Klein said.

Klein, who has headed the 1.1 million-student New York City school system since 2002, argued that the heart of the problem is a resistance to change. Although school administrators have sought and received increases in funding, these additional resources

have largely reinforced the status quo. This has led to a widening of the achievement gap between US students and their counterparts in other developed nations. In the Organization for Economic Cooperation and Development's Program for International Student Assessment, the United States ranking in math slipped from 18 of 27 in 2000 to 25 of 30 in 2006. In science, its ranking fell from 14 to 21. An OECD study on college-level graduation rates found that the nation's relative position slipped by 15 points from 1995 to 2005. (The United States was one of the only countries that experienced a decline in its graduation rate during this period.) Meanwhile, within the United States, the racial achievement gap has hardly budged. The National Assessment of Educational Progress found that the gap in eighth-grade reading ability between white students and black students closed by only 3 points, from 29 to 26, between 1992 and 2009. In math, the gap remains even wider.

"Where else would we settle for those kinds of numbers and that kind of downward trend?" asked Klein. "People say it's the kids, it's poverty, it's the families. I grew up in poverty, in public housing. I understand dysfunction in families." Though the challenges are daunting, Klein argues that it is still possible to achieve measurable progress: "We know that with the same kids in different schools—indeed the same kids in different classrooms—we can get dramatically different outcomes."

The data support this. NAEP results show dramatic differences in the performance of low-income students in school districts throughout the country. Analysis of 2009 NAEP data on fourth-grade math proficiency found that low-income black students in Boston scored 24 points higher than similar students in Los Angeles. In reading, there was a 36-point difference between Austin and Detroit. “That’s not family, that’s not ZIP code, that’s not skin color, that’s not poverty,” Klein said. “That’s the power of education—even at the school district level.”

The solution, Klein argued, is attracting and retaining effective teachers. Studies have shown that public schools essentially recruit teachers from the lowest-achieving third of high school students who attend college. The result has been a lack of effective teachers, and low-income neighborhoods have suffered disproportionately from this shortage. “Students with the greatest needs tend to have access to the least qualified and least effective teachers,” noted a study by the Aspen Commission. “If we don’t focus on the way we recruit, the way we retain, the way we reward, the consequences and accountability for teaching, we’re not going to succeed,” Klein said.

The current system mandates lockstep pay, seniority and lifetime tenure—what Klein called the “three pillars of mediocrity.” He described his difficulty in recruiting

math teachers as one of the most glaring examples of the system’s failure. Under current rules, physical education and math teachers must receive the same pay, even though math teachers are in far greater demand. “And when I’m short highly effective math teachers, which kids do you think are short: The kids on the east side of Central Park, the kids on the west side of Central Park [in the affluent areas], or those on the north side?” Klein said. It goes without saying it’s the poor kids on the north side in Harlem.

To ensure that teachers are effective, public schools need to focus on accountability, Klein said. No accountability system is perfect—all measurements have their flaws. They’re imprecise. They can be gamed. But the imperfect measures of progress are far better than nothing at all. Klein stressed the importance of choosing a method and committing to it. In recent years, New York has introduced a system of three key evaluation measures: progress reports, which compare schools based on student progress and performance; quality reviews, a measurement of performance management criteria; and school surveys, which assess information from teachers, parents, and students.

Corresponding to these evaluation measures are assessments that are designed to improve school performance, Klein said. These include periodic

