

# **Does Teacher Preparation Matter?**

**Evidence about Teacher Certification,  
Teach for America, and Teacher Effectiveness**

**Linda Darling-Hammond, Deborah J. Holtzman,  
Su Jin Gatlin & Julian Vasquez Heilig**

**Stanford University**

## ABSTRACT

Recent debates about the utility of teacher education have raised questions about whether certified teachers are, in general, more effective than those who have not met the testing and training requirements for certification, and whether some candidates with strong liberal arts backgrounds might be at least as effective as teacher education graduates. This study examines these questions with a large student-level data set from Houston, Texas that links student characteristics and achievement with data about their teachers' certification status, experience, and degree levels from 1995-2002. The data set also allows an examination of whether Teach for America (TFA) candidates – recruits from selective universities who receive a few weeks of training before they begin teaching – are as effective as similarly experienced certified teachers. In a series of regression analyses looking at 4<sup>th</sup> and 5<sup>th</sup> grade student achievement gains on six different reading and mathematics tests over a six-year period, we find that certified teachers consistently produce significantly stronger student achievement gains than do uncertified teachers. Alternatively certified teachers are also generally less effective than certified teachers. These findings hold for TFA recruits as well as others. Controlling for teacher experience, degrees, and student characteristics, uncertified TFA recruits are less effective than certified teachers, and perform about as well as other uncertified teachers. TFA recruits who become certified after 2 or 3 years do about as well as other certified teachers in supporting student achievement gains; however, nearly all of them leave within three years. Teachers' effectiveness appears strongly related to the preparation they have received for teaching. We discuss policy implications for districts' efforts to develop a more effective teaching force.

## **Introduction**

The relationship between teacher education and teacher effectiveness has been hotly debated in recent years in both research and policy circles (see, for example, Ballou & Podgursky, 2000; Darling-Hammond, 2000a; Darling-Hammond & Youngs, 2002; U.S. Department of Education, 2002). On the one hand, advocates of stronger preparation – especially for teachers in schools serving low-income students and students of color – have argued that such teachers need to understand how children learn and how to make material accessible to a wide range of students in order to be successful (National Commission on Teaching and America’s Future, 1996; Shulman, 1987). Studies finding positive effects of teacher education and certification on student achievement seem to support this perspective (Betts, Rueben, & Dannenberg, 2000; Darling-Hammond, 2000b; Ferguson, 1991; Fetler, 1999; Goe, 2002; Goldhaber & Brewer, 2000; Hawk, Coble, & Swanson, 1985; Monk, 1994; Strauss & Sawyer, 1986; Wenglinsky, 2000; Wilson, Floden, & Ferrini-Mundy, 2001).

On the other hand, opponents of teacher education and certification have argued that teacher effectiveness may be as much a function of general academic ability or strong subject matter knowledge as it is related to any specialized training in how to teach (Ballou & Podgursky, 2000; Finn, 1999; US Department of Education, 2002). Representing this view, the Secretary of Education argued in his 2002 report on teacher quality for the dismantling of teacher certification systems and the redefinition of teacher qualifications to include little preparation for teaching (US Department of Education, 2002). Stating that current teacher certification systems impose “burdensome requirements” for education coursework that make up “the bulk of current teacher certification regimes” (p. 8), the report argued that certification should be redefined to emphasize higher standards for verbal ability and content knowledge and to de-emphasize education training, making student teaching and most education coursework optional (p. 19).

The policy implications of these debates are far-reaching, affecting teacher education and certification policies as well as policies regarding school funding and educational rights. As teacher demand has increased and funding inequities have grown over the past 15 years, many urban and poor rural districts have hired a growing number of individuals on emergency permits or waivers who lack formal preparation for teaching.

These individuals typically teach low-income and minority students in the most disadvantaged schools (NCTAF, 1996; Shields et al., 2004). Such inequalities – and related disparities in funding and basic education materials -- have spawned lawsuits in more than a dozen states arguing that all students have the right to the resources needed to learn to state standards, including fully qualified teachers. They also sparked the “highly qualified teacher” requirements of No Child Left Behind. However, if it can be successfully argued that no special training is needed for teaching, as defendants in many of these lawsuits claim, the legal levers for redressing these inequalities would vanish as certification standards are diluted or ignored and students’ recourse is removed.

Cited in the Secretary’s report and at the center of many of these debates has been the Teach for America (TFA) program, which seeks to recruit academically able new college graduates, many of them from selective universities, into two-year teaching commitments in hard-to-staff districts. Following a summer program that provides several weeks of student teaching and basic coursework, recruits are placed in urban and poor rural schools on emergency teaching permits. Although recruits often taught without any further training in the early years of the program, states have increasingly required that they enter a teacher education program upon hiring and pursue coursework with supervision while they teach. Despite the increasing preparation the recruits receive both from TFA and from the formal teacher education programs most enter upon hiring, the program is often seen as an existence proof for the argument that bright, committed individuals can teach successfully without formal teacher education training. For example, Raymond and Fletcher (2002) suggest:

TFA corps members are an admittedly select group of college graduates, culled from the finest universities and often performing near the top of their class.... It’s possible that traditional certification programs and pedagogical training are less necessary for them than they are for the typical teacher (p. 68).

Several studies have sought to examine the effectiveness of TFA recruits, but none has explicitly compared the effectiveness of differently prepared or certified recruits using appropriate controls for students’ prior learning. Two studies have found evidence that TFA recruits’ students achieve comparable or better gains in student learning when compared to other similarly experienced teachers in similar schools (Raymond, Fletcher, & Luque, 2001; Decker, Mayer, & Glazerman, 2004), but in both of these studies the comparison group

teachers were also disproportionately untrained and uncertified teachers. Neither of these studies explicitly compared TFA teachers to teachers with standard training and certification, controlling for other student, teacher, and school variables. A study that examined the relative effectiveness of Teach for America teachers as compared to other new teachers with different levels of qualifications in Arizona found that the students of uncertified teachers, including TFA teachers, did significantly less well than those of comparably experienced certified teachers on mathematics, reading, and language arts tests (Laczko-Kerr & Berliner, 2002). However, the study did not use controls for prior achievement of students.

This study examines the question of how teacher preparation and certification influence teacher effectiveness for both TFA and other teachers. We use a newly constructed data set from Houston, Texas that allows us to link detailed certification data on teachers to background and achievement data on students, classrooms, and schools for 132,071 students who were in fourth and/or fifth grade from the 1996-1997 school year through the 2001-2002 school, and their 4,408 teachers. In this article, we report on the results for these students and teachers on several different achievement tests: the TAAS, the SAT-9, and the Aprenda .

### **Study Methods**

This study largely replicates the results of an earlier study of TFA recruits in Houston conducted for the Hoover Institution's CREDO center by Raymond, Fletcher, and Luque (2001). We recreate and go beyond their analyses to examine a wider range of achievement measures over a greater number of years with additional controls, and we include examination of teacher certification pathways more generally. The CREDO study examined the effect of TFA teachers on student achievement gains on the TAAS reading and mathematics tests from grades 3 through 8 between 1996 and 2000. The analyses pooled the data across these years and found that, in most estimates, after controlling for teacher experience, along with individual student, classroom, and school demographics, TFA recruits were about as effective as other teachers of comparable experience working in similar teaching settings. The study found statistically significant positive coefficients for TFA recruits in 2 of 10 estimates: when the students of TFA recruits were compared to those of beginning teachers with 0-1 years of experience on the TAAS mathematics test in 4<sup>th</sup> and 5<sup>th</sup> grades and when the students of TFA teachers were compared to those of other

teachers on the TAAS mathematics test in 6<sup>th</sup> through 8<sup>th</sup> grades. (Two other estimates, both in reading, were significant at the .10 level.)

Although the researchers stated that “students of Teach for America teachers learn more” than those of all other teachers, the analyses actually found that experienced teachers in Houston were significantly more successful than inexperienced teachers, including TFA teachers. In all of the equations, the variable representing teacher years of experience was strongly significant, and had a much larger effect on student outcomes than the TFA variable.

