THE NARROWING GAP IN NEW YORK CITY TEACHER QUALIFICATIONS
AND ITS IMPLICATIONS FOR STUDENT ACHIEVEMENT IN HIGH-POVERTY
SCHOOLS

Donald Boyd
Hamilton Lankford
Susanna Loeb
Jonah Rockoff
James Wyckoff

The Narrowing Gap in New York City Teacher Qualifications and its Implications for Student
Achievement in High-Poverty Schools
Donald Boyd, Hamilton Lankford, Susanna Loeb, Jonah Rockoff, and James Wyckoff
NBER Working Paper No. 14021
May 2008
JEL No. I21,J24,J45

ABSTRACT

The gap between the qualifications of New York City teachers in high-poverty schools and low-poverty
schools has narrowed substantially since 2000. Most of this gap-narrowing resulted from changes
in the characteristics of newly hired teachers, and largely has been driven by the virtual elimination
of newly hired uncertified teachers coupled with an influx of teachers with strong academic backgrounds
in the Teaching Fellows program and Teach for America. The improvements in teacher qualifications,
especially among the poorest schools, appear to have resulted in improved student achievement. By
estimating the effect of teacher attributes using a value-added model, the analyses in this paper predict
that observable qualifications of teachers resulted in average improved achievement for students in
the poorest decile of schools of .03 standard deviations, about half the difference between being taught
by a first year teacher and a more experienced teacher. If limited to teachers who are in the first or
second year of teaching, where changes in qualifications are greatest, the gain equals two-thirds of
the first-year experience effect.

Donald Boyd
Rockefeller Institute of Government
411 State Street
Albany, NY 12203
boydd@rockinst.org

Hamilton Lankford
School of Education, ED 317
University at Albany
State University of New York
Albany, NY 12222
hamp@albany.edu

Susanna Loeb
524 CERAS, 520 Galvez Mall
Stanford University
Stanford, CA 94305
and NBER
sloeb@stanford.edu

Jonah Rockoff
Columbia University
Graduate School of Business
3022 Broadway #603
New York, NY 10027-6903
and NBER
jonah.rockoff@columbia.edu

James Wyckoff
Curry School of Education
University of Virginia
P.O. Box 400277
Charlottesville, VA 22904-4277
wyckoff@virginia.edu

I. Introduction

What is the distribution of educational resources across schools and what effect do disparities in resources have on the achievement of poor and minority students? This question dates to the Coleman Report (1966), but continues to be hotly debated, involving the courts as well as federal, state and local governments. Arguably the most important educational resource is teachers. Disparities in teacher qualifications figure prominently in most educational policy discussions and are a central feature of the No Child Left Behind Act of 2001 (NCLB) which requires a "highly qualified teacher" in every classroom in a core academic subject. Many states and large districts also have policies in place to attract teachers to difficult-to-staff schools (Loeb and Miller, 2006).

The recent interest in teacher labor markets stems in part from recognition of the importance of teachers and from the recognition of substantial differences across schools in the qualifications of teachers. A consistent finding in the research literature is that teachers are important for student learning and that there is great variation in effectiveness across teachers (Sanders and Rivers, 1996; Aaronson, Barrow and Sander, 2003; Rockoff, 2004; Rivkin, Hanushek and Kain, 2005; Kane, Rockoff and Staiger, 2006). Thus, understanding what makes an effective teacher as well as how teachers sort by their effectiveness across schools is central to understanding and addressing student achievement gaps.

Prior studies have found substantial sorting of teachers across schools with the schools with the highest proportions of poor, non-white, and low-scoring students having the least qualified teachers as measured by certification, exam performance, and inexperience (Lankford, Loeb and Wyckoff, 2002; Clotfelter, Ladd and Vigdor, 2006a). Yet, there have been substantial changes in the educational policy landscape over the past five years. New laws, including NCLB, have changed requirements for teachers. Assessment-based accountability policies at the state-level have created standards and increased oversight of schools, especially those with low-achieving students. New routes into teaching, many with fewer requirements before teaching, have lowered the cost for individuals to enter the teaching profession. These changes have affected teacher labor markets profoundly.

In this paper we examine these changes, asking how the distribution of teachers has changed in recent years and what the implications of these changes are for students. We examine three questions:

- How has the distribution of teaching qualifications between schools with concentrations of poor students and those with more affluent students changed over the last five years?

- What effects are the changes in observed teacher qualifications likely to have on student achievement?
- And, what implications do these findings have for improving policies and programs aimed at recruiting highly effective teachers?

We address these questions using data on New York City teachers, students, and schools. While the findings could be specific to New York City, they may mirror changes in other large urban districts, many of which have seen similar policy changes over the past decade.

We find that measurable characteristics of teachers are more equal across schools in 2005 than they were in 2000. Schools with large proportions of poor students and students of color, on average, have teachers whose observable qualifications are much stronger than they were five years ago. Nonetheless, a meaningful number of schools with large proportions of poor students did not demonstrate such improvement. We find that changes in these observed qualifications of teachers account for a modest improvement in the average achievement of students in the poorest schools. More importantly our results suggest that recruiting teachers with stronger observed qualifications, e.g., math SAT scores or certification status, could substantially improve student achievement.

II. Background

A growing literature finds that teachers "sort" very unequally across schools, with the least-experienced teachers and those with the poorest academic records often found in schools with the highest concentrations of low-income, low-performing and minority students (See, for example, Betts, Reuben and Danenberg, 2000; Lankford, Loeb, and Wyckoff, 2002; Bonesrønning, Falch, and Strøm 2005; Clotfelter, Ladd and Vigdor, 2006; and Peske and Haycock, 2006). Across several different states and at least one other country, low-performing, poor, and minority students systematically are taught by teachers with the weakest credentials, such as certification status and exam scores, SAT scores, ranking of undergraduate college, and, importantly, teaching experience. As but one example, Lankford, Loeb and Wyckoff (2002) find systematic sorting of New York State's elementary school teachers in 2000. Non-white students were four times more likely than white students to have a teacher who was not certified in any of the courses he or she taught and 50 percent more likely to have a teacher with no prior experience. The sorting of teacher qualifications within districts can also be substantial. In New York City elementary schools in 2000, non-white students were 40 percent more likely to have a teacher who was not certified in any of the courses she taught and 40 percent more likely to have a teacher with no prior experience. This sorting resulted from teachers' choices about whether and where to start a

teaching career, whether and where to remain in teaching - choices made within the constrained labor market governed by administrator choices, teacher contracts, and state and district regulations (For a more complete discussion of teacher sorting see Boyd, Lankford and Wyckoff 2007).

There is agreement that teachers can significantly influence student achievement (Sanders and Rivers, 1996; Aaronson, Barrow and Sander, 2003; Rockoff, 2004; Rivkin, Hanushek and Kain, 2005; Kane, Rockoff and Staiger, 2006). Sanders and Rivers (1996) estimate that differences in teacher quality can provide up to a 50 percentile improvement in student achievement and that these improvements are additive and cumulative over subsequent teachers. Kane, Rockoff and Staiger (2006) estimate that the difference in effectiveness between the top and bottom quartile of teachers results in a .33 standard deviation difference in student gains over the course of a school year.

While there is consensus that more effective teachers produce dramatically greater student achievement than less effective teachers, there is much less consensus on the attributes of teachers responsible for these differences. Much, though not all, of the recent research examining teacher effectiveness concludes that some teachers' attributes, such as higher test scores and greater teaching experience, will produce students with higher achievement. However, the effects of most teacher attributes appear small in comparison to the substantial variation across students in how much they learn in a year, as measured by test score gains. Studies of teachers' value-added to student achievement use state or district administrative data and thus are usually limited to assessing the effects of teacher characteristics collected by these entities. Teacher experience and certification are among the most studied.

Students of first year teachers learn less, on average, than students of more experienced teachers. While some of this difference may be driven by differential attrition of the worst teachers (Hanushek et al., 2005; Krieg, 2006; Goldhaber et al., 2007; Boyd et al., 2007), studies that account for the effects of compositional change find that first-year teachers produce student achievement gains that are significantly lower than otherwise similar teachers with ten to fifteen years of experience (Rockoff, 2004; Rivkin, Hanushek and Kain, 2005; Kane, Rockoff and Staiger, 2006). Most of these gains from experience occur within the first four years of teaching.

Many studies examine the effect of teacher education and certification on student achievement. These studies differ, sometimes substantially, in their findings. Many studies do not adequately account for the systematic differences between the schools in which certified teachers and uncertified teacher typically work. However, several recent studies with strong research designs and good data are able to address how various teacher qualifications affect

student achievement (Boyd, Grossman, Lankford, Loeb and Wyckoff, 2006; Clotfelter, Ladd and Vigdor, 2006b; Goldhaber, 2006; Harris and Sass, 2007; Kane, Rockoff and Staiger, 2006). In general these studies find that individual characteristics of teachers and their qualifications have relatively small effects. In many cases these effects are two to four percent of a standard deviation. While not large[1], this is about half as large as the typical gain from the first year of teacher experience.

The studies described above address the effects of specific teacher attributes. The effects in most cases appear to be modest. However, the variation in teacher attributes across schools is not independent; schools with the highest proportion of first year teachers also tend to have the highest proportion of uncertified teachers and the lowest prior academic performance of teachers. Teacher attributes vary together, and thus they should be taken together when considering the true difference in the effectiveness of teachers serving different student populations. In this paper, we assess the total effect of the differences in measurable characteristics of teachers across schools. We trace changes in the distribution of teachers across schools in New York City from 2000 to 2005 and estimate the effects that these changes are likely to have had on students in the traditionally most difficult-to-staff schools.

III. Data

The analysis is divided into two sections. The first section examines how the sorting of teacher qualifications across schools, categorized by poverty status and the racial-ethnic composition of students, has changed between 2000 and 2005. We then estimate how the changing composition of teacher qualifications affected student achievement gains. The analyses draw on a rich database constructed from administrative data from the New York City Department of Education, the New York State Education Department, alternatively certified teacher programs, and the College Board.

New York State gives statewide student exams in mathematics and English language arts in 4th and 8th grade. In addition, the New York City Department of Education tests 3rd, 5th, 6th and 7th graders in these subjects. All the exams are aligned to the New York State learning standards and

---

[1] Boyd and others (2008) show that when effects are measured relative to gains in student achievement net of test measurement error, effect sizes for these teacher variables, estimated with New York City data, are four times as large. So typical teacher characteristics have effect sizes of 8 to 16 percent of the standard deviation in student achievement gains.

each set of tests is scaled to reflect item difficulty and are equated across grades and over time.[2] Tests are given to all registered students with limited accommodations and exclusions. Thus, for nearly all students the tests provide a consistent assessment of achievement for a student from grade three through grade eight.