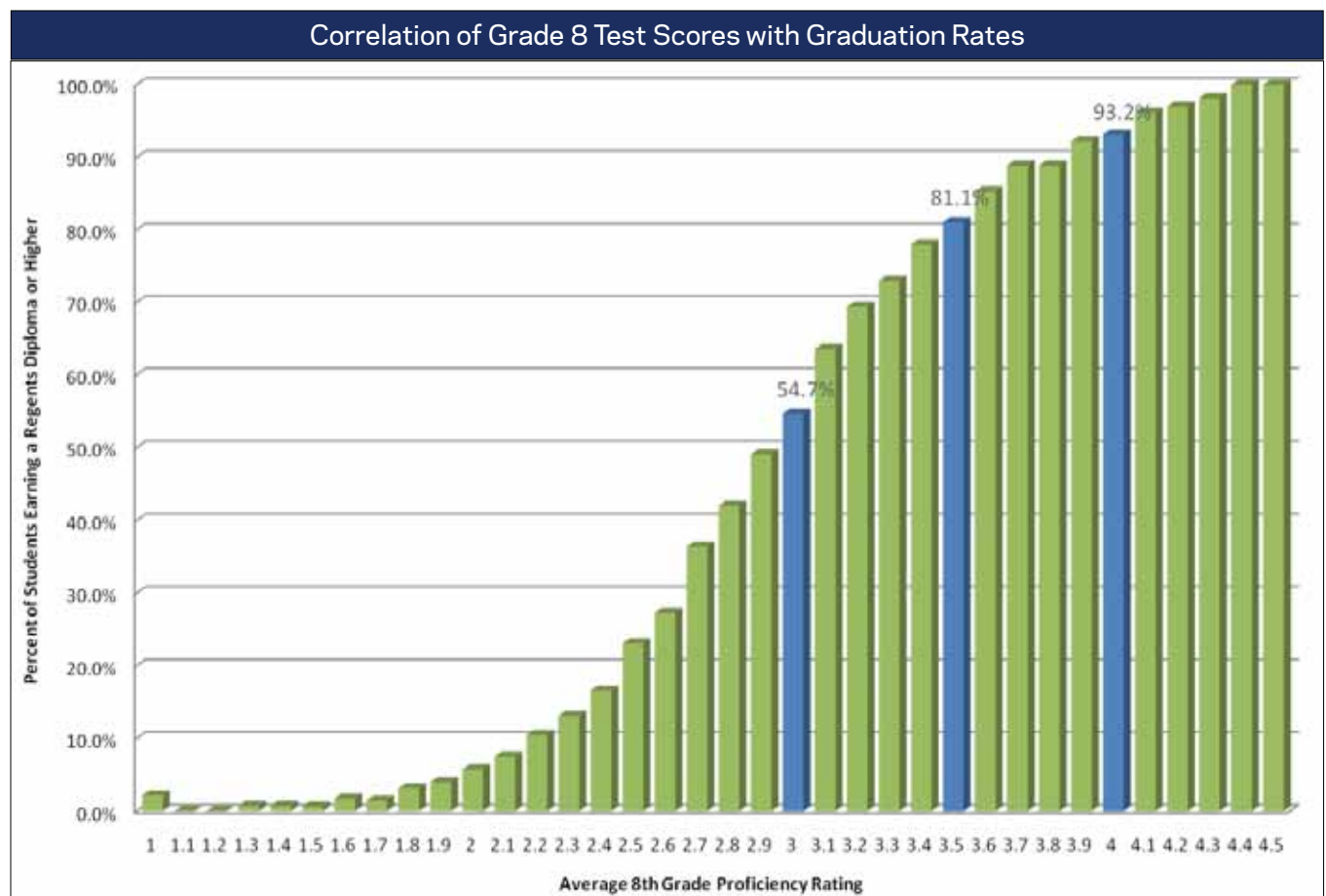


Figure 1: Eighth grade proficiency ratings are very predictive of high school Regents diplomas. From presentation by Joel Klein, chancellor of New York City’s Department of Education.



assessments, which track progress throughout the school year to enable midcourse corrections; teacher data reports, which focus accountability at the teacher-level; and the Achievement Reporting and Innovation System, known as ARIS, which Klein called “one of the most effective data systems I’ve seen, and not just in education.” Together, these measurements are the foundation of an evolving accountability system, Klein said.

Why do all of these measurements matter? Klein gave the example of proficiency tests for eighth-grade students. Research has shown that scores on these tests are highly predictive of Regents graduation rates, and graduation in turn translates into better jobs and higher incomes for New York’s students. (In New York State, most public school students must pass a series of subject exams to receive what is known as a Regents diploma.) Students scoring 3 on the proficiency test had a graduation rate of 55 percent. A score of 3.5 corresponds to a graduation rate of 81 percent. (See figure 1.)

ARIS, Klein said, has been particularly popular for families seeking information about student performance. Through the ARIS system—which several hundred thousand parents are currently using—users can get data about year-to-year performance at the classroom, student and school levels. “This is the most important piece for us, and it constantly evolves,” Klein said.

But perhaps the city’s most critical initiative is creating choice—by launching more than 400 new schools, including 82 charter schools, Klein said. Charter schools have sparked the most controversy. Critics argue that charters perform no better, on average, than traditional schools. Klein, a strong supporter of charters, said most of New York City’s charters perform favorably in comparison to the average school, and that much of the criticism stems from politics. (This view is supported by the New York Charter evaluation described in the following section.)

Charter schools are “schools of choice” that no child is required to attend, Klein said. (Charters in New York State are required to be free, and must enroll applicants without any screening process.) “One way to bring accountability to a system is to bring people real choice.” No parent in the city would tolerate a system without choice, he said: “There’s not a person who would allow me to assign their kid randomly to a public school in New York. Why shouldn’t everyone be allowed to vote with their feet?”

It is an ambitious agenda, and one likely to spur further controversies, Klein conceded. But he said he has never been more optimistic about improving public education and called on reformers to be bold. “The future of this country depends on leaders willing to take a position because it’s right,” Klein said, not because it’s safe or popular. “I think we know what is right, and I think our kids and our nation depend on it.”



Keynote Presentation: What Makes Charter Schools Effective?

Why are charter schools—and questions about their performance—so controversial? Every week brings articles and opinion pieces in major media outlets about the role charter schools should play in public education. In the span of one month, the *New York Times* published a front-page assessment of the success of charter schools, an op-ed challenging the relevance of test scores for charters, and a feature on the support of charters by hedge fund executives. Among education experts, there has similarly been sharp debate on whether charter schools outperform traditional public schools, and if so, why.