The CREDO study did not examine whether TFA teachers were differentially effective when compared to traditionally prepared and certified teachers in Houston, although the researchers noted that certification status was one of the variables in their data set. This question is an important one because TFA teachers in Houston were compared to an extraordinarily underqualified pool of teachers. TFA teachers were placed in schools serving high percentages of low-income and minority students, where most underqualified teachers in the district are placed and where, the study found, students lose ground in achievement from year to year. In 1999-2000, the last year of the CREDO study, about 50% of Houston’s new teachers (and one-third of all teachers) were uncertified, and the researchers reported that 35% of new hires lacked even a bachelor’s degree. (In a later paper, the lead author indicated that this statistic on degree status was incorrect and that the actual proportion of non-degreed teachers in Houston was likely lower.) The CREDO study pointed out that TFA teachers were more likely to hold a bachelor’s degree than other Houston teachers, but the study did not control for degree levels when evaluating TFA effects.

### **The Data Set**

Our study assembled a similar data set with the assistance of the Houston Independent School District (HISD).<sup>1</sup> The data set consists of information on all HISD teachers and students in grades 3 and higher from the 1995-1996 school year through the 2001-2002 school year (a total of 271,015 students and 15,344 teachers). We created a merged longitudinal data file from several files containing student-level data (e.g.,

---

<sup>1</sup> We could not access the CREDO data set for re-analysis as, we were told, it was a “proprietary” data set.

demographic characteristics and test scores on three sets of tests in reading and mathematics), teacher data (e.g., years of teaching experience, highest degree completed, certification information, and Teach for America participation), school data (e.g. student demographic information by school), and identifier data linking students with teachers by school year.

Students in the elementary grades were typically linked with a single teacher, presumably a teacher of a self-contained classroom teaching mathematics and reading, as well as other subjects. Students in secondary grades were linked with several different teachers both within and across subject areas, sometimes as many as 24 different teachers within a given year. Unfortunately, our data set did not allow us to match all teachers with the subjects they taught or to evaluate why there might be so many links for some students. The CREDO analysis eliminated many middle school students from their analysis because of these difficulties and created a “TFA intensity” ratio for the remaining students who had a TFA teacher as one of several teachers. We had a number of concerns about this methodology; consequently, we decided not to pursue an analysis of teacher effectiveness for grades 6 and above, and we limited our analysis to evaluating individual student gain scores linked to teacher characteristics in grades 3-5. Our analyses measured gains from spring of 3<sup>rd</sup> to spring of 4<sup>th</sup> grade and similarly from 4<sup>th</sup> to 5<sup>th</sup> grade, looking at effects associated with students’ 4<sup>th</sup> and 5<sup>th</sup> grade teachers.

We had a total of 406,036 records on students who were in grades 3 through 5 from 1995-96 through 2001-2002. (A student is represented by a distinct record for each year he or she is in the data base.) Links to teachers were available for 383,799 of the records. Most of the students without teacher links were coded as “no show” or “withdrawn” in the district records.<sup>2</sup> In each year from 1995-96 to 2001-02, about 55,000 students were linked with teacher records.<sup>3</sup>

---

<sup>2</sup> Of the 22,237 records without teacher links, 18,465 (83%) were coded with an enrollment status of “no show” or “withdraw.” In contrast, of the 383,799 records with teacher links, only 9.1% had a “no show” or “withdraw” enrollment status code.

<sup>3</sup> The numbers of grades 3-5 students with links to teachers were as follows: 53,181 in 1995-96; 54,425 in 1996-97; 54,655 in 1997-98; 52,699 in 1998-99; 55,164 in 1999-2000; 56,366 in 2000-01; 57,309 in 2001-02.

## Variables

**Outcome variables.** As measures of student achievement, we used student test scores in mathematics and reading on three separate standardized tests administered by Houston during the period studied: the Texas Assessment of Academic Skills (TAAS), the Stanford Achievement Test, 9<sup>th</sup> Edition (SAT-9), and the Aprenda.

The TAAS is a state-mandated, criterion-referenced test that was administered statewide each spring from 1994 through 2002. The examination was given in grades 3-8 and 10. A Spanish version was available for grades 3-6. For the TAAS, the data provided by HISD contained only two metrics: (1) whether or not students met minimum expectations, and (2) the Texas Learning Index (TLI), a derived continuous score that allows for cross-year and cross-grade comparisons.<sup>4</sup> Since we were interested in score gains, we used the TLI index. The TLI, however, was available only for the English TAAS and not for the Spanish TAAS.<sup>5</sup> Thus, our TAAS / TLI analyses apply only to students who took the English TAAS.<sup>6</sup>

Because of concerns raised by other researchers regarding potential score distortions on the high-stakes TAAS examinations (Klein, Hamilton, McCaffrey, & Stecher, 2000), we were interested in alternative measures of student achievement as well. Houston began to administer national norm-referenced tests in 1997-1998. The SAT-9 was administered to 1<sup>st</sup> through 11<sup>th</sup> graders who received reading and language arts instruction in English. A Spanish-language test, the Aprenda, was administered to 1<sup>st</sup> through 9<sup>th</sup> graders who received instruction in Spanish. The 1997-1998 administration was in the fall; administrations in subsequent years were in the spring. Because we wanted to look at growth over a single school year (e.g., spring to spring), we began our SAT-9

---

<sup>4</sup> The TLI range was approximately between 0 and 100 but differed by subject area and by grade each year. For example, for grade 5 math in spring 2000, the top score was 93, but for grade 5 reading in spring 2000, the top score was 101. More information on the TLI can be found on the Texas Education Agency website at <http://www.tea.state.tx.us/student.assessment/resources/techdig02/index.html> (see Chapter 10, "Scaling") and <http://www.tea.state.tx.us/student.assessment/reporting/freq/index.html>.

<sup>5</sup> In 1995-1996 and 1996-1997, large numbers of students who took the Spanish TAAS were, apparently incorrectly, assigned a TLI of 0. These students were not included in the analysis; we selected only students who took the English version.

<sup>6</sup> The data also contained a "score code" to indicate such things as exemptions due to absenteeism, disability, and LEP; a code of "S" meant that the score was suitable for inclusion in calculations. Because many of the scores for students with non-S codes were apparently invalid (e.g., they were often coded as the minimum TLI value for a given test in a given grade in a given year), and because of concern about how scores of non-S coded records could be interpreted, we included in our analyses only the scores of students who had a score code of "S."



and Aprenda analyses with the first spring administration, in 1998-1999. We used normal curve equivalent (NCE) scores to measure annual changes in student performance.<sup>7</sup>

**Control Variables.** A variety of individual, classroom, and school factors can affect student achievement, and we attempted to control for as many of these factors as possible in testing the influence of teacher certification and Teach for America status. To the control variables included in the CREDO study we added teacher degree levels and a proxy for class size. The full set included:

- **Student prior achievement:** We controlled for prior achievement by including in our regression models each student's prior-year test score. Because our data set begins with third grade students, we looked at student performance for students in fourth and fifth grades, controlling for each student's achievement on the same test a year earlier. The inclusion of the prior year score variable also means that our analyses begin with the second year of data for each achievement measure (1996-1997 for the TLI, and 1999-2000 for the SAT-9 and Aprenda), as the first year of achievement data is used as a control.
- **Student demographic characteristics:** HISD provided data on the following student-level variables: race/ethnicity (American Indian, Asian/Pacific Islander, African American, Hispanic, and white), eligibility for free/reduced price lunch,<sup>8</sup> and limited English proficiency (LEP). For LEP status, the HISD data set contained several different codes representing different levels of eligibility. After we conducted exploratory analyses with these finer categories and found that they behaved similarly, we collapsed LEP status into a binary yes/no variable.<sup>9</sup>
- **Teacher's years of experience and highest degree completed:** We used teachers' total years of teaching experience as a continuous variable.

---

<sup>7</sup> We would have preferred to use scaled scores for the SAT-9 and Aprenda analyses, but scaled scores were not included in the data we received from HISD.

<sup>8</sup> Students with all codes other than "paid" were classified as being eligible for free/reduced price lunch; students with a "paid" code and with no code at all (blank value) were classified as not being eligible for free/reduced price lunch.

<sup>9</sup> In exploratory analyses, we examined students whose LEP code indicated "former" or "tested but did not qualify" in addition to those coded "LEP" or "not LEP," but in the final analyses, these students were classified as being LEP since the variables behaved similarly to the "LEP" code in regressions and did not add any power to the estimates.

Highest degree completed has four levels, used as dummy variables: no degree, bachelor's degree, master's degree, and doctoral degree.