To analyze the relationship between teacher qualifications and student achievement, we create a student database with student exam scores, lagged scores and characteristics of students and their peers linked to their schools and to teachers and characteristics of those teachers. The student data, provided by the New York City Department of Education (NYCDOE), consists of a demographic data file and an exam data file for each year from 1998-99 through 2004-05. The demographic files include measures of gender, ethnicity, language spoken at home, free-lunch status, special-education status, number of absences, and number of suspensions for each student who was active in grades three through eight that year – approximately 450,000 to 500,000 students each year.

The exam files include, among other things, the year in which an exam was given, the grade level of the exam, and each student's scaled score on the exam. For most years, the file contains scores for approximately 65,000 to 80,000 students in each grade. The only significant exception is that the files contain no scores for 7th grade English language arts in 2002 because the New York City Department of Education is not confident that exam scores for that year and grade were measured in a manner that was comparable to the 7th grade English language arts exam in other years.

Using these data, we construct a student-level database where exam scores are normalized for each subject, grade and year to have a zero mean and unit standard deviation to accommodate any year-to-year or grade-to-grade anomalies in the exam scores. For this purpose, we consider a student to have value-added information in cases in which he/she has a score in a given subject (ELA or math) for the current year and a score for the same subject in the immediately preceding year for the immediately preceding grade. We do not include cases in which a student took a test for the same grade two years in a row, or where a student skipped a grade.

To enrich our data on teachers, we match New York City teachers to data from New York State Education Department (NYSED) databases and College Board databases, using a crosswalk file provided by NYCDOE that links their teacher file reference numbers to unique identifiers

---

[2] The mathematics exams in all grades are developed by CTB-McGraw Hill. New York State employs CTB-McGraw Hill for its 4th and 8th grade ELA exams. In 2003 New York City switched from CTB to Harcourt Brace for its 3rd, 5th-7th grade exams. At that time there was an equating study done to accommodate the switch in exams.

compatible with both databases. We draw variables for NYC teachers from these data files as follows:

- Teacher Experience: For teacher experience, we used transaction data from the NYCDOE Division of Human Resources payroll system to calculate experience in teaching positions in the New York City public school system.

- Teacher Demographics: We drew gender, ethnicity, and age from a combined analysis of all available data files, to choose most-common values for individuals.

- Undergraduate: We identified the institutions from which individual teachers earned their undergraduate degrees using the NYS Teacher Certification Database (TCERT) and combined it with the Barron's ranking of college selectivity to construct variables measuring the selectivity of the college from which each teacher graduated.

- Certification: We identified current certification areas from the NYS Teacher Certification Database (TCERT). New York State, like most states has several paths towards certification. Of particular note here, approved alternatively certified teacher programs in New York State receive a Transitional B teaching certificate. Thus while covered by this status teachers are certified.

- SAT scores: We obtained SAT scores for all individuals taking the SAT in New York State through 2002 from the College Board.

- Test performance: We drew information regarding the teacher certification exam scores of individual teachers and whether they passed on their first attempts from the NYS Teacher Certification Exam History File (EHF).

- Pathway: Initial pathway into teaching comes from an analysis of teacher certification applications plus separate data files for individuals who participated in Teach For America (TFA) and the Teaching Fellows Program, obtained directly from program officials.

- College Recommended: We obtained indicators for whether an individual had completed a college-recommended teacher preparation program and if so, the level of degree obtained (bachelor's or master's), from NYSED's program completers data files.

Finally, we match teachers and students to their schools, and incorporate data on those schools from the New York City Department of Education Annual School Report database, including:

- School-average performance on state and city standardized exams
- Poverty status as measured by the percentage of students eligible for free lunch
- Racial and ethnic breakdown of students
- Expenditures per pupil

The analysis of teacher sorting links teachers to schools and places schools into poverty groups based on the percentage of children eligible for free lunch in the first year a school appears in our database. We use a fixed school poverty group for each school so that it will not be influenced by year-to-year changes in reported free lunch percentages that sometimes appear spurious. In analysis that is not presented, we allowed the composition of quartiles to vary over

time as quartile boundaries and school poverty values change. These results are available from the authors. The results presented are not sensitive to this distinction. In defining groups, we weight each school by the number of teachers in our data, so that a school with many teachers will count more than a school with few teachers. The poverty groups are defined separately for elementary schools, middle schools, and high schools. In addition, for most of our analysis we only include schools present in both 2000 and 2005, so that the analysis will not be affected by changes in classifications of schools.[3] We take a similar approach for categorizing schools based on race and ethnicity.

IV. The Changing Distribution of Teacher Qualifications

The analysis below uses several indicators of teacher qualifications that researchers have previously employed to describe the teaching workforce. These measures include teaching experience, performance on state teacher certification exams, certification status and area, competitiveness of a teacher's undergraduate institution, pathway into teaching, and SAT scores. As discussed later, each of these measures appears likely to bear some relationship to student achievement, although the relationships are not always consistently significant or large in magnitude. We do not suggest that each of these measures, taken individually, has power to discriminate well between more and less effective teachers. However, taken as a group, we believe that they provide significant and substantial predictive power as well as useful insights, particularly because changes in teacher characteristics in the setting we study were driven by educational policies.

We analyze the distribution of teacher qualifications by the poverty status of students in the schools where these teachers work. There is substantial variation across the poverty groups in the percentage of students eligible for free lunch, as shown in the last row of Table 1. However, in New York City, even schools in the decile or quartile with the lowest percentage of free lunch-eligible students contain some students who are poor using this proxy. Thus, when we employ the terms affluent or rich in describing schools, this is a relative concept. By these measures, the distribution of teachers in 2000 was unequal. For example, Figure 1 shows that high-poverty schools were far more likely to have novice teachers: 25 percent of teachers in schools in the highest-poverty group (top 10 percent) were in their first two years of teaching, compared with 15 percent of teachers in the lowest-poverty group (bottom 10 percent). These patterns held across other available measures of teacher qualifications (Table 1). Teachers in the highest-poverty

---

[3] We also examine teacher sorting for all schools and with the exception of a somewhat larger narrowing of the gap in the percentage of novice teachers across poverty quartiles, the results are insensitive to this change. These results are available from the authors.

schools: had much lower scores on the SAT exams, were five times more likely to be uncertified, were much more likely to have graduated from the least-competitive colleges and had much lower scores on SAT exams and failed the Liberal Arts and Sciences Test (LAST), a state teacher certification exam that measures general knowledge, nearly three times as frequently as teachers in low-poverty schools.

The Narrowing Gap

Between 2000 and 2005 there was a remarkable narrowing in the gap in teacher qualifications between high-poverty schools and low-poverty schools. In particular, the high-poverty schools improved considerably while the low-poverty schools either did not improve or did so only modestly. As but one example, Figure 2 illustrates the narrowing of the gap with respect to math SAT scores. In 2000, teachers in the highest-poverty quartile of schools had math SAT scores that averaged just below 450, while those in the lowest-poverty quartile averaged 482, for a gap of 32 points. By 2005, math SAT scores for the teachers in the highest-poverty quartile rose to 472, while the lowest-poverty quartile rose to 488, so the gap narrowed by 16 points, or half its level five years earlier. Figure 3 shows a similar trend for teacher experience. In 2000, just over 25 percent of teachers in the highest-poverty quartile of schools had less than three years of experience, compared with slightly more than 17 percent in the lowest-poverty quartile of schools, for a gap of eight percentage points. By 2005, the gap had narrowed modestly as 22 percent of teachers in the highest-poverty schools were novices, reducing the gap to six percent points. Very similar changes in teacher qualifications occur when schools are categorized by the proportion of the school's students who are black or Hispanic, the proportion of students in the school who failed to reach proficiency on the state fourth grade math exam, and the proportion of students in the school who failed to reach proficiency on the state fourth grade reading exam. In each case, the initial gaps and the closing of the gaps in teacher qualifications from 2000 to 2005 follows the same pattern as shown for the schools arrayed by poverty. These results are shown in Appendix tables 1 through 3.

Table 2 shows that the same basic pattern held with every other measure of teacher qualifications, including the percentage of uncertified teachers, the percentage of teachers who failed the LAST teacher certification exam on their first attempt, and the percentage of teachers who attended least-competitive colleges. In general, the gap between the lowest and highest-poverty schools narrowed as a result of substantial improvements in the highest-poverty schools. Table 2 also shows expenditures per pupil and average teacher salaries (which are available for

2005 but not for 2000).[4] Expenditures per pupil were higher in high-poverty schools than in low-poverty schools in both years, and the difference actually increased between 2000 and 2005. Although total spending was higher in high-poverty schools, average teacher salaries are higher in the low-poverty schools. The differences in teacher salaries reflect the remaining difference in teacher experience between low and high poverty schools.

There are similar trends in teacher qualifications across schools by grade levels; however elementary schools experienced the greatest narrowing in the teacher qualifications gap. For example, as shown in Appendix Table 3a, the novice experience gap between high-poverty and low-poverty elementary schools in 2000 was 12 percentage points. By 2005 that had diminished to 5.6 percentage points. Similarly, the gaps in passing the LAST exam and SAT scores were reduced by 50 percent. Although middle schools also had a novice experience gap of over 11 percentage points in 2000, there was no meaningful reduction by 2005 (Appendix Table 3b). A much smaller percentage of middle school teachers failed the LAST exam initially than was the case for elementary school teachers and the middle school failure rate declined only modestly between 2000 and 2005. Finally, high schools experienced some meaningful improvement between 2000 and 2005 (Appendix Table 3c). On most measures the narrowing of the gap in qualifications fell between those of elementary schools and middle schools.

Not all poor schools experienced an improvement in teacher qualifications over this period. Three quarters of the schools in the poorest decile experienced an increase in average math SAT scores. However, the remaining 25 percent of the poorest schools experienced a decrease, although in many cases the decrease was small. Similar results hold for the other measures of teacher qualifications. Nonetheless, a substantial proportion of the high-poverty schools shared in the improved qualifications of teachers.

Explaining the Change

To further understand the recent change in teacher sorting it is worth asking to what extent the change is driven by new hires as opposed to the behaviors of more experienced teachers. Little of the change in teacher qualifications among poverty quartiles between 2000 and 2005 is attributable to the transfer and quit behavior of teachers. Figure 4 shows how the average math SAT scores for those teaching in 2000 change over time as that group moves across schools or leaves teaching in New York City. In 2000, the difference between the lowest and highest

---

[4] School-level expenditures include "direct services" such as classroom instruction, instructional support such as counseling and evaluation, school leadership and support, building services such as custodial services, and ancillary services such as food and transportation. They do not include expenditures on regional or system-wide services. On average, direct services amount to about 80 percent of total expenditures.

poverty quartiles of math SAT scores is 35 points. While both quartiles lose teachers with higher SAT scores, thus causing mean SAT scores in each to fall, the gap actually grows modestly to 38 points. Similar results hold for other measures of teacher qualifications. It is evident that the transfer and quit behavior of teachers had little to do with the reduced gap in teacher qualifications.