At least among New York's charters, there is compelling evidence that—at least at their current scale—they're doing a better job of educating kids than the city's public schools. Caroline Hoxby, the Scott and Donya Bommer Professor of Economics at Stanford University, has conducted a rigorous analysis of the performance of New York City charter schools. She obtained her data through a lottery-based study. About 88 percent of students who attend charter schools in the United States, and 97 percent of those in New York, are admitted by lottery. In a lottery-based study, researchers compare students who attend charter schools to students who applied but did not get in and therefore remain in traditional public schools.

These "lotteried-out" students serve as the perfect control group, which is important because students who attend charter schools tend to be different from public school students, Hoxby said. For example, charter school students are more likely to be urban, poor, and either black or Hispanic. (This is true not only in New York, but across the United States.) Charter school students also share distinct traits that cannot be easily observed or measured. "Many people surmise—perhaps accurately, but we don't know—that parents who send their kids to charter schools are more motivated, given their poverty and other disadvantages, than parents who do not apply to charter schools," Hoxby explained. "We also know that kids who fit very well into traditional public schools and are thriving there do not tend to apply to charter schools." In a lottery-based study, the subjects, on average, have equal levels of motivation, have experienced similar difficulties fitting into traditional public schools, and have completed the same application process, Hoxby said.

Studies that rely on the "matching" method, which requires a researcher to find a traditional school student whose attributes match those of a charter school student, are less reliable: you can never be sure that you've found the "right" match. After all, you're comparing students who chose to apply to a charter school versus those who chose not to, so questions about differences in their families, their emphasis on educational opportunities, and many other factors that can't be accounted for in the matching

Measure of Charter School's Effects

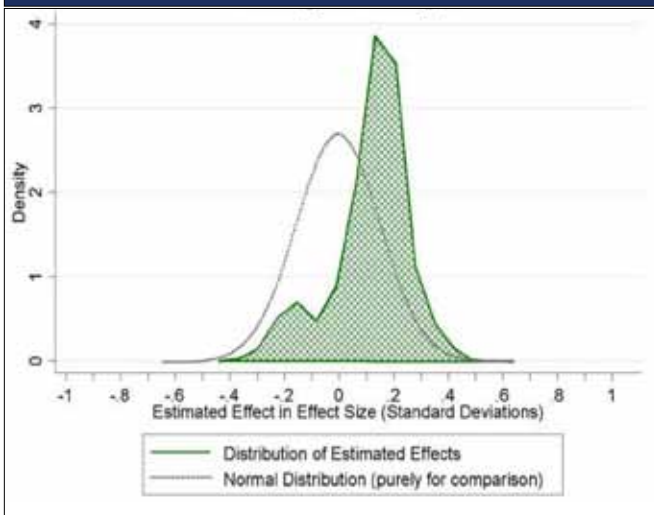


Figure 2: Distribution of charter school's effects on mathematics tests for grades three through eight. *From the presentation by Caroline Hoxby, the Scott and Donya Bommer Professor of Economics at Stanford University.*

process remain. At least some of the controversy surrounding charter schools stems from this inconsistent quality in research, which has made it difficult for both the media and the public to judge accurately whether charter schools are successful, Hoxby said.

In Hoxby's study, she analyzed the performance of New York City charter schools from the year 2000 onward. She focused on statewide exam scores and high school graduation rates to determine the per-year effect of attending a charter school. Hoxby found that charter schools improve students' performance by 3 to 4 points a year on Regents exams (on a 100-point scale). And for every year that a student spends in charter schools, the student is 7 percent more likely to graduate. "These effects are moderate if we think of a single year," Hoxby said. "But attending a charter school for multiple years can close the achievement gap."

This becomes evident when one takes a closer look at Regents scores. Based on their characteristics (family, test scores, etc.), Hoxby's analysis predicts that an average student who applies to a charter school will, on average, get a Regents exam score of 60—below the passing grade of 65. If these applicants spend two years in a charter school, they are likely to move above the passing threshold, Hoxby's research showed. And a student who spends four years in a charter school is 28 percent more likely to graduate with a Regents diploma than a student who was lotteried out.

"Attending charter schools has a cumulative effect," Hoxby said. "And it's that accumulation that can really close the gap." In fact, Hoxby's research showed that students who attend charter schools from the third through eighth grades tightened the achievement gap by about 30 points. That is almost enough to completely close the 35 point gap that exists between students in the affluent, suburban

school district of Scarsdale, New York, and students in the low-income, urban neighborhood of Harlem.

But what makes these charter schools so successful? "It must be things the charters have in common, because on average charters have a positive effect," Hoxby said. (See figure 2.) Common traits among charter schools include a system of governance similar to those of nonprofits, rather than of school boards; layers of accountability to various levels of government; accountability to parents, who can remove their children from the school; autonomous management, including discretion over the hiring and firing of teachers; routine assessments of students' performance; and school uniforms or strict dress codes.