- Classroom level variables: The data obtained from HISD did not contain classroom-level information, but we were able to create classroom-level variables by aggregating up from the individual student-level data, using the teacher identifier to group students. Among the classroom-level variables we included in our analyses is the average prior year score (for all of the grades 3-5 students in the class with prior year scores). This provides an indication of the teaching context and accounts in part for the influence of peers in the learning environment. We also included the number of students in grades 3-5 in the class as a proxy for class size.<sup>10</sup> We considered a variable characterizing the socioeconomic make up of the classroom (the proportion of free / reduced price lunch students), but we found that this variable was highly collinear with both the individual free / reduced price lunch variable ( $r=.65$ ) and the school free/reduced price lunch variable ( $r=.83$ ), as well as being strongly related to the prior year achievement scores on each test. Thus, we kept the individual and school-level variables, but not the classroom level variable.
- School level demographics: The HISD data contained some school-level student demographic variables, including the school's percentage of African American students, the percentage of Hispanic students, and percentage eligible for free/reduced price lunch. We used these three variables to capture features of the school that may be relevant both to teaching context and to community characteristics.

**Teacher Preparation Pathway and Certification.** In exploring the influence of preparation and certification on teacher effectiveness, we had a complex set of variables to examine, represented by more than 100 certification and license codes in use in Houston.

---

<sup>10</sup> There were some "classes" that appeared to have extraordinarily small class sizes, as low as one or two students. These could have represented classrooms that largely served other grade levels (hence the number of 3<sup>rd</sup>, 4<sup>th</sup>, or 5<sup>th</sup> grade students' records tied to a teacher's record was very small), that were primarily tutorial situations for students with particular needs, or that were very small special education classes. Because these would result in teacher effectiveness being evaluated on very small sample sizes, we restricted the classes we included in the estimates to those with a class size of 15 or above.

These represented both certification categories and subject matter and specialty areas in which teachers held a license. Although we would have liked to have had direct measures of teacher preparation, such as coursework and program measures, we could use these codes to identify many aspects of the pathways teachers pursued into teaching: for example, whether teachers began teaching with a credential or entered as an emergency certified teacher before attaining a standard credential, whether they entered through an alternate certification program, whether they had certification in multiple areas (including specialty areas like bilingual education or reading), and whether they entered teaching through Teach for America. We also examined the interaction between TFA status and certification, since most TFA teachers became certified during their tenure.

Teach for America status is a straightforward yes/no variable. Certification status, on the other hand, is considerably more complex. Houston provided us with two different files containing data on teacher certification, each of which had a different certification coding scheme. One file contained eight different certification types, and the other contained 13 distinct certification codes. (Any given teacher was in one of these files only.<sup>11</sup>) These different classifications were related to changes in the state and local certification systems over the study years. In addition to the types of certification, the files included the areas of certification (e.g. elementary, bilingual, reading, music, counselor, etc.). We found that, for elementary teachers, these areas of certification were very similar across teachers with standard certification (for example, most teachers secured a reading endorsement along with an elementary teaching certificate), and after exploring the file, we concluded that adding these additional details would not contribute to our explanation of differences in teacher effectiveness.

Finally, our data included a certification date attached to the records of more than 75% of the teachers in grades 3-5. The certification date was frequently not the year of initial employment, since many teachers enter Houston schools without certification and secure some form of training that leads to a certification or permit later. Since we were concerned with the amount and kind of training a teacher might have in a specific year, we created a year-by-year certification code for each teacher. After investigating the data, we

---

<sup>11</sup> An additional 293 teachers of grades 3-5 were not included in either file. No certification data was available for these teachers, so their certification is categorized as “unknown.”

learned that most teachers' certification dates were in the spring and summer months prior to a given school year. We coded teachers who were not certified by the start of a given school year (defined as August 30) as being uncertified for that year; thus, a teacher's certification code does not "kick in" until the school year following his or her certification date. If this decision were to bias the results, it would be on the side of understating the differences in effectiveness between certified and uncertified (or "not yet certified") teachers, rather than overstating them.

We coded the 1558 teachers in grades 3-5 who had a certification code but no certification date as possessing their certification for all of the years in which they are in our data. However, we created a new control variable, "certification date unknown," which takes a value of 1 for the teachers without a certification date and 0 for teachers with a certification date. In this way, we control for the fact that for this group of teachers, we do not know whether they actually possessed their certification in any given year.

After researching Texas' credentialing system and evaluating our data, we collapsed the certification types into the following categories:

- *Standard* includes the standard, provisional, professional, and out-of-state certificates, all of which require the completion of an approved educator preparation program and passage of the appropriate certification examinations.<sup>12</sup>
- *Alternative* includes the Texas Alternative Certification Program (ACP) and probationary certificates, which are issued to individuals who have a bachelor's degree, have passed subject matter tests, and are accepted into approved alternative certification programs in Texas.<sup>13</sup> These alternative and probationary certificates are renewable for up to 3 years while the individual completes the requirements for a standard certificate.
- *Emergency/temporary* includes the categories of emergency permit uncertified, permit (teacher aide), temporary certificate, and recognition. These categories,

---

<sup>12</sup> The provisional and professional certificates were the Texas lifetime certificates granted before 1999. The professional certificate was granted to individuals with a postbaccalaureate degree. After 1999, these were replaced by the standard certificate, granted for a 5 year renewable term. Out-of-state certificates are granted to individuals who hold the equivalent standard certificate in other states. The specific requirements for each credential can be found in the Texas Administrative Code, Title 19, Part 7, Chapter 230 and Chapter 232. [http://info.sos.state.tx.us/pls/pub/readtac\\$ext](http://info.sos.state.tx.us/pls/pub/readtac$ext).

<sup>13</sup> In Houston, about half of ACP holders are enrolled in the HISD intern program. The others attend programs at universities (e.g. Prairie View A&M, University of Houston), through the Regional office, or are enrolled in something called a "deficiency plan," an individualized program to make up specific needs for the credential.

used either by the state or the Houston district (which has some of its own categories of permits), are issued on a temporary basis to individuals who have not undertaken teacher education and who are supposed to receive mentoring and training while they are teaching.

- *Certified out-of-field* includes individuals who are “emergency permit certified” – that is, already certified but teaching out of their field of certification – and those who hold a temporary classroom permit to teach in a field other than their field of preparation.
- *Certified no-test* includes individuals holding the school district permit or non-renewal permit granted to those who have completed preparation but not passed the state test.
- *Uncertified* is the designation we gave later-certified individuals during the years before they secured some kind of state or local certification or permit.<sup>14</sup>
- *Certification unknown* includes individuals for whom our files held no certification code.

In some analyses, we further collapsed these categories into standard certification, alternative or other nonstandard certification, and uncertified.

Although the files included up to eight separate certification codes for each teacher, an analysis of the codes indicated that, for most teachers with multiple codes, all of the codes were in the same category (within the seven classifications listed above).<sup>15</sup>

Therefore, we based our certification variable on the first certification code assigned to each teacher.

---

<sup>14</sup> We actually created six separate “certified later, but not yet” categories: one for each of the certification codes, such that we could tell not only that a teacher was certified later, but specifically which category of certification they later received. In analyses using these variables, we found that the categories all behaved similarly. Thus, in the analyses published here, all of the “certified later, but not yet” teachers were grouped into a single category of “uncertified.”

<sup>15</sup> Of the 15,344 total teacher cases in our file, 8756 had only a single certification code. Of the remaining 6588, 2084 had the same certification code for their second certification area as for their first and 4504 had codes that were different only because of the change in state terminology for the standard certificate before and after 1999. The “standard,” “provisional,” and “professional” codes are the different terms given to certificates for fully prepared teachers before and after a major certification reform in 1999. Of the remaining 460, many had non-standard codes that were also equivalent versions of different state and local terms used before and after 1999 (e.g. 67 teachers coded as emergency uncertified/school district permit and 61 coded as ACP/probationary). Thus, we used only a single certification code for each teacher. Among 2104 teachers of grades 3-5 who had first code of “provisional,” 1488 had a second certification code. For 781 of these teachers, the second code is also “provisional”; for 694, the second code is “professional.” In line with the state rules, we categorize all of these as “standard” certificates.

TFA teachers fell into the categories of uncertified, alternatively certified, and standard certified teachers in different years of their teaching careers. Because our goal was to look not only at the effects of TFA controlling for certification, but also to look specifically at the interaction between TFA status and certification, we examined TFA teachers in each category of certification in comparison to non-TFA teachers in each category of certification.