As illustrated by Figure 4, the reductions in the teacher-qualifications gap have been driven primarily by changes in the qualifications of newly hired teachers and the ways in which they vary with the poverty status of schools. Figure 5 shows that the average math SAT scores of newly hired teachers increased over the 2000-05 period, but that the increase in the poorest quartile of schools was so dramatic that by 2005, SAT scores were higher in these schools than in the lowest-poverty quartile. A similar convergence occurred for LAST failure rates and the percentage of uncertified teachers, but not for the competitiveness of colleges attended by teachers. One additional factor which may have also helped contribute to these changes is a considerable increase in the salaries of teachers in New York City, particularly for new teachers. The starting salary for a teacher with no experience and a bachelor's degree rose from $33,186 in 2000 to $39,000 in 2003. This salary schedule applies to teachers in all schools, regardless of poverty, and thus it is difficult to establish any direct link between salaries and the sorting of teachers. However, it is quite plausible that higher salaries for new teachers aided the recruitment and retention of Teaching Fellows and other highly qualified individuals choosing to teach in high-poverty schools.

Three policy changes were important in the improving and converging qualifications of new teachers. (1) In 1998 the New York State Board of Regents recommended abolishing temporary licenses for uncertified teachers effective September 1, 2003 and—except for limited waivers in New York City for 2004 and 2005—this was accomplished. (2) In 2000 the Regents created alternative certification routes that would allow school districts to hire teachers who are participating in approved alternative certification programs to become teachers as long as they were able to pass required teacher certification exams. (3) In collaboration with The New Teacher Project, the New York City Department of Education developed the Teaching Fellows program and in 2000 selected its first cohort of Fellows.

The shift in the entry pathway of teachers has had a large impact on the distribution of teacher qualifications for two reasons. First, Teaching Fellows and TFA teachers on average have test scores and prior academic experiences that are stronger than those of other teachers, and much stronger than those of the temporarily licensed teachers they replaced. For example, newly hired Teaching Fellows/TFA teachers in 2005 had average math SAT scores of 541, while newly

hired traditional teachers averaged 493. In 2002, the last year with a large number of newly hired temporarily licensed teachers, they averaged 460 on the math SAT, which is eighty percent of a standard deviation below that of the Teaching Fellows and TFA teachers who replaced them.[5] Second, newly hired Teaching Fellows and TFA teachers are placed disproportionately in high-poverty schools, as were their temporarily licensed predecessors. Teaching Fellows and TFA teachers grew from about three percent of newly hired teachers in the poorest quartile of schools in 2000 to 43 percent of all new teachers in those schools in 2005. Over the same period, temporarily licensed teachers fell from 63 percent of new hires to less than one percent of newly hired teachers in the poorest quartile of schools, while teachers entering through other pathways, including traditional teacher preparation programs increased from 34 percent to 56 percent. In sum, temporary license teachers were replaced by certified teachers, primarily alternatively certified teachers and to a lesser extent teachers from traditional teacher preparation programs and other routes.

These changes in the composition of teachers can be translated into improved teacher qualifications by decomposing the overall change in a qualification into the shares attributable to each pathway. For example, the 59 point improvement in average math SAT scores of entering teachers in the poorest quartile of schools between 2000 and 2005 can be decomposed into the portion attributable to changes in the proportion of teachers and changes in their math SAT scores attributable to each source of teachers (Teaching Fellows and TFA; Other pathways, primarily traditional teacher preparation, and temporary license teachers). Based on this decomposition, 65 percent of the 59 point increase in math SAT scores of newly hired teachers in the poorest quartile of schools between 2000 and 2005 is attributable to the increased share and scores of Teaching Fellows and TFA teachers; 35 percent is attributable to the increased share and scores of teachers from other pathways, primarily traditional teacher preparation programs. Thus, the increase in Fellows and TFA teachers accounts for the majority of improvement in math SAT scores over the period but teachers from other pathways also contributed in meaningful ways. Similar results occur for other measures of qualifications with the exception of experience. As a result, the policy changes that led to the substitution of the Teaching Fellows and TFA teachers for temporary licensed teachers resulted in improved measures of teacher qualifications, with the exception of experience, in the poorest quartile of schools.

---

[5] Even more dramatic differences exist between Teaching Fellows/TFA teachers and temporarily licensed teachers on other measures of qualifications, including undergraduate college ranking, certification status and the percent who failed the LAST exam on the first attempt.

The reduction in the proportion of novice teachers is not attributable to the policy changes noted. The dramatic increase in Teaching Fellows over the 2000 to 2005 period likely increased the proportion of novice teachers, as these teachers typically have no prior teaching experience. Rather, the reduction in the proportion of novice teachers appears to have resulted from a reduction in hiring of new teachers over this period. As we describe below, the vast majority of improved student achievement resulting from measured teacher qualifications results from qualifications other than experience.

V. The Relationship between Teacher Qualifications and Student Achievement

Over the same period in which the gap in teacher qualifications narrowed, the gap in the proportion of students failing to meet proficiency standards also narrowed. In the 2000 school year, 30 percent of students in the lowest-poverty group failed to meet state proficiency standards on the grade 4 ELA exam, while 74 percent of students in the highest-poverty schools failed to meet the state standard, for a gap of 44 percentage points. Between 2000 and 2005, failure rates declined in all poverty groups, as shown in Table 3, but they declined most in the highest-poverty schools so that the gap between low and high-poverty groups narrowed to 32 points.

Table 3 shows that this narrowing of the percentage of students reaching proficiency between high and low-poverty schools occurred across all four major state exams – ELA for grade 4 and 8, and math for grade 4 and 8, although middle school tests showed only a slight closing of the gap. We also have examined other measures of the achievement gap, including average test scores by school, and the percentage of a school's students scoring at Level 4 (the highest level of performance). By all measures except the Level 4 percentage for 8th grade ELA, the achievement gap between high-poverty and low-poverty schools narrowed between 2000 and 2005. In general, achievement in high-poverty schools has improved and come closer to that of low-poverty schools, though in some cases the changes were not large.

Even though the narrowing of student achievement across poverty groupings of schools occurred concurrently with the narrowing of the teacher-qualifications gap across these groupings, the causal relationship between the two trends is not clear. The change in teachers may have caused the change in student outcomes; the change in student outcomes may have caused the change in teachers; a third factor may have led to both changes; or, alternatively, they may have separate though simultaneous causes. Whether teacher qualifications played a role in this narrowing is an open question. While we can not determine the complete causal mechanism, we can predict how much a change in measurable characteristics would, on average, affect student outcomes. The prediction may under or over estimate the effects of the changes in teacher sorting on student achievement depending on how unmeasured characteristics of teachers

changed during this same time period. If the extent of teacher sorting declined with respect to positive unmeasured, as well as measured, characteristics then the estimates will underestimate the teacher effects; if teacher sorting increased on positive unmeasured characteristics, then we will overestimate the total teacher effect.

Estimating the Effects of Measured Teacher Characteristics

It is not easy to estimate how the achievement gains of students are affected by the qualifications of their teachers because teachers are not randomly sorted into classrooms. For example, if teachers in schools in which students perform best in math are more likely to be certified in math, one might be tempted to conclude that being certified to teach math contributes to higher student achievement. The causal relationship, however, may operate in the other direction; that is, more qualified teachers may be in schools where students perform well in math because they prefer to teach good students and because employers want to staff their courses with in-field certified teachers. Analysts need to be careful not to attribute the test-score gains associated with sorting to the attributes of teachers. Unfortunately, there is not a specific agreed-upon methodology for answering this question in a non-experimental framework. Because of this, we choose to run a number of different specifications in order to test the robustness of the estimated effects.

Equation 1 summarizes our base model for estimating the effects of teacher attributes.

$$A_{isgty} - A_{is'g(g-1)t'(y-1)} = \gamma_0 + \gamma_1 S_{iy} + \gamma_3 C_{ty} + \gamma_4 T_{ty} + \pi_i + \pi_g + \pi_y + \varepsilon_{isgty} \quad (1)$$

Here the standardized achievement gain score of student *i* in school *s* in grade *g* with teacher *t* in year *y* is a linear function of time varying characteristics of the student *S*, characteristics of the other students in the same grade with the same teacher in that year *C*, and the teacher's qualifications *T*. The model also includes student, grade and time fixed effects and a random error term. The time-varying student characteristic is whether the student changed schools between years. Class variables include proportion of students who are black or Hispanic, the proportion who receive free or reduced price school lunch, the class size, the average number of student absences in the prior year, the average number of student suspensions in the prior year, the average achievement scores of students in the prior year, and the standard deviation of student test scores in the prior year. Teaching experience is measured by separate dummy variables for each year of teaching experience up to a category of 21 and more years. Other teacher qualifications include whether the teacher passed the general knowledge portion of the certification exam on the first attempt, certification test scores, whether and in what area the

teacher was certified, the Barrons ranking of the teacher's undergraduate college, math and verbal SAT scores[6], the initial path through which the teacher entered teaching, e.g., a traditional college recommended program or the New York City Teaching Fellows program, and an interaction term of the teacher's certification exam score and the portion of the class eligible for free lunch. The standard errors are clustered at the teacher level to account for multiple student observations per teacher. We also estimate the model with student achievement level as the dependent variable, the previous year's achievement and its square as independent variables along with all other independent variables and a school fixed effect, omitting the student fixed effect, and obtain results that are remarkably similar to those presented for student fixed effects. The effect of employing this model in assessing the effect of teacher observables on student achievement is presented below; a full set of coefficient estimates is available from the authors.

Student achievement gains are measured as the difference between the student's test score in a given year and his or her test score in the prior year. Student achievement gains are computed after normalizing test scores to have zero mean and unit standard deviation for each year and grade. Based on the differential pattern of teacher sorting between elementary and middle schools described above and earlier research that finds differences in the determinants of student achievement across grade levels (Boyd et al. 2006), we estimate four models: separate models for math and ELA, and separate models for students in 4th or 5th grades and those in 6th through 8th. We discuss only the math results; the effect of observed qualifications on student achievement in ELA in both grade groupings is very small.