From this list, we can get suggestions about what may work in charter schools, Hoxby said. She emphasized that researchers cannot prove that any of these common traits are directly linked to success. (Without random testing of the policies, researchers cannot draw any firm conclusions about whether these attributes cause charter success or whether they are the results of other unobserved differences between the schools, Hoxby explained. For example, routine assessment might be critical, or it could just be one indication of organized and effective leadership.) Hoxby posited that some charter school policies are likely contributing to the positive effect, including a longer school year and school day, a greater amount of time spent on reading and math, and a disciplinary system that focuses on small rewards and small punishments. (See figure 3.) Surprisingly, a school's longevity does not seem to matter. "Schools that have

Charter Schools' Policies Are Different

	Average for NYC Charter Students	Average for NYC Regular Public Students
Days in the school year	192	178
Hours in the school day	8	5.5
Saturday school	57%	0%
Minutes of reading each day	112	80
Long math period (90+ minutes)	50%	5%*
Routine internal assessments	92%	10%*
Parent contract	52%	0%
Small reward/punishment discipline	22%	-
Uniforms or strict dress code	89%	3%*

Figure 3: Caroline Hoxby, the Scott and Donya Bommer Professor of Economics at Stanford University, compared charter schools' policies to other New York City public schools.

* Hoxby's estimation of policies for regular NYC public schools.



been open for four years are doing just as well as schools that have been open for 10 years," Hoxby said.

Is the success limited to New York? There's no evidence that suggests the performance of the city's charters is exceptional, Hoxby said. Unfortunately, very few charter schools in other cities have been evaluated using the lottery method, so it is difficult to make comparisons.

New York City may have some advantages relative to others around the country. First, they may be benefiting from their relatively high concentration. "Charters learn from one another, both formally, in that they imitate one another when something is going well in another school, and also informally, because they hire a lot from one another," Hoxby said. "Teachers tend to spread good advice." At the same time, schools in the city face particular challenges, including the difficulties and expense of securing space.

"There is no obvious reason why the supply of effective charter schools will run out or why other cities cannot achieve the same success as New York," Hoxby argued. Yet unless researchers and policy experts raise the standard of evidence for assessing charter schools, the controversies are likely to persist, she said. "Testing charter school-type policies in traditional public schools would help," Hoxby said. "Without testing, we will still be here in 20 years, dependent on anecdotes about what works and why."

Using Teacher Evaluation to Improve School Performance

On all sides of the charter schools debate, reformers and educators stress the importance of making teachers more effective. But what is meant by "effective"? As Jonah Rockoff, the Sidney Taurel Associate Professor of Business at Columbia Business School, explained in his presentation with Douglas Staiger, the John French Professor in Economics at Dartmouth College, this question is often answered through an input- or an output-based perspective. Those taking an input view base their judgments on a teacher's qualifications and actions, while those taking an outcomes view turn instead to the performance of the teacher's students.

In recent years, most researchers have leaned toward the outcomes view. Economists with this perspective often take what is called a value-added approach. In this approach, researchers measure the performance of students on standardized tests, and compare this performance to a counterfactual, or hypothetical, expectation. For example, suppose 80 percent of students in a class pass an exam. To determine whether the students' teacher was effective, researchers need to estimate the percentage of these students that would have passed the exam had they been taught by the average teacher in a school district or city. If the answer

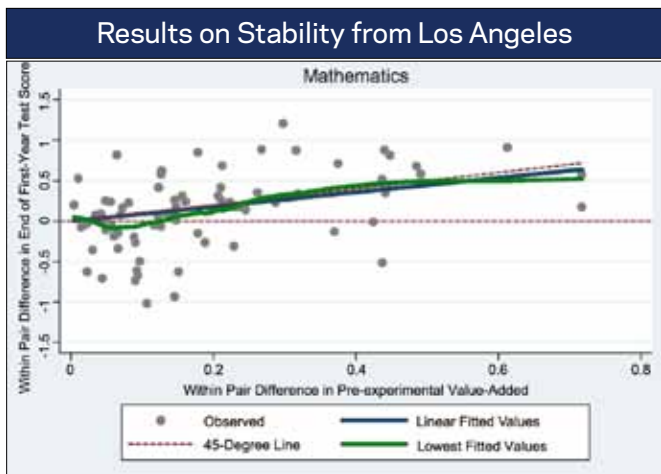


Figure 4: Figure taken from the paper, “Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation,” by Thomas J. Kane and Douglas O. Staiger (NBER Working Paper 14607). From the presentation of Jonah Rockoff, the Sidney Taurel Associate Professor of Business at Columbia Business School.

is less than 80 percent, then the student’s teacher was more effective than average; if the answer is more than 80 percent, the teacher was less effective.

The challenge is estimating the counterfactual expectation: How would a particular student have fared if taught by the average teacher? To make this prediction, researchers first find all of the students in the district who had the same test scores and similar socioeconomic characteristics as the selected student on the previous year’s exam. The researchers then see how these similar students performed on this year’s exam, and use their average scores to represent how the selected student would have performed on this year’s exam if taught by the average teacher.

Rockoff, Staiger, and other scholars have found that the value-added approach—using a range of methodologies and data from an array of cities and states—reveals significant variation in the effectiveness of teachers. There is much variability within each school, Rockoff said. “It’s not that all the effective teachers are in this school and the ineffective teachers are in that school,” he said. “Each school has its share of effective and ineffective teachers.”