## **Methods**

After merging our several data files and cleaning the resulting data set, we ran a series of descriptive analyses of the characteristics of students and teachers and examined the distribution of teachers to students of different kinds. Then we developed ordinary least squares regression estimates of the predictors of six sets of student test scores (three tests with two subject areas for each), both with data pooled across all of our study years and for each year individually. In each set of estimates, conducted at the individual student level, we controlled for prior year test scores, student race / ethnicity, poverty, and language status; teacher years of experience and degree level; class size and class average previous year's test score; and school demographics (the percentage of students who were African American, the percent Hispanic, and the percent eligible for free/ reduced price lunch). Key independent variables were teacher certification status and TFA status.

## **Descriptive Findings**

Our data set closely reflects both published data about the Houston Independent School District and the CREDO data. We summarize here data on the 4<sup>th</sup> and 5<sup>th</sup> grade students and teachers who are the focus of our analysis. Table 1 shows 4<sup>th</sup> and 5<sup>th</sup> grade student enrollments by race/ ethnicity, language background, and socioeconomic status for each of the years in our study. Hispanic students in Houston comprise a majority of the population, followed by African American students, who comprise about a third of the student body. About three-quarters of the students are eligible for free or reduced price lunch. Before 1998, about 12% of students were classified as limited English proficient, but the proportion jumps to over one-third from 1998-99 on. Our conversations with

Houston staff suggested that this was due to a change in classification practices between 1997-98 and 1998-99, rather than a change in student demographics.

	<b>1996-97</b>	<b>1997-98</b>	<b>1998-99</b>	<b>1999-00</b>	<b>2000-01</b>	<b>2001-02</b>
<u>Total N</u>	37,396	37,536	34,589	36,622	38,015	38,928
American Indian	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%
Asian/Pacific Islander	2.6%	2.7%	2.8%	2.9%	2.9%	2.8%
African American	33.8%	33.7%	33.8%	33.7%	32.8%	32.6%
Hispanic	51.1%	51.9%	52.4%	52.1%	53.3%	54.8%
White	12.4%	11.7%	10.9%	11.3%	10.9%	9.7%
Free/Reduced Lunch	74.7%	72.7%	76.7%	71.1%	72.5%	75.0%
Limited English Proficient	11.0%	12.2%	34.4%	37.5%	40.3%	42.1%

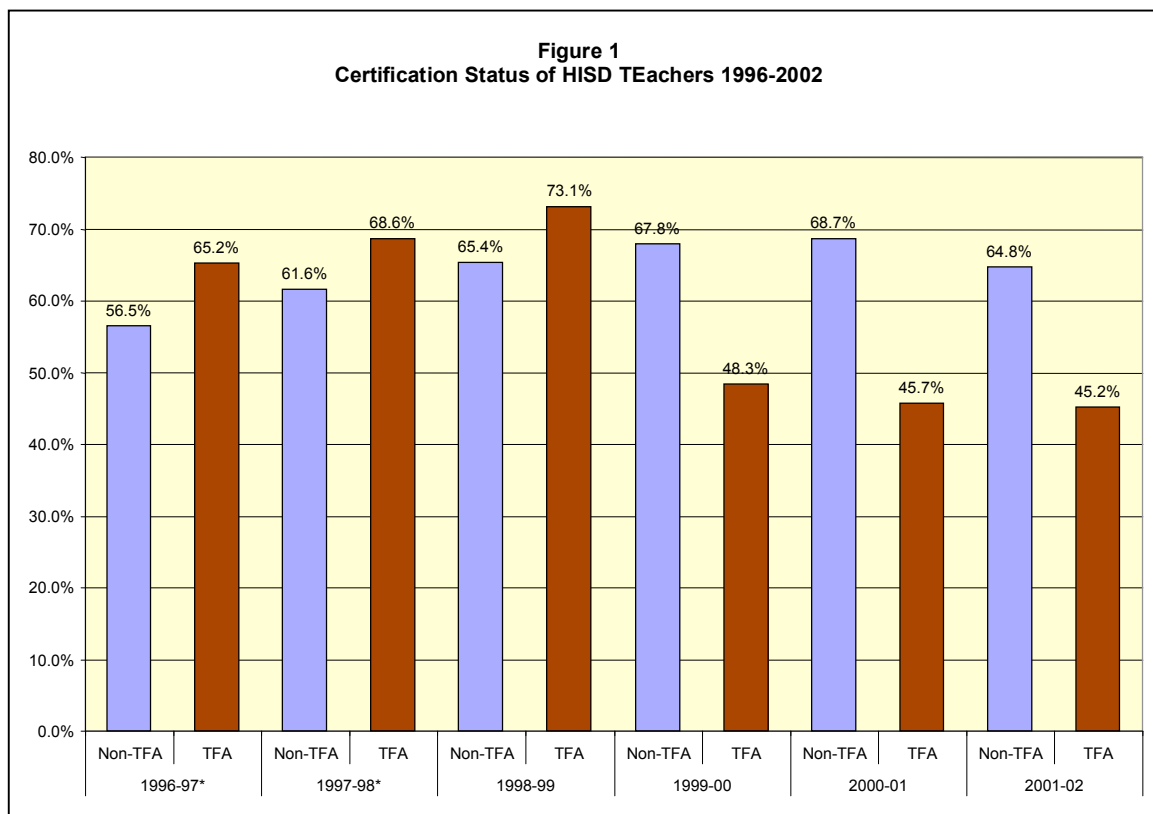
Houston's teaching force for these years was more inexperienced than most. (See Table 2.) Nationally, the teaching force averaged about 15 years of experience during these years, and about 5 percent of all teachers were brand new to teaching. In Houston, beginning teachers (with less than 2 years of experience) were a large and growing share of the teaching force, increasing from 14% to 23% of teachers from 1996 to 2002. Teachers with six or more years of experience decreased from 64% to 57% during that time.

	<b>1996-97</b>	<b>1997-98</b>	<b>1998-99</b>	<b>1999-00</b>	<b>2000-01</b>	<b>2001-02</b>
<u>Total Teaching Experience</u>						
0-1 year	14.2%	15.0%	19.0%	18.4%	19.2%	22.7%
2-5 years	21.6%	19.8%	16.8%	18.8%	19.9%	20.5%
6-10 years	19.0%	18.0%	18.4%	17.7%	16.5%	15.3%
11 or more years	45.2%	47.2%	45.8%	45.1%	44.4%	41.5%
<u>Highest Degree Earned</u>						
Less than Bachelors	0.5%	0.2%	3.5%	3.8%	5.5%	0.6%
Bachelors	67.6%	69.2%	66.4%	68.7%	67.4%	75.0%
Masters	30.9%	29.9%	28.9%	26.4%	25.5%	23.2%
Doctoral	0.8%	0.7%	1.1%	1.1%	1.6%	1.2%
Total N	1841	1802	1617	1770	1805	1794

Note: Includes all teachers who taught any 4<sup>th</sup> or 5<sup>th</sup> grade students.

Most teachers held bachelors degrees, with the proportion of such teachers increasing from 68% to 75% from 1996 to 2002 and the share with masters degrees declining at the same rate. The data show an unusual bump in the proportion of teachers with less than a bachelors degree between 1998 and 2000 (increasing from less than 1 percent to 5.5% in 2000-01 and then dropping back down to 0.6% in 2001-02).<sup>16</sup>

We also looked at the certification status of TFA and non-TFA teachers for each year of the study. (See Figure 1, below, and Table A-1 in the Appendix.) Strikingly, in each year of the study, the proportion of teachers in these grade levels teaching without standard certification ranged from one-third to nearly half. In the early years of the study (1996-1998), Teach for America teachers appeared to be as likely as other teachers in these grade levels to hold standard certification.<sup>17</sup> From 1999-2001, Teach for America teachers were noticeably less likely to hold standard certification than other teachers.<sup>18</sup>



<sup>16</sup> As the CREDO research team suggested about the anomaly in the recorded number of teachers with less than a bachelors degree in these years, we believe it is possible that there were data coding errors in the HISD data set.

<sup>17</sup> We note, however, because of the quality of available certification data for the earlier years of the study, a noticeable proportion of teacher records lacked certification dates in the personnel file covering the years prior to 1998-99, which makes us unwilling to draw strong conclusions about when teachers acquired their certification in these years.

<sup>18</sup> The personnel files provided by Houston covering the post-1999 period had certification dates for virtually all teachers, giving us greater confidence that teachers' year to year certification status is accurate.