Many of the measures of teachers' qualifications are highly correlated with each other in our sample. The LAST certification exam score and the verbal SAT are correlated at 0.68; attending a most competitive undergraduate college is correlated with the verbal SAT at 0.35; and

---

[6] We impute values for SAT scores and the LAST certification exam for all teachers with missing values. We observe SAT's for every person who took the SAT in New York from 1980 until 2000. Thus we may be missing SAT scores for three groups: those who took the SAT prior to 1980 and thus are likely to be more experienced teachers; those who took the SAT in another state, and those who never took the SAT. We do not observe SAT scores for about 53 percent of the teachers in our sample. Two-thirds of the teachers for whom we are missing SAT scores were born prior to 1963 and thus were younger than 17 in 1980, when our SAT data begin. Our imputations are guided by a growing literature (see for example Cameron and Travidi, 2005). Consistent with this literature, we employ a model based approach to imputing SAT and LAST scores for missing observations. As shown in our results presented below, we have examined several alternative models to explore the robustness of our results to the imputation of SAT and LAST scores.

Finally, New York State switched teacher certification exams from the Educational Testing Service (ETS) general knowledge exam to an exam designed for New York State by National Evaluation Systems (NES) in 1995. Because our sample includes teachers who took the ETS exam, we create a dummy variable that indicates if a teacher passed either exam the first time they took it. In addition, we impute values of the LAST for those who did not take it.

certification to teach math and entering teaching through the New York City Teaching Fellows program have a correlation of 0.30. As a result, including them all in one large regression may understate the importance of individual qualifications in affecting student achievement. For example, as shown in Table 4, while teacher experience is statistically significant and appears important, few of the other measures of teacher qualifications are significant, even though if entered alone they would have been.

The gains to teacher experience can serve as a benchmark against which to judge the effect size of other teacher qualifications. As discussed above, the coefficient estimates for experience in Table 4 may provide misleading estimates of the gains that accrue to teacher experience. These results are a combination of teacher improvement with experience and teacher attrition. Figure 7 shows the gains to experience for math achievement in a model that employs teacher fixed effects and thus increments to value added are identified only from teachers who persist from one year to the next. As shown, teachers continue to improve the achievement outcomes of their students over the first 3 to 5 years of their careers. The effect of moving from being completely inexperienced to having a full year of experience is the largest gain and in our sample of 4th and 5th grade math achievement is about 0.06 standard deviations.

Other measures of teacher qualifications also are related to student achievement gains. Not being certified at the time a teacher taught the course reduces student achievement by 0.042— roughly two-thirds the size of the gain of the first-year of teaching experience, which most observers agree is important. A similar size effect results from improving math SAT scores by one standard deviation which improves student achievement by 0.043. Having a teacher who attended a competitive undergraduate college improves performance relative to one who attended a less competitive college, but the effect appears small (.014).

The Combined Effect of Teacher Characteristics

Although some of the individual qualifications described above affect student outcomes in important ways, often the effects are relatively small in magnitude when compared with the variation in student learning over a school year. However, the rather substantial changes in teacher qualifications in the poorest schools during the 2000 to 2005 period occurred across a variety of measures. The effects of these joint changes are likely to be greater than changes in a single measure holding other attributes constant. In order to estimate the combined effect of the change, we use the coefficient estimates for the teacher variables presented in Table 4 and the actual qualifications of teachers in the poorest and most affluent deciles of schools in 2001 and 2005 to predict the student achievement gains attributable solely to changes in teacher

qualifications. To insure stability in the predictions, we employed records reflecting the qualifications of teachers working in 2000 and 2001 (labeled 2001) and the teachers working in 2004 and 2005 (labeled 2005).

As shown in Figure 8, the improvement in qualifications increased predicted student achievement in the poorest decile, shifting the overall distribution to the right between 2001 and 2005. On average, the change in the observed qualifications of teachers employed in our value-added model increased student achievement by 0.029 standard deviations in the decile of schools with the highest concentration of students in poverty. We estimated the improved value-added of teachers in low and high-poverty schools in 2001 and 2005 by predicting the value-added of each student resulting from all of teacher variables in Table 4, holding constant all other variables. We then averaged across all students within low and high-poverty schools for both years. The predicted student gains in the most affluent decile of schools improved by 0.007. Therefore, as a result only of the changes in observed teacher qualifications, the gap in gains resulting from observed teacher qualifications between the poorest and richest deciles declined by 0.022, from .089 to .067 (Table 5). Said differently, improvements in the measured teacher qualifications in the poorest decile of schools reduced the gap resulting from observed differences by 25 percent.

As noted above, the change in teacher sorting has been driven almost exclusively by new teachers. Many teachers in a school remain unchanged over any five year period and thus when examining the effect of changes in teacher qualifications, these observations do not contribute to improved student achievement (except for the net gains to experience). The prior analyses predict student achievement based on the full sample of teachers. The results are predictably stronger if we look only at teachers in their first or second year of teaching. As shown in the second column of Table 5, achievement predicted only from the observable qualifications of first and second-year teachers in the poorest decile of schools improves by 0.044 from 2001 to 2005— about two thirds of the gain estimated to accrue to teachers after their first year of teaching. The gap in student achievement between poor and more affluent schools was reduced by 0.041.

The reduction in the gap in achievement gains resulting from improved teacher qualifications is robust to several alternative specifications. As shown in Table 5, if instead of imputing the SAT and LAST exams, we drop the math and verbal SAT variables and omit observations that are missing the LAST, the poorest decile shows greater improvement and the gap closes by more than in our base model. If instead, we include the SAT variables and omit observations with missing values, gains to the poorest decile and the closing of the gap are much greater. Finally we estimate a model similar to the Base model that employs current achievement levels as the dependent variable with lagged student achievement and school fixed effects instead

of a gain model with student fixed effects. In these estimates the gap closes by 0.029. These results suggest that our findings are robust to several different assumptions regarding the estimation of the effects of teacher attributes on student value-added.

One way of summarizing these results is to examine what portion of the original achievement gap resulting from observed teacher qualifications between the most affluent and poorest deciles each model would predict is eliminated as a result of improved teacher qualifications. As shown in the last row of Table 5, across four quite different specifications, the percentage of gap reduction attributable solely to observed teacher qualification varies between 20 and 28 percent, with the base model predicting 25 percent. Thus, the changes in teacher qualifications alone that occurred in New York City's poorest schools between 2000 and 2005 had a meaningful effect on 4th and 5th grade math achievement. These predicted effects include the effect of the reduction in the teacher experience gap. If that effect were held constant, there would still be a narrowing of the gap in student achievement gains of .018 in the Base model, as shown in the last column of Table 5. Thus, about 80 percent of the reduction in the original gap between schools with poor and more affluent students is attributable to qualifications other than experience. If unobserved measures of teacher qualifications, such as motivation, are systematically correlated with the observed measures, this would contribute to the effects we document. From a recruitment perspective, however, the end result for improved student achievement is not altered.

In addition to explaining a moderate proportion of the change in achievement across schools, the results show that there is a substantial difference between the teachers in predicted student achievement gains based solely on observable qualifications. As is apparent in any of the achievement distributions in Figure 8, there are meaningful achievement differences between higher and lower performing teachers solely attributable to observed teacher qualifications.

Consider only 4th and 5th grade teachers whose students are in the quartile of schools with the highest rates of student poverty. The difference between the average value-added attributable solely to teacher qualifications for those teachers in the top and bottom quintiles of this distribution is 0.16—roughly three times the effect of the gains attributable to the first year of teacher experience. Table 6a shows how these values change over the quintiles of value-added for teachers in the poorest quartile of schools. It also shows the average qualifications of teachers in each of these quintiles. There are important differences in qualifications between teachers who produce the highest and lowest value added students, even among teachers working in poorest quartile of schools. Those with the weakest value added tend to be inexperienced, have failed the LAST certification exam the first time taken, be uncertified at the time they teach the class, and

17

have low math SAT scores. As might be expected, differential experience plays a role in accounting for the differences in student value-added. However, as shown in Table 6b, when we omit experience from the prediction (assign all students a novice teacher) there remains an 11 percent of a standard deviation difference in achievement gains between the top and bottom quintiles—about twice the size of the gains associated with the first year of teaching experience. As is shown in Table 6b, the actual teachers of these students have substantially different qualifications—e.g., differences in teacher certification status of about 70 percentage points and math SAT scores that differ by more than 150 points.

The conclusion arising from this analysis is clear. *The performance of students in 4th and 5th grade math can be substantially increased across all stratifications of students by recruiting and hiring better qualified teachers.*

The effects of observed teacher qualifications on student achievement are more modest for middle school math. Figure 9 shows the how the narrowing of differences in teacher qualifications from 2001 to 2005 corresponds to improvement in student achievement of 0.015 for the poorest decile, but to virtually no change in the gap between the poorest and the most affluent deciles. If limited to only teachers in their first or second year of teaching, the poorest decile improves by 0.020 standard deviations. The smaller effects for middle school achievement are fully consistent with the smaller changes in teacher qualifications noted above and in Appendix Table 3b. Nonetheless, there are meaningful within-decile differences in the predicted effects of observed teacher qualifications of the least and most effective teachers, and thus, again, recruiting more qualified teachers could meaningfully improve achievement outcomes.

VI. Conclusions

The gap between the qualifications of New York City teachers in high-poverty schools and low-poverty schools has narrowed substantially since 2000. For example, in 2000 teachers in the highest-poverty decile of schools had math SAT scores that on average were 43 points lower than their counterparts in the lowest poverty decile of schools. By 2005 this gap had narrowed to 23 points. The same general pattern held for other teacher qualifications such as the failure-rate on the Liberal Arts and Sciences (LAST) teacher certification exam, the percentage of teachers who attended a "least competitive" undergraduate college, and verbal SAT scores. Most of this gap-narrowing resulted from changes in the characteristics of newly hired teachers, rather than from differences in quit and transfers rates between high and low-poverty schools.

The gap-narrowing associated with new hires has been largely driven by the virtual elimination of newly hired uncertified teachers coupled with an influx of teachers with strong

academic backgrounds from alternative certification programs and to a lesser extent traditional teacher preparation programs. Only five percent of newly hired Teaching Fellows and TFA teachers in 2003 failed the LAST exam on their first attempt, while 16.2 percent of newly hired traditional teachers failed the LAST exam, and fully 32.5 percent of uncertified teachers failed the LAST exam. In 2005, 43 percent of all new teachers in the quartile of schools with the poorest students were Teaching Fellows or TFA teachers.

The improvements in teacher qualifications, especially among the poorest schools, appear to have resulted in improved student achievement. By estimating the effect of teacher attributes using a value-added model, the analyses above predict that observable qualifications of teachers resulted in average improved achievement for students in the poorest decile of schools of 0.03 standard deviations, about half the difference between being taught by a first year teacher and a more experienced teacher. If limited to teachers who are in the first or second year of teaching, where changes in qualifications are greatest, the gain equals two-thirds of the first year experience effect.