Staiger’s study of public school students in Los Angeles illustrates this point. (See figure 4.) Staiger paired random sets of teachers, measured their value-added, and plotted the differences in each pair’s value-added on a graph. The following year, the teachers were randomly assigned to classrooms, and Staiger compared their students’ performance. For those pairs of teachers that had a relatively large difference in value-added, the difference in the performance of their students that year was also large, he found. Similarly, for those pairs of teachers with a relatively small difference in value-added, there was little difference in student performance.

The value-added approach has no shortage of critics. How does this approach account for the unlucky teacher who

always gets the difficult students? Critics also point to the imprecision of the data and the process, and the risk of linking teacher evaluations—which are often tied to pay and promotions—to a measure that is inexact and varies from year to year.

However, Rockoff and Staiger argue that concerns about the approach are swamped by their value in holding teachers accountable. “While bias and imprecision are issues, the value-added approach appears to contain real power to predict teacher effectiveness as measured by student achievement,” Rockoff said. “The stability of the value-added estimates across years is enough to appear useful in teacher evaluation. And while bias can certainly matter for individual teachers, the results are predictive on the level of a district or city.” To illustrate this point, Rockoff, Staiger and colleague Tom Kane from the Harvard Graduate School of Education divided New York City teachers into quartiles based on two years of student performance. The researchers then compiled data from the two following years. Students who were assigned to the bottom quartile of teachers performed far worse on average than those assigned to the top quartile of teachers (see figure 5, page 9).

In fact, the value-added approach seems particularly promising because it is so difficult to otherwise predict whether a teacher will be effective, Rockoff explained. Decades of studies have shown that qualifications such as a teacher’s certification, graduate education, and exam scores have only a modest link with student performance. Studies by Rockoff, Staiger, and other economists have shown that while factors such as a teacher’s personality, cognitive skills, and knowledge of content have moderate predictive power when taken as a whole, there is no single, easily observable trait that predicts whether a teacher will be effective. And while there is consistent evidence that direct observation of teachers can predict gains in student achievement, this subjective method elicits the same

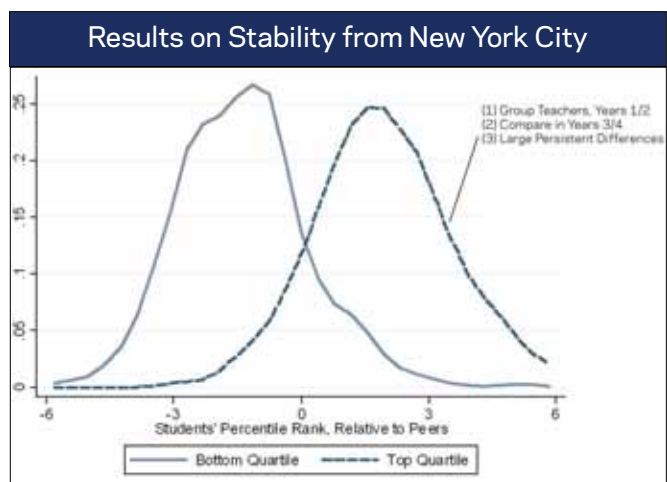


Figure 5: The figure above shows relative student performance in math for bottom- and top-quartile teachers in the third and fourth years in which they are observed teaching. Quartiles are defined based on estimates of teacher value-added in the first and second years in which they are observed teaching. From the presentation of Jonah Rockoff, the Sidney Taurel Associate Professor of Business at Columbia Business School.

concerns as the value-added approach, such as the risks of bias and imprecision.

As an alternative, Rockoff offered what he called a modest proposal: providing information gleaned from the value-added approach to principals, who can combine this data with insights gained through direct observation. "Principals are entrusted with making personnel decisions," Rockoff said. "And at least in some areas of the country, they are responsible for the performance of their schools." New York City is piloting a randomized trial of this technique, and Rockoff and his research partners are following the results.

"There is a way to bring more rigorous, outcome-based analysis to teacher evaluation—teacher evaluation as we conceive it as someone in the back of a classroom observing a teacher," Rockoff said. "So it wouldn't just be a researcher at a computer, taking a value-added approach, or someone just saying whether teachers are doing a good job without also consulting the data."

Staiger pursued this idea further in his study on how to use value-added measurements as a tool for hiring and retaining teachers. While there are concerns that value-added estimates are too imprecise to be used in high-stakes decisions, the researchers "often use imprecise measures to improve things," Staiger said. "The question is how these measurements could be used, given their statistical properties, and how aggressive we should be using them."

Drawing on data from New York City and Los Angeles, Staiger simulated the potential gain of screening teachers using value-added calculations. His objective was to devise a formula that maximized student achievement by screening out ineffective teachers early in their careers. In his model, he calculated teachers' value-added at the end of their first year of teaching. Then he calculated the percentage of teachers that should be dismissed.