These certification patterns are related to the experience levels of TFA recruits, who make a two-year commitment to teaching. Most TFA recruits are placed in a teacher education program upon arrival (with few exceptions, this is the Houston alternative certification program), and most are certified after one or two years. In our data set, TFA recruit tended to be certified by their second or third year of teaching.<sup>19</sup>

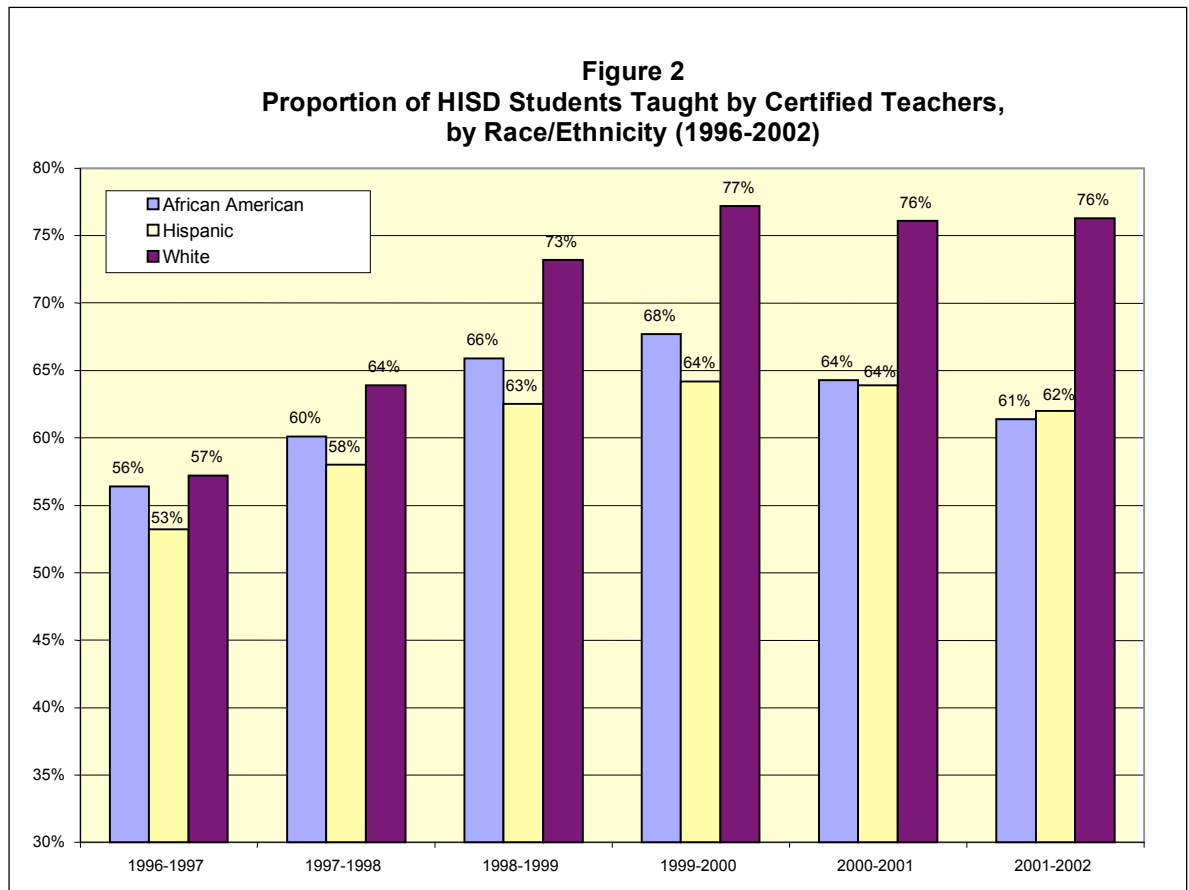
The Houston program is designed to provide beginning teachers with weekly training sessions run by the district and a mentor at their school, as well as release time to observe other teachers once a month. The recruits take teacher education courses at a local university; in recent years the University of St. Thomas has offered up to six courses in the certification program and a more expanded master's degree at a discounted rate for those who want to pursue additional study. The program is designed to be finished in a year; however, many TFA recruits in our data base did not become certified until their third year of teaching. This may be a function of taking additional time to complete courses or pass the requisite certification tests, or because some recruits did not complete the program. We found that teachers without standard certification, including TFA teachers, were disproportionately likely to be teaching African American and Latino students and low-income students. Although the percentages of Houston students being taught by standard-certified teachers rose substantially over the years covered by this study (peaking in 1999-2000), the racial/ethnic and economic disparities associated with students' access to certified teachers also increased substantially. In 1996-1997, for example, 56% of black students and 57% of white students were taught by standard certified teachers, a difference of less than 1%. By 2001-2002, 76% of white students had standard-certified teachers, while only 61% of black students did – a difference of nearly 15%. (See Figure 2.)

Similarly, in 1996-97, 54% of low-income students (those eligible for free or reduced price lunch) had teachers who held standard certification, as compared to 57% of students not eligible for free or reduced price lunch, whereby 2001-02, the proportions were 61% and 72%, respectively.

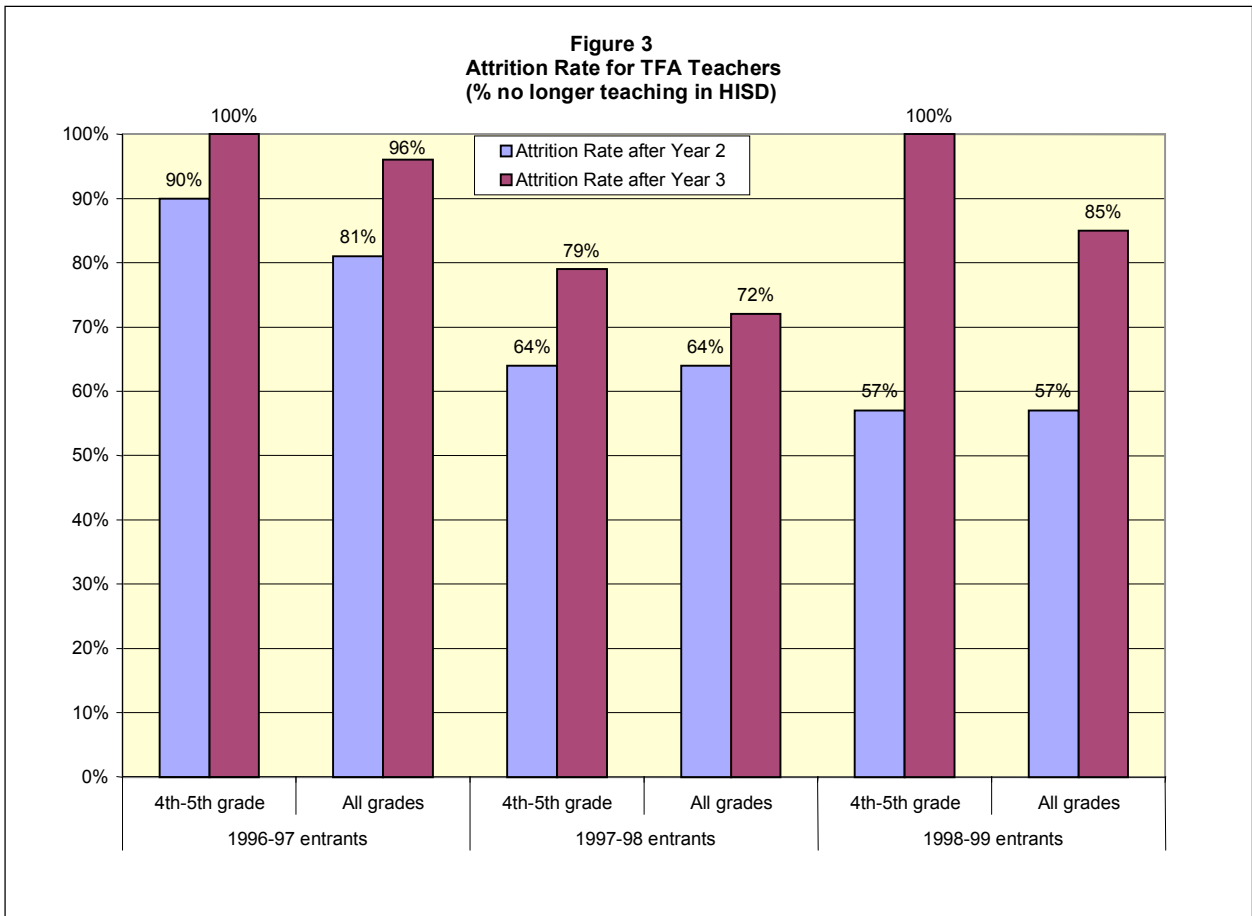
---

<sup>19</sup> The trend for many TFA teachers to become certified by their second or third year of teaching is clear in the data we received. However, in Houston's records, only a few of the TFA recruits were coded in the "alternative certification" category during their first year of teaching, although most were placed in Houston's district-run alternative certification program. It appears that recruits received a code of "standard certification" when they finished their preparation program, but did not have a separate "alternative" certification code prior to that time. Thus, in our table 2, most are coded as uncertified until they received the standard certification.