Many of these changes resulted from policy interventions that changed the qualifications of the teachers of poor, minority and low achieving students in New York City. In particular, most of the changes, other than the reduced proportion of novice teachers, can be attributed to the New York State policy that eliminated uncertified teachers and the New York City policy that established the Teaching Fellows program and, to a lesser extent, employed Teach for America teachers. The sorting of the least qualified teachers to the students most in need of better teachers is not destiny, but it requires forceful action by policy makers and a commitment by local hiring authorities to attract more highly qualified teachers.

 Perhaps most intriguing, much larger gains could result if teachers with strong teacher qualifications could be recruited. Among teachers teaching 4th and 5th grade math students in schools with the highest proportions of students in poverty, we found there are substantial differences in student achievement solely attributable to differences in observed teacher qualifications. The top quintile has value-added that differs from the bottom quintile by an effect size of 0.11, about three times the effect accruing to the first year of experience. Thus, recruitment can substantially change outcomes for students.

A new paper by Boyd, Grossman, Lankford, Loeb and Wyckoff (2008) shows that effect sizes as typically measured, including those in this paper, understate how teacher attributes and other factors affect actual gains in student achievement. Judging such effects relative to the dispersion in achievement gains, not the dispersion of achievement, and netting out that portion of the dispersion in test score gains attributable to measurement error result in effect sizes being

larger by a factor of four. This has important implications for the results presented in this paper. Rather than having an effect size of 0.03, as reported above, the effect of observed teacher qualifications on the true gain in achievement of students is 12 percent of a standard deviation. Similarly, the potential improvement of recruiting more qualified teachers is more than 40 percent of a standard deviation, net of the role of experience.

Improving student achievement, especially among students in low-performing schools will likely result from several complementary strategies. A large proportion of the variation in teacher effectiveness in improving student achievement is not related to measurable teacher characteristics such as test scores or certification. Because of this, policies that enable school leaders to better understand the strengths and weaknesses of each teacher so that they can target professional development and effectively utilize the due-process system to continually improve the teacher workforce are likely to be important. However, this paper suggests that selection of teachers with stronger qualifications has made an important difference in New York City public schools and that recruitment and retention of teachers with stronger measurable characteristics can lead to improved student learning.