The answer generated by his calculations appeared extreme when Staiger first ran the numbers: Dismiss 80 percent of teachers after the first year. Only the most effective 20 percent should go on to teach a second year. "When we first saw the results, we thought this couldn't be right," Staiger said. "But differences in teacher effectiveness are large and persistent, relative to short-lived costs of hiring a new teacher. And the cost of retaining an ineffective teacher outweighs the costs of dismissing an effective teacher." (See figure 6.)

Such a policy presents obvious challenges, ranging from high turnover rates (including the risk that retained teachers could be poached by other schools), larger costs than the model assumes for the hiring and firing of teachers, and the much higher pay that may be required to offset job insecurity. Staiger examined whether requiring a second or even a third year of evaluation would limit these costs. The results surprised him. "Principals do worse if they wait, because they're keeping around teachers they know they want to get rid of," he said. However, allowing principals the option to evaluate some teachers for a second or third year can prove valuable. "For two-thirds

of teachers, principals know after the first year," Staiger said. "It's very unlikely they'll get more information down the road that will change their mind." That is, value added estimates have enough precision that they can pinpoint very poor teachers after a year in the classroom.

Despite the challenges, Staiger's model showed the 80 percent rule would lead to improvements in student performance similar to those achieved by significantly reducing class size or in successful charter schools. "There may be practical reasons limiting success of this strategy," Staiger said. "But the real message from this, which I didn't expect, is the importance of selecting only the most effective teachers—and doing it quickly. You want to act fast." Staiger is not advocating firing 80 percent of new hires in schools across America; however, this is what is required to improve schools until a better way can be found to identify good teachers at the entry-level or until proven models can be developed to improve teachers once they're in the system. If a way to move one of these three levers is not figured out, the United States will remain saddled with mediocre and declining schools.

In the discussion following the presentation, Stanford University's Caroline Hoxby asked if Staiger's research implied that education schools are failing to help future teachers predict if they will be effective. "Most teachers take five years to get their basic certification," Hoxby said. "There's no reason why we should wait until they're rookies [already certified] to try to learn some of this information."

Staiger agreed. Gains could be tripled if effectiveness were measured before the hiring stage, he said. "There may be a way to measure this without having teachers in a classroom, such as in a student-teaching environment or in summer school," Staiger said. "This is a push toward trying to get some signal—even a little signal—as early as possible."

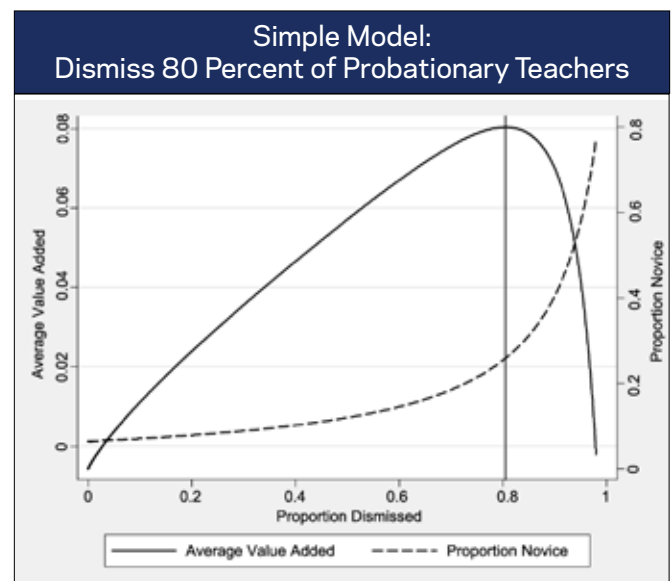


Figure 6: Does firing ineffective teachers improve student performance? From the presentation of Douglas Staiger, the John French Professor in Economics at Dartmouth College.



Ray Fisman, the Lambert Family Professor of Social Enterprise and director of the Social Enterprise Program at Columbia Business School, asked what Staiger's proposal would mean for the culture of teaching. The answer: a big boost in salary and prestige. "If you have a profession in which only 20 percent make it to the hiring stage and get to keep their jobs, and you make people get certification before they can even take the risk of trying to get and keep that job, you're going to have to pay them a lot more money," Rockoff said. But perhaps the change in status would be even more significant, he said. "This would move us somewhat closer to other countries—those in which teachers are drawn from the absolute highest performing academic achievers."

Scaling Up—Building Successful Schools in New York City and New York State

There is compelling evidence that New York City's charter schools are significantly outperforming the average public school. With so many city-run schools either failing or generating mediocre results, why not vastly increase the presence of charters?

Unfortunately, the city and state must contend with the realities of politics and with bureaucratic structures that have resisted change for decades. Therefore, while administrators are working to transform the system, they must also improve student outcomes within the existing structure. This is the argument put forth by James Liebman, the Simon H. Rifkind Professor of Law at Columbia Law School and former chief accountability officer of New York City's Department of Education, and John King, the senior deputy commissioner for P-12 education for New York State's Department of Education. In their presentations, Liebman and King both stressed the need for increasing accountability and effectiveness in the short term. "We've waited too long, and I don't think there are particularly good models yet for taking charter schools to scale," Liebman said. "We have to do something for the kids that we have."