**Figure 2**  
**Proportion of HISD Students Taught by Certified Teachers,**  
**by Race/Ethnicity (1996-2002)**



Like the CREDO study, we found high rates of attrition for TFA teachers. Raymond and colleagues reported that across the four years of their study, TFA teachers had left teaching in Houston by their third year at rates of between 60% and 100%. We found similar attrition rates: Between 57% and 90% of TFA recruits had left teaching in Houston after their second year, and between 72 and 100% of recruits had left after their third year. (See Figure 3.) Thus, although a substantial proportion of TFA recruits became certified within two or three years, few stayed in the district after they had completed their initial preparation for teaching.



### Analyses of Teacher Effectiveness

#### The General TFA Effect

Although our data set differed in minor ways from that used by the CREDO research team, we were able to largely replicate the results of their TFA study for the teachers of 4<sup>th</sup> and 5<sup>th</sup> graders. That is, when we examined the effects of TFA status on student achievement gains on the TAAS for the pooled years 1996-97 to 2001-02 -- controlling for prior achievement, student demographic characteristics, and other teacher characteristics such as experience and degree level -- we found a similar pattern of results: TFA teachers exerted a significant positive effect on achievement on the TAAS/ TLI in math and a nonsignificant effect on the TAAS/ TLI in reading. (Table A-2 in Appendix.)

However, when we looked at other test measures and for individual years, the results were quite different. On the SAT-9, TFA teachers had a significant negative effect on student scores in math; on the Aprenda (the test given to Spanish-speaking students),

the TFA effect was significantly negative in both reading and math. The TFA coefficient was negative on five of these six tests (Table 3.)

**Table 3**  
**TFA Effects**  
(All Years of Test Data Pooled)

	TAAS/ TLI Math	TAAS/ TLI Reading	SAT-9 Math	SAT-9 Reading	Aprenda Math	Aprenda Reading
TFA teacher status	.696*** (4.02)	-.056 (-.265)	-.840** (-2.60)	-.575~ (-1.86)	-2.39* (-2.39)	-2.37* (-2.39)

~p < .10; \*p < .05; \*\*p < .01; \*\*\*p < .001

Equations control for: Student's previous year's test score, student race/ethnicity, free/reduced price lunch status, and LEP status; teacher years of experience, degree level, certification date unknown; class number of students, class average previous year's test score, school demographics (racial/ethnic composition and poverty).

Furthermore, the effects varied by year on all of the tests. On the TAAS / TLI math tests, the positive effects found in the overall analysis by both us and CREDO held only for the years 1996-97 through 1998-99. Starting in 1999-2000, when TFA teachers were noticeably less likely to be certified than other teachers in these grade levels, the TFA effect in math became non-significant, with negative coefficients in two of the three years. In reading, where the overall effect had been negative but non-significant, the TFA coefficients become significant and negative in two of the three years from 1999-00 to 2001-02. (See Table 4.)

**Table 4**  
**TFA Effects, By Test and Year**

	TLI Math	TLI Reading	SAT-9 Math	SAT-9 Reading	Aprenda Math	Aprenda Reading
<b>1996-97</b>	1.49** (2.79)	-.063 (-.100)	...	...	...	...
<b>1997-98</b>	1.37** (2.90)	.302 (.573)	...	...	...	...
<b>1998-99</b>	2.81*** (5.23)	2.41*** (3.73)	...	...	...	...
<b>1999-00</b>	-.785~ (-1.75)	-1.21* (-2.16)	-.699 (-1.06)	-1.21~ (-1.91)	-1.25 (-.894)	.975 (.692)
<b>2000-01</b>	.312 (1.04)	-.655 (-1.53)	.212 (.374)	.468 (.882)	-3.74* (-2.22)	-2.23 (-1.34)
<b>2001-02</b>	-.354 (-1.47)	-.959** (-2.70)	-1.63** (-3.37)	-.908~ (-1.92)	-6.32** (-3.31)	-8.48*** (-3.95)

~p < .10; \*p < .05; \*\*p < .01; \*\*\*p < .001

On the SAT-9 in math, the TFA effect is non-significant in all years except 2001-02, when it is significantly negative. On the SAT-9 in reading, the TFA effect is significant and negative at the .10 level in two of three years. On the Aprenda in math, the TFA coefficients are negative in all three years, and are significantly negative in 2000-01 and 2001-02. On the Aprenda in reading, the TFA effect is non-significant in 1999-00 and 2000-01, but it is negative in 2001-02.

Thus, examined across measures and individual years, the TFA effect is positive only on the TAAS/ TLI and only in the first three years of the six years of data we analyzed. It is non-significant or negative on other measures and in the three later years when TFA teachers were noticeably less likely to be certified than other teachers. It is worth noting that the differences we observed in TFA effects across tests may be a function of the years in which the tests were offered, since the SAT-9 and Aprenda were administered and analyzed in the years after 1998-99, when TFA recruits were, as a group, less well qualified, and were also found to be less effective on the TAAS.

### **The Effects of Certification**

Finally, we examined the effects of certification and TFA status on teacher effectiveness. First we looked at the effects of different certification categories on teacher effectiveness, irrespective of TFA status. Then we looked at the interaction of certification and TFA status.

We found that, relative to teachers with standard certification, uncertified teachers and those in most other substandard certification categories generally had negative effects on student achievement, after controlling for student characteristics and prior achievement, as well as teacher experience and degrees. Uncertified teachers showed negative effects across all six tests, and five of the coefficients were strongly significant. Similarly, teachers who were certified without passing the state teacher certification tests (a special permit category in HISD) showed negative effects across all six tests, and four of the coefficients were strongly significant. (See Table 5).

Alternatively certified teachers had negative effects on achievement on five tests, three of them significant. On the Aprenda in reading, alternatively certified teachers had a significant positive effect. Since the Houston alternative certification program enrolls a

substantial number of Hispanic teachers, it may be that more of these teachers are Spanish-speaking and able to support the literacy progress of Spanish-speaking students who take the Aprenda. Teachers on emergency or temporary certificates showed negative effects on student achievement on five of the six tests (three of these significant at the .10 level or above); however, they showed a positive effect on the SAT-9 in reading. Teachers with missing certification codes (who are likely to be uncertified) also showed significant negative effects on student achievement on four of six tests.

**Table 5**  
**Teacher Certification Status and Student Achievement Gains**  
 Unstandardized Coefficients (T-Value)

	<b>TLI Math</b>	<b>TLI Reading</b>	<b>SAT-9 Math</b>	<b>SAT-9 Reading</b>	<b>Aprenda Math</b>	<b>Aprenda Reading</b>
Uncertified	-.533*** (-8.130)	-.573*** (-7.114)	-.454*** (-3.608)	-.579*** (-4.806)	-1.772*** (-5.704)	-.122 (-.394)
Alternative certification	-.899* (-2.021)	-.809 (-1.474)	-2.322** (-3.312)	-1.431* (-2.131)	-.202 (-.187)	3.183** (2.955)
Emergency/ temporary certification	-.707*** (-3.881)	-.685** (-3.050)	-.611 (-1.613)	.819* (2.261)	-2.380~ (-1.931)	-1.030 (-.839)
Certified, but Out-of-field	.546* (2.050)	.914** (2.794)	-1.825*** (-4.408)	.220 (.556)	-3.765* (-2.506)	3.806* (2.555)
Certified, without test	-.156 (-.546)	-.662~ (-1.891)	-4.404*** (-6.916)	-2.217*** (-3.633)	-3.272** (-3.387)	-4.536*** (-4.749)
Certification code missing	.113 (.548)	-.639* (-2.520)	-2.253*** (-4.622)	.127 (.272)	-5.710*** (-4.563)	-4.438*** (-3.557)
<b>R Squared</b>	.43	.40	.62	.68	.42	.43
<b>N of Students</b>	105511	103122	60488	60607	11437	11436

~p<.10, \* p < .05, \*\* p < .01, \*\*\*p< .001; Reference group is standard-certified teachers.