Figure 1: Percentage of New York City Teachers With Less than 3 Years of Experience, By Poverty Grouping of School's Students, 2000
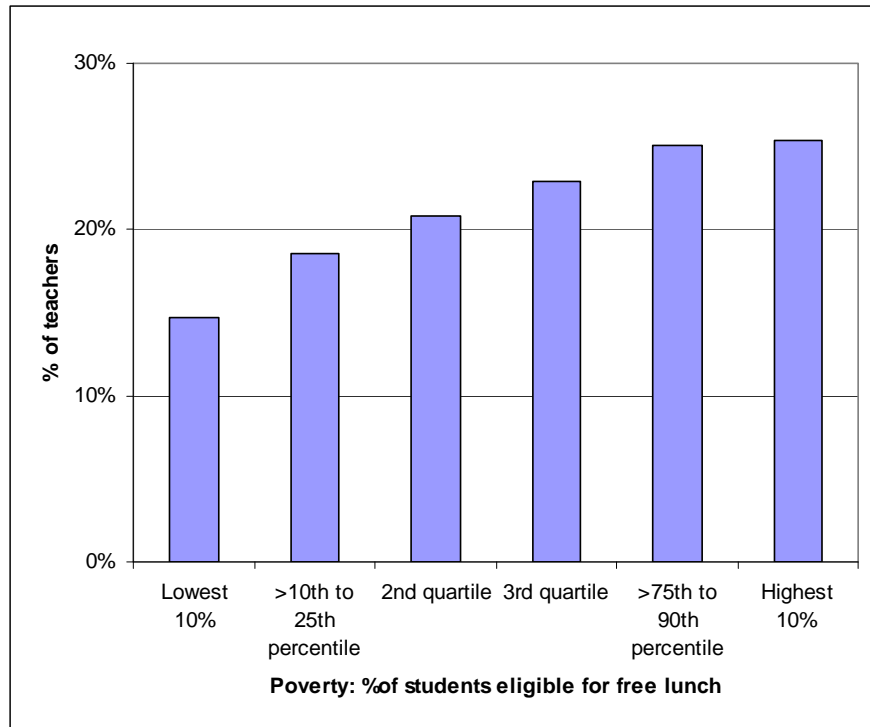
Figure 2: Average Math SAT Scores All New York City Teachers by Poverty Quartile of School's Students, 2000-2005
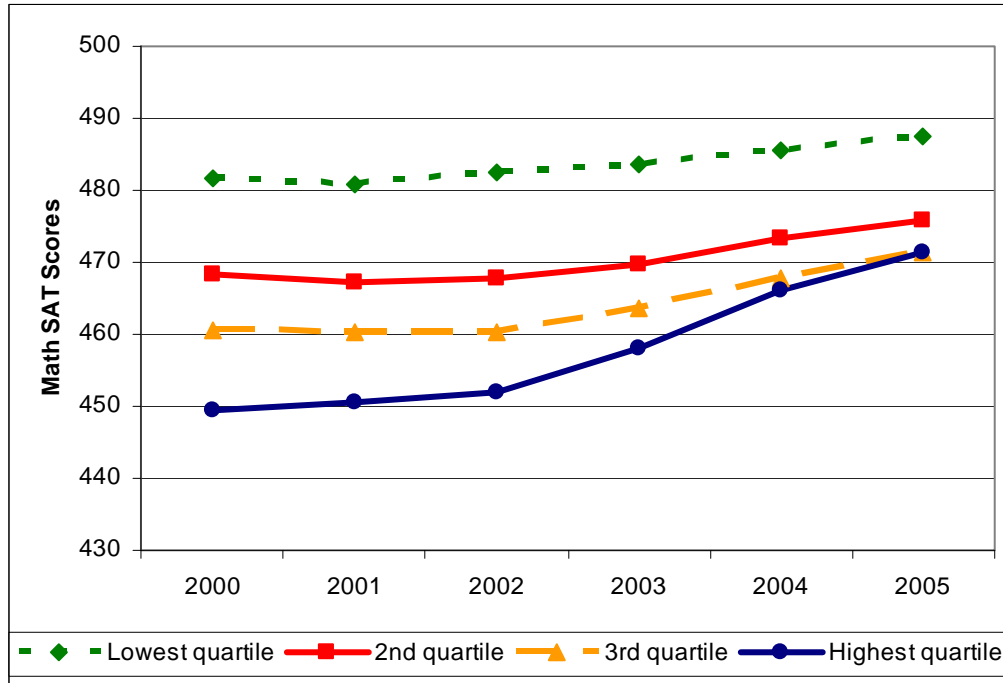
Figure 3: Percent of All New York City Teachers Who are Novices by Poverty Quartile of School's Students, 2000-2005
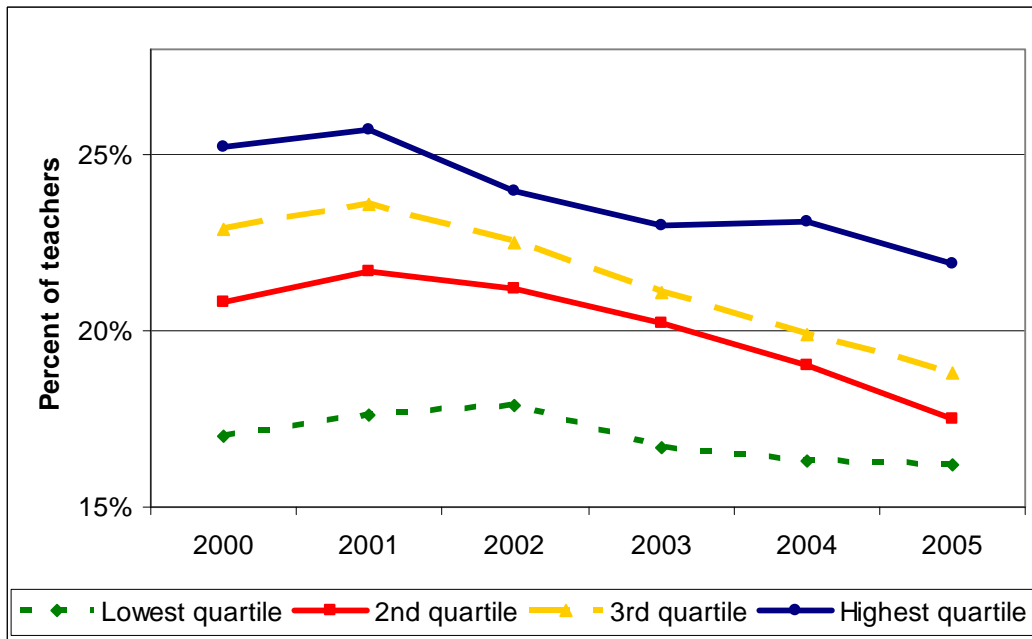
Figure 4: Average Math SAT Scores of Those Teaching in 2000 and the Effect of Their Transfers and Quits Over Time
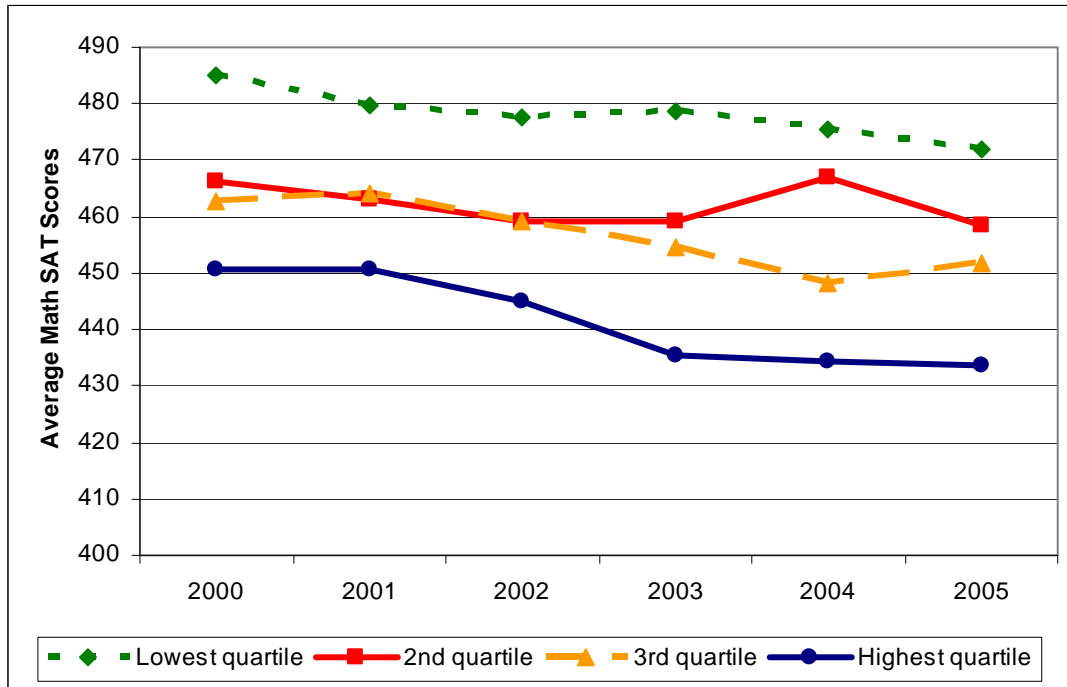
Figure 5: Average Math SAT Scores of New Teachers by Poverty Quartile of School's Students, 2000-2005
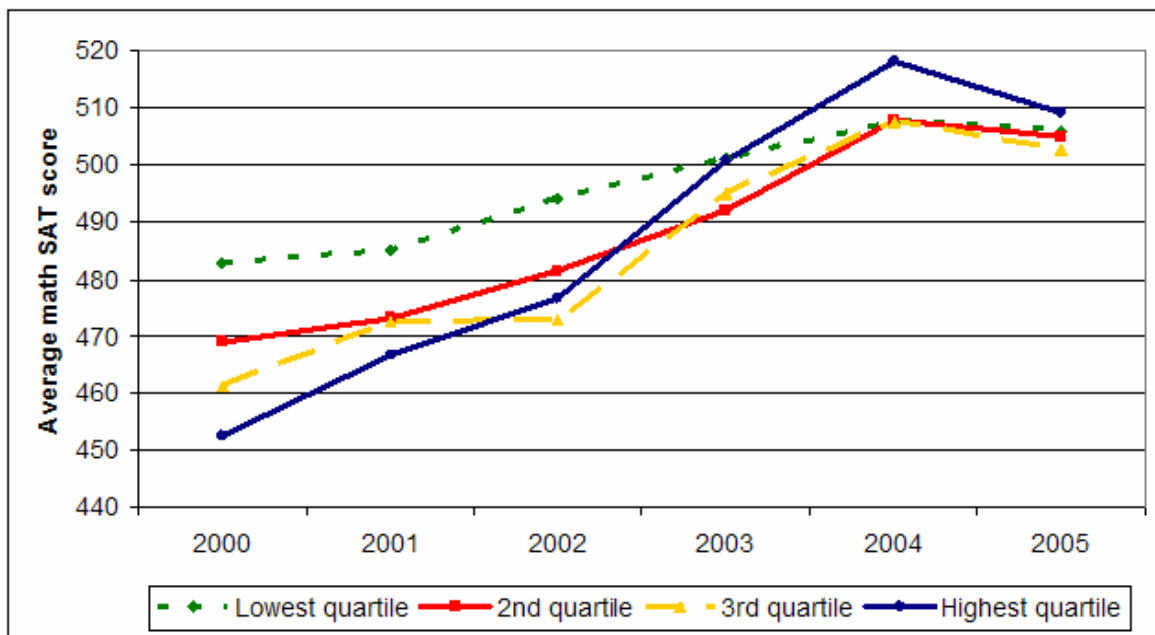
Figure 6: Number of New Teachers by Pathway, 2000-2005
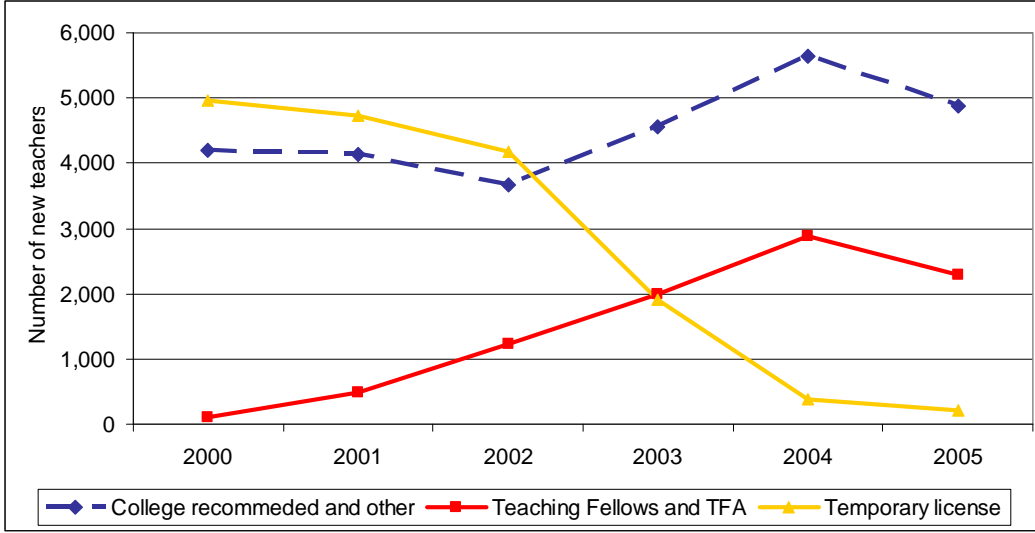
Figure 7: Improvements in Math Student Achievement Attributable to Additional Teacher Experience
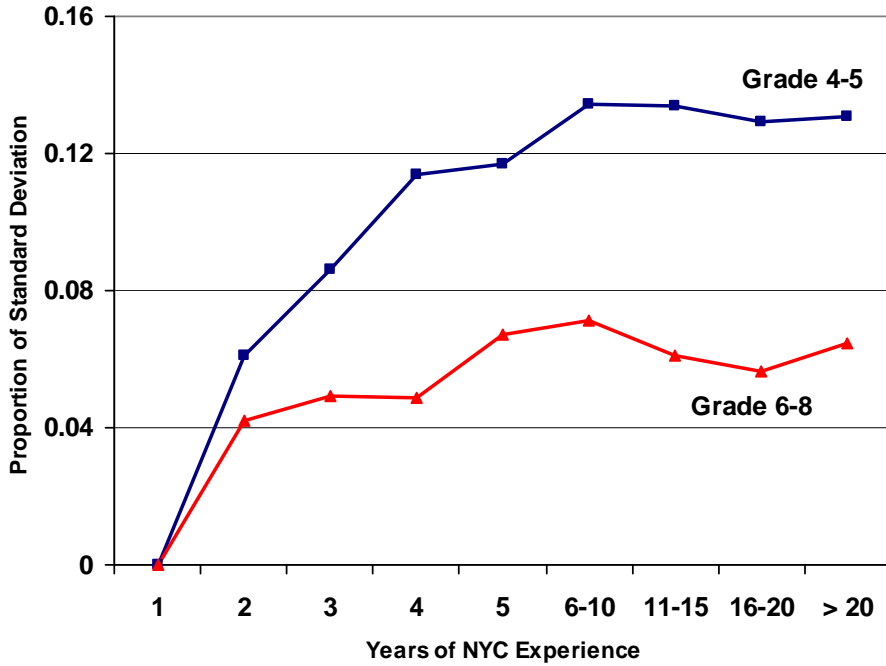
Figure 8: Effect of All Observed Teacher Qualifications on Students in Grades 4 & 5 Math Achievement, Most Affluent and Poorest Deciles of Schools, 2001 and 2005
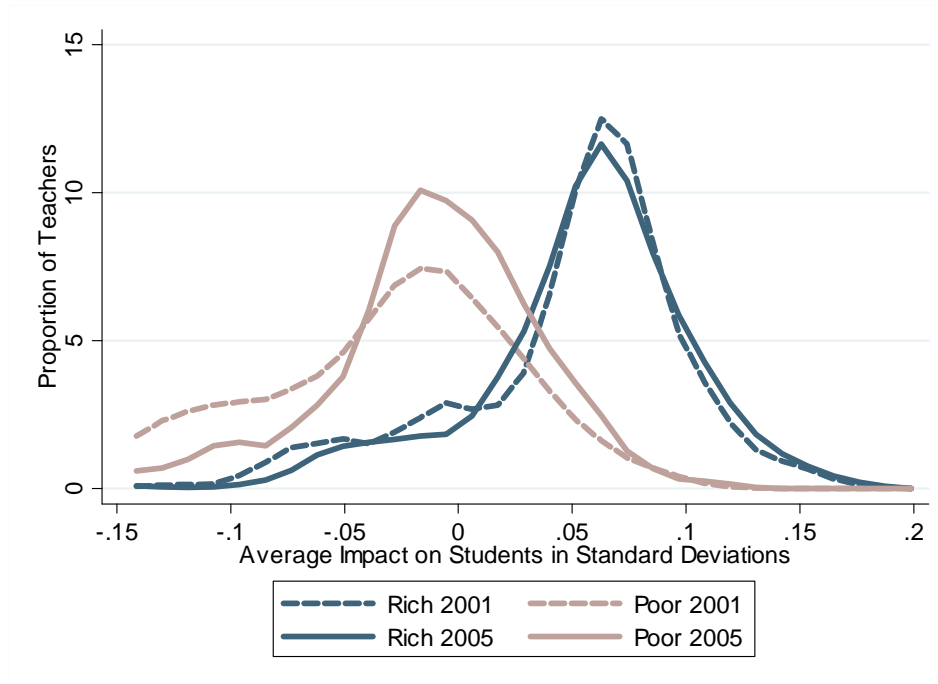
Figure 9: Effect of All Observed Teacher Qualifications on Students in Grades 6 - 8 Math Achievement, Most Affluent and Poorest Deciles of Schools, 2001 and 2005
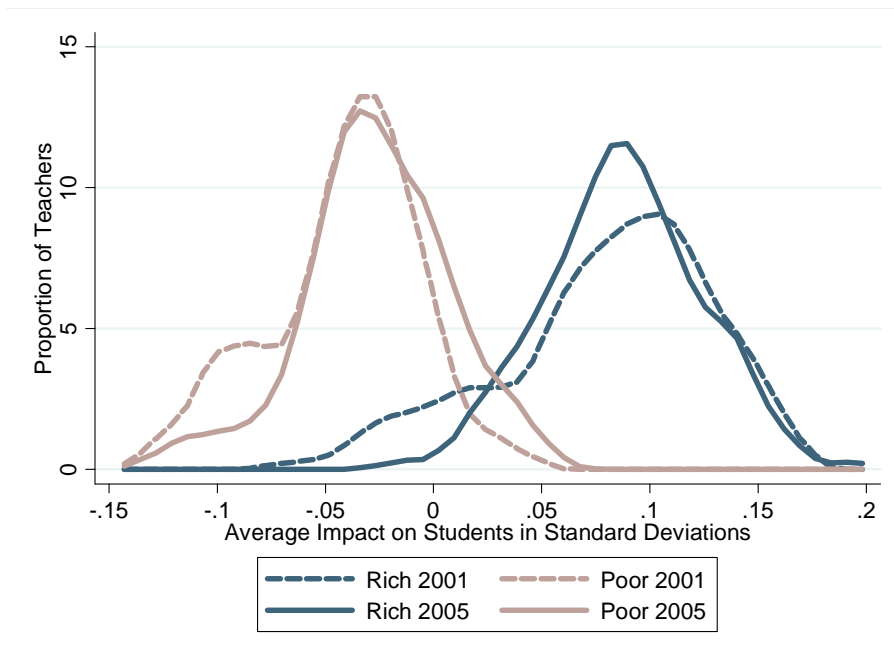
Table 1
Qualifications of Teachers by Poverty Status of Schools in Which They Taught in 2000