In recent years, New York City-run schools have emulated some of the strategies of charter schools, Liebman said.



This has led to significant performance improvements. From 2006 to 2010, for example, the percentage of eighth graders with scores indicating they would likely graduate college-ready four years later without need of remedial coursework increased from 10 percent to 60 percent in math, and climbed from 43 percent to 57 percent in English language arts.

Though some critics say the higher scores reflect easier tests, the data suggests otherwise. New York City students are closing the gap with the rest of the state, from 22 percent to 13 percent in English language arts, and from 25 percent to 7 percent in math. And, after more than 15 years of little change, the percentage of New York City students graduating with a Regents diploma has climbed 33 percent from 2002 to 2009. (See figure 7, page 13.)

How has the city accomplished this? Liebman cited the importance of leadership. “One of the messages from charter schools is that decisions should be made by the people who are right there with the kids,” he said. “Principals need control over their staffs and budgets. They need per capita funding, so they get the resources they need.”

Yet merely decentralizing schools from a central bureaucracy—in an attempt to replicate the autonomy of charter schools—is not the answer, Liebman

argued. “Most decentralization is a disaster. You need accountability so there is an incentive on the part of schools to move kids forward. Empowerment and accountability are mutually reinforcing principles.” Principals should have the authority and resources to make decisions; in return, they must be held accountable for their students’ performance.

The city’s evaluation and empowerment resources include progress reports, quality reviews, periodic assessments, and ARIS, which Joel Klein discussed in his presentation. If tradeoffs must be made between a structure that motivates principals and teachers and one that perfectly measures performance, Liebman advocates focusing on motivation and comprehensibility. In the past, the city rewarded principals based on regression data, a method that principals could not replicate themselves. The city now uses calculations that principals can easily check. In the process of verifying their ratings, principals can learn what they must do to improve, Liebman said. “Very often principals will find that 30 kids in their school made the difference between the grade their school got and the grade above.”

He also stressed the importance of an accountability system in which teachers are evaluated by peers. “What you want to do, over the long term, is move from a system in which there’s a stark, external accountability system to one in which educators are holding each other

accountable.” To provide a way for teachers to examine each other’s work with the goal of improving student outcomes, New York City has created “inquiry teams” in each school, staffed by teachers and school leaders; currently, 60 percent of all teachers serve on at least one team.

All of these changes are designed to move from an environment of secrecy and blame to one of trust. Within the education system, there is a long tradition of factions assuming the worst about the motivations of others—whether teachers, administrators, politicians, or reformers—and this has stifled progress. “There’s been an effort to either hide or excuse failure,” Liebman said. “We need educators to be clear about the outcomes for students, and to feel responsible for those outcomes, child by child.”

John King, who is leading New York State’s reform efforts, complemented these views with his remarks, and emphasized the need to provide tools that enable teachers to become more effective. Like Liebman, he stressed the need for rapid change in the existing structure. “We are in great peril as a country of losing our prosperity because of how dysfunctional our school systems are. If anything, we understate the scope of our challenge.”

About 75 percent of first-year students in CUNY’s community colleges are in remedial courses—and many

of these are students who graduated with Regents diplomas. “Even the kids who are at the top of the spectrum aren’t being very well prepared.” King, who co-founded a charter school in Boston and was a managing director for a nonprofit organization that operates charter schools in New York and New Jersey, argued that improving teacher-student interaction is critical. “If we don’t drive change at the interaction between teacher and student, then we’re not getting at the essence of education.”

At the state level, reformers are working on three primary goals: raising current standards; creating new standards that measure college preparedness; and designing a curriculum with a clear scope, sequence and consistency. To achieve these objectives, the state needs more data on school performance. New York City is far ahead in collecting and analyzing data, King conceded. “With every educational reform we take, we ought to be building in how we’re going to evaluate whether or not it’s working.”

Equally important is improving teacher preparation, King said. “Out in the rural, western parts of the state, teachers are pretty much who they’re going to be.” A high churn rate—as suggested by some reformers—would not be manageable. “It’s unlikely that we’ll have a supply of newly minted teachers for those communities. So preparation matters a lot, and we need to do a much better job of it.” Teachers of education should know if their students go