Interestingly, teachers who were already certified but credentialed to teach out of field had significant positive effects on two out of three reading tests (the TAAS / TLI and Aprenda, with a positive coefficient on the SAT-9 as well), but significant negative effects on two of three math tests (the SAT-9 and Aprenda). The coefficient for these teachers on the TAAS/ TLI – math test was positive and significant.

Overall, teachers with standard certification were found to be significantly more effective in raising student test scores than teachers without certification or with substandard certification in 22 of 36 estimates (p<.10). Coefficients for teachers without standard certification were negative in 28 of the estimates.

**The Combined Effects of TFA and Certification Status**

In Table 6 we display the results for Teach for America teachers within different certification categories considered in relation to non-TFA standard certified teachers and other differently certified non-TFA teachers. Again we found that uncertified teachers and those with less than standard certification – whether TFA or non-TFA – exert negative effects on student achievement relative to teachers with standard certification.<sup>20</sup>

**Table 6**  
**Teacher Certification, TFA Status, and Student Achievement Gains**  
 4<sup>th</sup> and 5<sup>th</sup> Grade Students, All Available Years Pooled  
 Unstandardized Coefficients (T-Value)

	<b>TLI Math</b>	<b>TLI Reading</b>	<b>SAT-9 Math</b>	<b>SAT-9 Reading</b>	<b>Aprenda Math</b>	<b>Aprenda Reading</b>
TFA teacher, uncertified	-.251 (-.966)	-1.215*** (-3.813)	-1.630*** (-3.702)	-.966* (-2.295)	-6.120*** (-3.739)	-6.117*** (-3.759)
TFA teacher, alternatively certified	-.734 (-.700)	-1.358 (-1.047)	-5.757** (-3.317)	-2.616 (-1.573)	.704 (.304)	-2.747 (-1.195)
TFA teacher, standard certified	1.203*** (5.028)	.429 (1.464)	.015 (.031)	-.429 (-.931)	-2.450~ (-1.663)	.841 (.574)
Non-TFA teacher, uncertified	-.495*** (-7.439)	-.542*** (-6.635)	-.382** (-2.977)	-.586*** (-4.772)	-1.694*** (-5.407)	-.063 (-.203)
Non-TFA teacher, nonstandard certified	-.279 (-1.601)	-.601** (-2.797)	-1.508*** (-4.517)	-.012 (-.038)	-2.484*** (-3.579)	-1.178~ (-1.709)
<b>R Squared</b>	.43	.40	.62	.68	.42	.43
<b>N of Students</b>	105511	103122	60488	60607	11437	11436

~p<.10, \* p < .05, \*\* p < .01, \*\*\*p< .001

Uncertified TFA teachers have significant negative effects on student achievement in five of six estimates (and the sixth also has a negative coefficient.) The same is true of uncertified teachers who were not members of Teach for America. Alternatively certified TFA teachers – a very tiny share of the already small TFA sample -- have negative coefficients on five of six estimates, one of which is statistically significant. For non-TFA

<sup>20</sup> All TFA teachers had one of three certification types: uncertified, alternatively certified, and standard certified. We grouped all other teachers into comparable categories: uncertified, alternative plus all other nonstandard certification categories, and standard certified. We used the same control variables as in all of the previous analyses: student prior year test score, race/ethnicity, LEP status, and free / reduced price lunch status; teacher experience and highest degree; class size and class average previous year’s test score; and school racial/ethnic and income composition.

teachers, the category including alternative certification, along with emergency, temporary and other substandard certifications, also shows negative influences on achievement that are significant at the .10 level or below in four of six estimates.

Relative to other teachers with standard certification, TFA teachers with standard certification did about as well, with only two significant differences. As signaled in the CREDO analyses and ours, TFA teachers' students did significantly better than those of other standard certified teachers on the TAAS in mathematics. In year by year analyses, we found that the effect of standard certified TFA teachers was significantly positive in three years and non-significant in the other three years. On the other hand, TFA students did marginally worse than those of other standard certified teachers ( $p < .10$ ) on the Aprenda in mathematics.

In another analysis of these data, reported separately (Holtzman, 2005), hierarchical linear modeling (HLM) techniques were used to take into account the nesting of students within classrooms and of classrooms within schools. This analysis also found that 4<sup>th</sup> and 5<sup>th</sup> grade students taught by TFA teachers without standard certification made significantly less progress on each of the six tests than the students of standard certified teachers (controlling for teacher experience and degrees and student / school characteristics). In fact, in most estimates, the students of TFA teachers without certification did less well than the students of other uncertified teachers. In these analyses, where both fixed and random effects were modeled, students of certified TFA teachers made greater gains than those of other certified teachers in 4 estimates, smaller gains in 2 estimates, and comparable gains in 12 estimates.

Overall, then, the effects of TFA teachers on student achievement appear to depend, as they do for other teachers, primarily on the level of preparation these teachers have had, as reflected in their certification status.

### **Discussion**

Although a number of studies have found that students taught by fully certified teachers appear to achieve at higher levels, few have been able to examine individual student-level data over multiple years on multiple measures with appropriate controls. Previous studies of Teach for America, as a specific pathway into teaching, have either



failed to control for certification status and teachers' degrees or for students' prior achievement in examining the outcomes of this program on teachers' effectiveness. Our ability to look at these questions using a large data set that represents these teacher variables and a range of student, classroom, and school controls has provided a unique opportunity to evaluate how teacher education and pathways into teaching may influence teacher effectiveness.

Certification is, of course, only a proxy for the real variables of interest that pertain to teachers' knowledge and skills. These include knowledge of the subject matter content to be taught and knowledge of how to teach that content to a wide range of learners, as well as the ability to manage a classroom, design and implement instruction, and work skillfully with students, parents, and other professionals. In Texas, teachers who have achieved standard certification must have passed tests of core academic skills in communications and mathematics, specialized subject matter knowledge, and pedagogical knowledge. They must also have completed an approved teacher education program which includes specified courses in the content area(s) to be taught as well as coursework in: teaching and learning; instructional methods and strategies; classroom management; curriculum; measurement and evaluation of student learning; human growth and development; multicultural education; the education of special needs students; legal and ethical aspects of teaching; organization of schools; technology; and the teaching of reading (Texas Administrative Code, Title 19, Part 7, Rule 230.191, 2004).

This array of requirements, in combination, appears to make a difference in teacher effectiveness. Like other studies cited earlier, we find that 4<sup>th</sup> and 5<sup>th</sup> grade teachers in Houston who hold full certification – the professional or standard certificate Texas awards to recruits who have graduated from an approved teacher education program – are more effective than other teachers in stimulating student achievement gains on three different test measures in both reading and mathematics over a multi-year period. This relationship holds whether the teachers are recruited through Teach for America or through other pathways. Those who have completed the training that leads to certification are more effective than those who have not. And while Teach for America teachers who have achieved standard certification appear to be more effective than other certified teachers on one of the six measures we examined (the TAAS mathematics test) and somewhat less

effective on another (the Aprenda mathematics test), they generally perform on a par with other certified teachers, after controlling for degrees and experience, as well as a variety of student and school factors.

There is no instance where uncertified Teach for America teachers perform as well as standard certified teachers. On 5 of 6 tests, uncertified TFA teachers showed a significant negative effect on student achievement gains relative to standard certified teachers. (The sixth coefficient is also negative but non-significant.) Alternatively certified TFA teachers (of whom there were very few in our data base) showed negative coefficients on 5 of 6 tests, but only one of these was significant.

### **The Successes and Limitations of Teach for America in the Houston Context**

These data, in combination with the findings of the CREDO study, suggest both some successes and limitations of the Teach for America program in Houston, depending on the perspective used to assess effects. Across the country, Teach for America operates only in districts that, for a variety of reasons, hire many uncertified teachers. During the years studied, Houston was such a district, although there were improvements in the recruitment of certified teachers over the years studied (from about 56% of non-TFA teachers in 1996 to 67% in 2001).

The Teach for America organization often notes that its goal is to bring stability for at least one or two years to classrooms in poor and minority schools that might otherwise have a parade of substitute teachers, and argues that its recruits do as well as other teachers these students might have. Given the likelihood that these students would otherwise have equally inexperienced and uncertified teachers, this claim seems to be supported by our data. Entering TFA teachers appear to perform about as well as other uncertified teachers in Houston, after controlling for experience, degree status, and student characteristics. Most of them stay for two years, which may provide a modest degree of stability to schools that otherwise experience an even more quickly revolving door for teachers in and out of classrooms.