	Lowest 10%	>10th to 25th percentile	2nd quartile	3rd quartile	>75th to 90th percentile	Highest 10%
% with less than 3 years of NYC teaching experience	14.7%	18.6%	20.8%	22.9%	25.1%	25.4%
SAT math score	490	477	468	461	451	447
SAT verbal score	506	487	481	472	465	461
% who failed LAST exam on first attempt	12.20%	16.80%	23.50%	29.60%	35.30%	34.20%
% Not certified to teach	4.00%	8.20%	11.50%	17.00%	21.00%	21.90%
% who attended least competitive undergraduate institutions	23.50%	22.90%	23.50%	25.30%	27.50%	27.40%
Expenditures per pupil*	$8,002	$8,335	$8,338	$8,738	$9,093	$9,479
% Eligible for free lunch	21.6	50.4	67.6	81.6	90.5	96.3

Table 2: Average School Qualifications of Teachers by Student Poverty, 2000 and 2005

	2000			2005			Change from 2000 to 2005		
	Lowest 10%	Highest 10%	Gap: Highest 10% - Lowest 10%	Lowest 10%	Highest 10%	Gap: Highest 10% - Lowest 10%	Lowest 10%	Highest 10%	Change in Gap
% with less than 3 years of NYC teaching experience	14.7%	25.4%	10.7%	15.1%	21.7%	6.6%	0.4%	-3.7%	-4.1%
SAT math score	490	447	-43	495	471	-23	5	24	19
SAT verbal score	506	461	-45	503	485	-18	-3	23	26
% who failed LAST exam on first attempt	12.2%	34.2%	22.0%	13.4%	24.7%	11.3%	1.2%	-9.5%	-10.7%
% Not certified to teach	4.0%	21.9%	17.9%	1.5%	3.3%	1.8%	-2.5%	-18.6%	-16.1%
% who attended least competitive undergraduate institutions	23.5%	27.4%	3.9%	26.7%	24.3%	-2.4%	3.2%	-3.1%	-6.3%
Number of Teacher Absences	na	na		10.0	10.8	0.7	na	na	na
Expenditures per pupil*	$8,002	$9,479	$1,477	$9,711	$11,866	$2,155	$1,709	$2,387	$677
Teacher Salary	na	na		$59,314	$53,830	-$5,484	na	na	na

* All 2000 dollars adjusted to 2005 school-year dollars using CPI.

Table 3: Percentage of New York City Students Failing to Meet Proficiency on Achievement Exams by Test and Poverty Decile, 2000 and 2005

	2000			2005			Change from 2000 to 2005		
	Lowest 10%	Highest 10%	Gap: High 10% - lowest 10%	Lowest 10%	Highest 10%	Gap: High 10% - lowest 10%	Lowest 10%	Highest 10%	Change in gap
Percent failing to meet proficiency									
ELA grade 4	29.6	73.7	44.2	18.1	50.5	32.4	-11.5	-23.2	-11.8
Math grade 4	24.3	71.1	46.8	7.7	29.2	21.5	-16.6	-41.8	-25.2
ELA grade 8	37.5	78.4	40.9	41.3	76.2	35.0	3.7	-2.2	-5.9
Math grade 8	51.9	85.6	33.7	38.9	69.4	30.5	-13.1	-16.2	-3.2
Percent achieving highest level									
ELA grade 4	25.6	2.8	-22.8	32.0	8.1	-23.9	6.4	5.4	-1.1
Math grade 4	26.8	2.5	-24.3	56.0	19.3	-36.7	29.2	16.8	-12.4
ELA grade 8	18.9	3.1	-15.8	13.7	1.6	-12.1	-5.2	-1.5	3.8
Math grade 8	10.1	1.2	-8.9	15.0	2.4	-12.6	4.9	1.2	-3.7
Mean test scores									
ELA grade 4	665.3	620.3	-44.9	679.8	643.4	-36.4	14.5	23.1	8.6
Math grade 4	657.7	617.3	-40.4	684.6	651.4	-33.2	26.9	34.1	7.3
ELA grade 8	710.3	668.3	-42.1	706.0	681.2	-24.8	-4.4	12.9	17.3
Math grade 8	711.8	676.2	-35.6	725.0	698.3	-26.7	13.2	22.1	8.9

Table 4: Base Model for Math Grades 4 & 5 with Student Fixed Effects, 2000-2005

Variable	Coefficient	Variable	Coefficient	Experience	Coefficient	Variable	Coefficient
Constant	0.17147 [1.51]	SD ELA score t-1	-0.02332 [1.91]	14	0.1263 [8.21]**	Not certified	-0.04235 [5.72]**
Student changed schools	-0.03712 [6.60]**	SD math score t-1	-0.11722 [8.27]**	15	0.1252 [6.82]**	Barrons undergrad college	
Class Variables		**Teacher Variables**		16	0.12464 [6.36]**	Most competitive	0.01498 [1.48]
Proportion Hispanic	-0.4576 [12.89]**	Experience		17	0.08298 [3.10]**	Competitive	0.01426 [2.24]*
Proportion Black	-0.57974 [16.16]**	2	0.06549 [10.61]**	18	0.14161 [4.02]**	Least Competitive	0.00686 [1.25]
Proportion Asian	-0.07711 [1.75]	3	0.1105 [16.56]**	19	0.13686 [2.62]**	Imputed Math SAT	0.00043 [9.05]**
Proportion other	-0.56887 [3.95]**	4	0.13408 [17.91]**	20	0.24658 [2.50]*	Imputed Verbal SAT	-0.00034 [6.06]**
Class size	0.002 [3.36]**	5	0.117 [14.24]**	21 or more	0.38977 [3.89]**	SAT missing	-0.01535 [2.94]**
Proportion Eng Lang Learn	-0.42941 [14.16]**	6	0.13365 [14.58]**	Cert pass first	0.00657 [0.94]	Initial path into teaching	
Proportion home lang Eng	-0.02902 [1.16]	7	0.12307 [12.27]**	Imputed LAST score	0.00025 [0.57]	College Recommended	0.03108 [4.95]**
Proportion free lunch	-0.00181 [0.01]	8	0.11898 [10.81]**	LAST missing	0.00188 [0.26]	NYC Teaching Fellows	0.01173 [1.10]
Proportion reduced lunch	0.10521 [3.40]**	9	0.12433 [10.04]**	Certified Math	0.07086 [1.30]	Teach for America	0.02364 [1.20]
Mean absences t-1	-0.01367 [15.10]**	10	0.13693 [9.85]**	Certified Science	-0.04852 [0.95]	Individual evaluation	0.00866 [1.00]
Mean suspensions t-1	0.14069 [2.78]**	11	0.12592 [9.41]**	Certified special ed	0.01086 [1.05]	Other	-0.00138 [-0.09]
Mean ELA score t-1	0.33811 [31.29]**	12	0.10209 [7.66]**	Certified other	-0.00521 [0.62]	Teacher LAST* class proportion free lunch	-0.00024 [0.49]
Mean math score t-1	-0.88479 [58.78]**	13	0.11831 [8.23]**			Observations	578,630

Table 5: Effect of Observed Teacher Qualifications on Student Grades 4 & 5 Math Achievement, Most Affluent and Poorest Deciles of Schools, 2001 and 2005 for Various Model Specifications*

		Imputed SAT and LAST		Drop SAT	Drop Missing	School FE	No
		All Obs	Exp < 3	Variables	SAT Obs		Experience
Most affluent decile							
	2001	0.049	-0.011	0.093	0.129	0.074	0.050
	2005	0.056	-0.008	0.102	0.125	0.077	0.048
Change		0.007	0.003	0.009	-0.004	0.004	-0.002
Poorest decile							
	2001	-0.040	-0.106	-0.053	-0.083	-0.047	-0.032
	2005	-0.011	-0.062	-0.015	-0.027	-0.014	-0.016
Change		0.029	0.044	0.038	0.056	0.033	0.016
Gap between most affluent and poorest decile							
	2001	0.089	0.095	0.146	0.212	0.121	0.082
	2005	0.067	0.054	0.117	0.152	0.091	0.064
Change		-0.022	-0.041	-0.029	-0.060	-0.029	-0.018
Percentage reduction in gap		24.8	43.0	19.7	28.4	24.3	21.9

* Base model is as shown in Table 4; Exp < 3 includes only teachers in their first two years of teaching; Drop SAT variables omits the SAT variables from the estimation; Drop Missing SAT obs omits any teacher for whom we do not observe SAT scores, which has the effect of eliminating about 45 percent of the observations; School Fixed Effect substitutes school fixed effects for student fixed effects in the Base Model; No Experience is the base model with teacher experience omitted from the predictions.