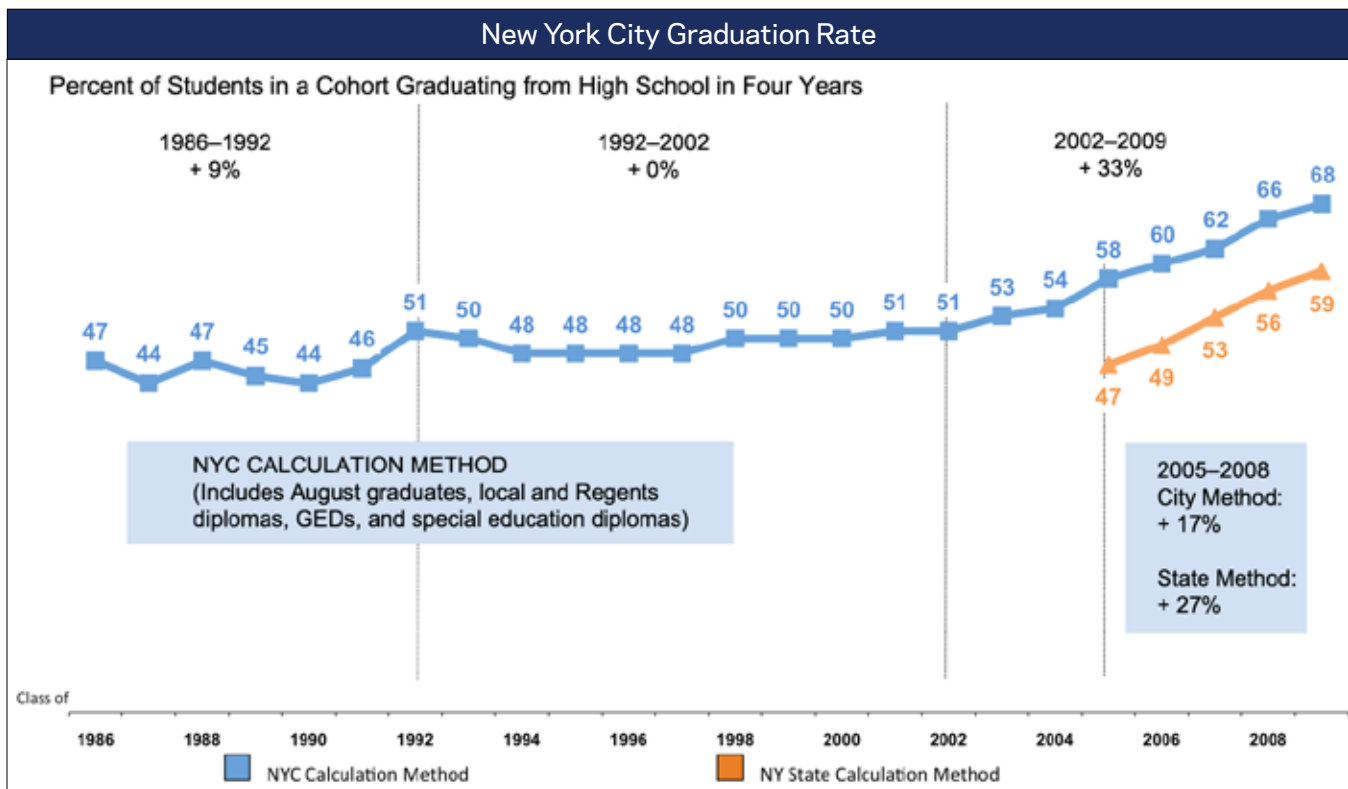


Figure 7: After remaining nearly flat for 16 years, the NYC graduation rate has increased by 33 percent since 2002. **Notes:** New York City’s traditional calculation includes local and Regents diplomas, GEDs, special education diplomas, and August graduates. It does not include disabled students in self-contained classrooms or District 75 students. The New York State calculation, used since 2005, includes Local and Regents Diplomas and all disabled students. It does not include GEDs and Special Education diplomas. *From the presentation of James Liebman, the Simon H. Rifkind Professor of Law at Columbia Law School and former chief accountability officer of New York City’s Department of Education.*



on to become effective teachers, King said. He called for expanding the use of data to make this information available. He also argued for allowing less traditional institutions to train teachers. "Why do we think that the institutions that have had a 130-year monopoly on teacher training are the only ones that can do it?" he asked. "Given their results, I think there's some question about that."

As part of its agenda for short-term change, the state plans to consider student achievement data when awarding permanent certification. "Teachers will not get a permanent license to teach unless they demonstrate value-added gains for the students they taught early in their careers," King said. "It's really taking the research that Doug Staiger and Jonah Rockoff were talking about and translating that into policy." Other initiatives include training principals how to evaluate teachers and moving to a four-point evaluation scale from the current two-point scale, in which nearly all teachers receive a positive rating.

But questions remain: What should happen when a teacher receives an unsatisfactory evaluation? Can teachers improve? "That's where research is so critical," King said. "There's a lot of work to do around professional development, teacher support, and figuring out how we can help teachers to get better. We don't have the answers now."

King concluded with a call for honesty: the public needs to recognize the extent of the problem, and reformers should acknowledge when questions cannot yet be answered. In these highly divisive debates, the focus should be on students, not on politics. In fact, this is a time to take advantage of resources. The federal government is investing more in education than it has before, and philanthropic contributions are soaring. "We have an extraordinary opportunity," King said. "We must not lose this moment."

 **Columbia Business School**

Social Enterprise Program
Uris Hall
3022 Broadway, Room 203
New York, NY 10027
T: 212-854-1060
F: 212-854-0016
Email: socialenterprise@gsb.columbia.edu
www.gsb.columbia.edu/socialenterprise

 **COLUMBIA UNIVERSITY**
IN THE CITY OF NEW YORK