Table 6a: Average Qualifications of Teachers in Poorest Quartile of Schools by Math Achievement Quintiles
Predicted Solely from Teacher Qualifications, 2000-2005 <u>With Experience</u>

VA Quintile	Mean VA	Years Experience	LAST Pass First	Not Certified	LAST Score	Math SAT	Verbal SAT	Barrons Ranking of Undergraduate College			
								Most Competitive	Competitive	Less Competitive	Not Competitive
1	-0.103	2.054	0.653	0.626	238	423	478	0.135	0.136	0.442	0.287
2	-0.033	5.324	0.638	0.272	242	421	466	0.102	0.096	0.493	0.308
3	-0.003	6.867	0.715	0.063	244	433	469	0.078	0.095	0.516	0.312
4	0.021	6.546	0.777	0.022	247	446	461	0.105	0.153	0.415	0.327
5	0.059	5.944	0.872	0.007	252	489	459	0.162	0.229	0.389	0.219
Range	0.162	3.890	0.219	-0.619	14	66	-18	0.027	0.093	-0.052	-0.068

Table 6b: Average Qualifications of Teachers in Poorest Quartile of Schools by Math Achievement Quintiles
Predicted Solely from Teacher Qualifications, 2000-2005 <u>Without Experience</u>

VA Quintile	Mean VA	LAST Pass First	Not Certified	LAST Score	Math SAT	Verbal SAT	Barrons Ranking of Undergraduate College			
							Most Competitive	Competitive	Less Competitive	Not Competitive
1	-0.068	0.460	0.731	227	355	440	0.036	0.065	0.548	0.351
2	-0.032	0.656	0.141	239	414	467	0.052	0.069	0.539	0.340
3	-0.010	0.779	0.076	245	423	462	0.094	0.130	0.440	0.336
4	0.010	0.851	0.031	252	450	470	0.156	0.196	0.374	0.274
5	0.045	0.908	0.013	254	512	474	0.245	0.249	0.354	0.152
Range	0.113	0.448	-0.718	27	157	34	0.208	0.184	-0.193	-0.199

31

Appendix Table 1

Average School Qualifications of Teachers by Percent of Students in School Who are Black or Hispanic, 2000 and 2005

	2000			2005			Change from 2000 to 2005		
	Lowest 10%	Highest 10%	Gap: Highest 10% - Lowest 10%	Lowest 10%	Highest 10%	Gap: Highest 10% - Lowest 10%	Lowest 10%	Highest 10%	Change in Gap
% with less than 3 years of NYC teaching experience	14.4%	26.3%	11.9%	14.7%	19.8%	5.1%	0.3%	-6.5%	-6.8%
% who failed LAST exam on first attempt	13.9%	37.0%	23.1%	15.0%	28.3%	13.3%	1.1%	-8.7%	-9.8%
% who attended least competitive undergraduate institutions	28.4%	30.1%	1.7%	30.1%	29.4%	-0.7%	1.7%	-0.7%	-2.4%
SAT verbal score	490	458	-33	493	472	-20	2	15	-12
SAT math score	480	440	-40	487	457	-30	6	17	-10
Number of absences	na	na		10.6	10.6	0.0	na	na	na
Expenditures per pupil*	$8,140	$8,923	$783	$10,940	$11,675	$735	$2,800	$2,752	-$49
Teacher salary	na	na		$59,472	$54,019	-$5,453	na	na	na

* All 2000 dollars adjusted to 2005 school-year dollars using the CPI.

Appendix Table 2

Average School Qualifications of Teachers by Percent of Students in School Who Scored at Level 1 On 4th Grade ELA Exam, 2000 and 2005

	2000			2005			Change from 2000 to 2005		
	Lowest 10%	Highest 10%	Gap: Highest 10% - Lowest 10%	Lowest 10%	Highest 10%	Gap: Highest 10% - Lowest 10%	Lowest 10%	Highest 10%	Change in Gap
% with less than 3 years of NYC teaching experience	18.3%	31.5%	13.2%	16.4%	18.0%	1.6%	-1.9%	-13.5%	-11.6%
% who failed LAST exam on first attempt	13.6%	39.1%	25.5%	15.4%	25.7%	10.3%	1.8%	-13.4%	-15.2%
% who attended least competitive undergraduate institutions	22.8%	30.9%	8.1%	22.3%	28.3%	6.0%	-0.5%	-2.6%	-2.1%
SAT verbal score	490	458	-32	494	475	-19	4	16	-12
SAT math score	475	440	-35	486	458	-28	11	18	-7
Number of absences	na	na		10.6	10.9	0.2	na	na	na
Expenditures per pupil*	$8,135	$10,124	$1,989	$10,197	$14,410	$4,214	$2,062	$4,287	$2,225
Teacher salary	na	na		$57,941	$54,566	-$3,375	na	na	na

* All 2000 dollars adjusted to 2005 school-year dollars using the CPI.

Average School Qualifications of Teachers In Elementary Schools by Student Poverty, 2000 and 2005

	2000			2005			Change from 2000 to 2005		
	Lowest 10%	Highest 10%	Gap: Highest 10% - Lowest 10%	Lowest 10%	Highest 10%	Gap: Highest 10% - Lowest 10%	Lowest 10%	Highest 10%	Change in Gap
Elementary % with less than 3 years of NYC teaching experience	15.7%	27.6%	11.9%	14.4%	20.0%	5.6%	-1.3%	-7.6%	-6.3%
% who failed LAST exam on first attempt	10.4%	37.7%	27.3%	12.3%	26.7%	14.4%	1.9%	-11.0%	-12.9%
% who attended least competitive undergraduate institutions	24.7%	29.2%	4.5%	27.2%	26.3%	-0.9%	2.5%	-2.9%	-5.4%
SAT verbal score	502	452	-50	496	474	-22	-6	22	-28
SAT math score	482	435	-47	486	459	-27	4	24	-20

Appendix Table 3b

Average School Qualifications of Teachers In Middle Schools by Student Poverty, 2000 and 2005

	2000			2005			Change from 2000 to 2005		
	Lowest 10%	Highest 10%	Gap: Highest 10% - Lowest 10%	Lowest 10%	Highest 10%	Gap: Highest 10% - Lowest 10%	Lowest 10%	Highest 10%	Change in Gap
Middle School % with less than 3 years of NYC teaching experience	16.7%	28.0%	11.3%	15.2%	26.4%	11.2%	-1.5%	-1.6%	-0.1%
% who failed LAST exam on first attempt	15.7%	32.0%	16.3%	14.6%	27.2%	12.6%	-1.1%	-4.8%	-3.7%
% who attended least competitive undergraduate institutions	24.2%	27.7%	3.5%	30.7%	24.8%	-5.9%	6.5%	-2.9%	-9.4%
SAT verbal score	501	473	-28	497.8	489.3	-9	-3	16	-19
SAT math score	517	483	-34	492.8	475.0	-18	-24	-8	-16

Appendix Table 3c

Average School Qualifications of Teachers In High Schools by Student Poverty, 2000 and 2005

	2000			**2005**			**Change from 2000 to 2005**		
	Lowest 10%	Highest 10%	Gap: Highest 10% - Lowest 10%	Lowest 10%	Highest 10%	Gap: Highest 10% - Lowest 10%	Lowest 10%	Highest 10%	Change in Gap
High School									
% with less than 3 years of NYC teaching experience	10.7%	18.2%	7.5%	16.6%	22.9%	6.3%	5.9%	4.7%	-1.2%
% who failed LAST exam on first attempt	15.7%	32.0%	16.3%	15.5%	17.7%	2.2%	-0.2%	-14.3%	-14.1%
% who attended least competitive undergraduate institutions	19.8%	22.0%	2.2%	22.5%	18.6%	-3.9%	2.7%	-3.4%	-6.1%
SAT verbal score	522.0	485.9	-36	526.7	513.8	-13	5	28	-23
SAT math score	516.5	482.8	-34	520.4	504.8	-16	4	22	-18

References

Aaronson, D., L Barrow, W Sander (2003). Teachers and Student Achievement in the Chicago Public Schools. Federal Reserve Bank of Chicago WP-2002-28.

Betts, J., K. Rueben, K. Danenberg, (2000). *Equal Resources, Equal Outcomes? The Distribution of School Resources and Student Achievement in California*, Public Policy Institute of California.

Bonesrønning, Falch, and Strøm (2005). "Teacher sorting, teacher quality, and student composition" *European Economic Review*, Vol. 49, 457-83.

Boyd D., P. Grossman, H. Lankford, S. Loeb & J. Wyckoff (2006). "How changes in entry requirements alter the teacher workforce and affect student achievement," *Education Finance and Policy* 1(2): 176-216.

Boyd D., P. Grossman, H. Lankford, S. Loeb & J. Wyckoff (2008). "Measuring Effect Sizes, the Effect of Measurement Error" working paper.

Boyd D., P. Grossman, H. Lankford, S. Loeb & J. Wyckoff (2007). "Who Leaves? Teacher Attrition and Student Achievement" working paper.

Boyd D., Lankford & J. Wyckoff (2007). "Closing the Student Achievement Gap by Increasing the Effectiveness of Teachers in Low-Performing Schools" (with D. Boyd, and H. Lankford) in *Handbook of Research in Education Finance and Policy*, H. Ladd and E. Fiske eds., Routledge (2008), pp. 535-550.

Camerson, A. and P. Trivedi (2005). Microeconometrics: Methods and Applications. Cambridge: Cambridge University Press.

Clotfelter, C., H. Ladd and J. Vigdor, (2006a). "Teacher-Student Matching and the Assessment of Teacher Effectiveness" National Bureau of Economic Research Working paper 11936.

Clotfelter, C., H. Ladd and J. Vigdor (2006b). "How and Why do Teacher Credentials Matter for Student Achievement," Working paper, Duke University.

Coleman, J S. (1966). Equality of Education Opportunity Study, Washington, DC: U.S. Department of Health, Education, and Welfare, Office of Education/National Center for Education Statistics.

Goldhaber, D., B. Gross and D. Player (2007). "Are Public Schools Really Losing Their "Best"? Assessing the Career Transitions of Teachers and Their Implications for the Quality of the Teacher Workforce," working paper.

Goldhaber, D. (2006). Everyone's Doing It, but What Does Teacher Testing Tell Us About Teacher Effectiveness? Manuscript.

Hanushek, E., J. Kain, D. O'Brien and S. Rivkin (2005). "The Market for Teacher Quality," NBER working paper.

Hanushek, E., J. Kain and S. Rivkin (2004). "Why Public Schools Lose Teachers," *Journal of Human Resources* 39(2) 326-354.

Harris, D. and T. Sass (2007). "Teacher Training and Teacher Productivity" working paper, Florida State University.

Kane, T. J., J. E. Rockoff and D. O. Staiger (2006). "What Does Certification Tell Us About Teacher Effectiveness? Evidence from New York City" NBER Working Paper 12155, April 2006.

Krieg, J.M. (2006). "Teacher Quality and Attrition," *Economics of Education Review* 25(1) pp. 13-27.

Lankford, H., S. Loeb, and J. Wyckoff (2002). "Teacher Sorting and the Plight of Urban Schools: A Descriptive Analysis," *Educational Evaluation and Policy Analysis*, Vol. 24, No. 1, pp. 37-62

Loeb, S., and L. Miller (2007). *A Review of State Teacher Policies: What Are They, What Are Their Effects, and What Are Their Implications for School Finance.* Technical Report, Getting Down to Facts Project: Stanford University.

Peske, Heather and Kati Haycock (2006). *Teaching Inequality: How Poor and Minority Students are Shortchanged on Teacher Quality*, The Education Trust, June 2006.

Rivkin, S., E. Hanushek, and J. Kain (2005). "Teachers, Schools, and Academic Achievement," *Econometrica*, 73(2), 417-458.

Rockoff, Jonah (2004). "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data," American Economic Review 94 (2): 247-252.

Sanders, W.L. and Rivers, J.C. (1996). "Research Project Report: Cumulative and Residual Effects of Teachers on Future Student Academic Achievement," University of Tennessee Value-Added Research and Assessment Center.