It might also be argued that the reputedly strong liberal arts background of TFA teachers may contribute to their relatively better showing on the TAAS mathematics tests. On this one test – though not on the other two mathematics tests used in Houston (the SAT-9 and the Aprenda) – the students of certified TFA recruits performed significantly

better than the students of other certified teachers. Given the longstanding concerns about the mathematics background of many elementary school teachers, it would not be implausible that candidates who have attended relatively selective colleges would have a stronger basic mathematics background than the average elementary teaching candidate. That this effect did not hold up on the SAT-9 and the Aprenda may be a function of differences in what the tests measure or of the strength of the TFA cohorts in the later years of our study, when the SAT-9 and Aprenda were administered.

The strength of TFA cohorts may also differ from year to year, as the program's recruitment and training practices fluctuate. The strongest positive TFA effects were in 1998-99, when TFA candidates were much more likely to be certified than the average Houston teacher. The most negative year was 2001-02, when TFA recruits had significant negative effects on student achievement on 5 of 6 tests. In addition to the fact that TFA teachers were much less likely than other Houston teachers to be certified in this year, there may have been other selection or training effects operating.

The limitations of teachers without preparation, including TFA teachers, are also illuminated by this study. It is clear that, across the board, Houston students achieved stronger achievement gains in both reading and mathematics when they were taught by standard certified teachers rather than uncertified teachers, including uncertified TFA teachers. Uncertified teachers were particularly disadvantageous to their students in reading (as measured by the TAAS and the SAT-9), a finding that makes sense given the specialized training that is typically needed to teach reading effectively to a wide range of students (National Reading Panel, 2000). This is consistent with the findings of other studies showing that teachers with more intensive teacher preparation in the teaching of reading produce greater achievement gains for their students (International Reading Association, 2003) and that alternatively certified teachers who have had less intensive teacher training produce significantly lower achievement gains in language arts than do traditionally-prepared teachers (Gomez & Grobe, 1990).

Uncertified teachers, both TFA and non-TFA, also did particularly poorly with Spanish-speaking students who took the Aprenda mathematics test. Even certified TFA teachers had a marginally negative effect on student learning gains on this measure (at the .10 level). This might be a function of the specialized knowledge needed to teach English

language learners that may be more consistently acquired in teacher education programs that have the time to teach not only the basics of classroom management and lesson planning, but also the strategies for teaching content to students who have particular language needs.

Finally, although students taught by TFA recruits (and other uncertified teachers) were slowed in their academic progress in the first year of the recruits' teaching efforts, one of the accomplishments of Teach for America and HISD was the development of means for enabling recruits to participate in preparation and become certified for teaching in their 2<sup>nd</sup> or 3<sup>rd</sup> year of teaching. However, Houston students did not reap long-term benefits from these efforts, as the vast majority of the TFA recruits left after their second or third year of teaching. High turnover of beginning teachers is both extremely costly for school districts (Benner, 2000) and counterproductive for students, as teacher effectiveness typically increases markedly after about the second year of teaching (Hanushek, Kain, & Rivkin, 1998).

A challenge for states, school districts, and teacher preparers is how to develop and expand the reach of strong, efficient, and affordable preparation routes that enable teachers to be competent when they enter teaching and that retain teachers as they become more effective. The literature includes examples of urban teacher education programs that have strong records of preparing capable teachers who stay in the city schools (see for example, Darling-Hammond, Chung, & Frelow, 2002; Darling-Hammond & Macdonald, 2000; Snyder, 2000; Zeichner, 2000). Increasing the availability of such programs could help stem turnover, as several recent studies have found that teacher attrition is strongly related to the extent of preparation teachers have had upon entry (Henke, Chen, Geis, & Knepper, 2000; NCTAF, 2003). However, analyses of urban districts that have resolved teacher shortages indicate that additional state and local policies are needed to create the labor market conditions required to hire and retain an adequate supply of prepared teachers. These include competitive salaries, reasonable working conditions, and supportive administrators, as well as training subsidies, recruitment incentives, and hiring reforms (Darling-Hammond & Sykes, 2003; Murnane et al., 1991; NCTAF, 2003).

## References

- Ballou, D. and Podgursky, M. (2000). Reforming teacher preparation and licensing: What is the evidence? Teachers College Record, 102 (1), pp. 1-27.
- Benner, A. D. (2000). The cost of teacher turnover. Austin, TX: Texas Center for Educational Research.
- Betts, J.R., Rueben, K.S., Danenberg, A. (2000). Equal resources, equal outcomes? The distribution of school resources and student achievement in California. San Francisco: Public Policy Institute of California.
- Darling-Hammond, L. (2000a). Reforming teacher preparation and licensing: Debating the evidence. Teachers College Record, 102 (1), pp. 28-56.
- Darling-Hammond, L. (2000b). Teacher quality and student achievement: A review of state policy evidence. Educational Policy Analysis Archives, 8, (1). <http://epaa.asu.edu/epaa/v8n1>.
- Darling-Hammond, L., Chung, R., & Frelow, F. (2002). Variation in teacher preparation: How well do different pathways prepare teachers to teach? Journal of Teacher Education, 53 (4): 286-302.
- Darling-Hammond, L., & MacDonald, M. (2000). Where there is learning there is hope: The preparation of teachers at the Bank Street College of Education. In L. Darling-Hammond (Ed.), Studies of excellence in teacher education: Preparation at the graduate level (pp. 1-95). Washington, DC: American Association of Colleges for Teacher Education.
- Darling-Hammond, L. & Sykes, G. (2003). Wanted: A national teacher supply policy for education: The right way to meet the 'highly qualified teacher' challenge. Educational Policy Analysis Archives, 11 (33). <http://epaa.asu.edu/epaa/v11n33/>.
- Darling-Hammond, L. & Youngs, P. (2002). Defining "highly qualified teachers:" What does "scientifically-based research" actually tell us? Educational Researcher, 31 (9): 13-25.
- Decker, P.T., Mayer, D.P., & Glazerman, S. (2004). The Effects of Teach For America on Students: Findings from a National Evaluation. Princeton, NJ: Mathematica Policy Research, Inc.
- Ferguson, R.F. (1991, Summer). Paying for public education: New evidence on how and why money matters. Harvard Journal on Legislation 28(2), pp.465-498.
- Fetler, M. (1999). High school staff characteristics and mathematics test results. Education Policy Analysis Archives, 7(9): <http://epaa.asu.edu/epaa/v7n9/>

Finn, C.E. (1999). Foreword. In M. Kanstoroom & C.Ed. Finn, Jr. (eds.), Better teachers, better schools. Washington, DC: The Thomas B. Fordham Foundation.

Goe, L. (2002). Legislating equity: The distribution of emergency permit teachers in California. Educational Policy Analysis Archives, 10 (42): <http://epaa.asu.edu/epaa/v10n42/>

Goldhaber, D.D., & Brewer, D.J. (2000). Does teacher certification matter? High school teacher certification status and student achievement. Educational Evaluation and Policy Analysis, 22 (2): 129-146.

Gomez, D. L., and Grobe, R. P. (1990). Three Years of Alternative Certification in Dallas: Where are we? Paper presented at the Annual Meeting of the American Educational Research Association, Boston, MA.

Hanushek, E., Kain, J., & Rivkin, S. (1998). *Teachers, schools, and academic achievement*. (Working Paper 6691). Cambridge, MA: National Bureau of Economic Research.

Hawk, P.P., Coble, C.R., & Swanson, M. (1985). Certification: It does matter. *Journal of Teacher Education*, 36(3), 13-15.

Henke, R.R., Chen, X., Geis, S., & Knepper, P. (2000). Progress through the teacher pipeline: 1992-93 college graduates and elementary/secondary school teaching as of 1997. NCES 2000-152. Washington, DC: National Center for Education Statistics.

Holtzman, D. (2005). Teach for America, Teacher Certification, and Student Achievement in Houston, Texas in 2001-2002: A Multilevel Analysis. Stanford, CA: Stanford University.

International Reading Association (2003). Prepared to make a difference: Research evidence on how some of America's best college programs prepare teachers of reading. Newark, NJ: International Reading Association.

Klein, S.P., Hamilton, L.S., McCaffrey, D.F., & Stetcher, B.M. (2000). What do test scores in Texas tell us? Santa Monica: The RAND Corporation.

Laczko-Kerr, I., & Berliner, D. (2002). The effectiveness of Teach for America and other under-certified teachers on student academic achievement: A case of harmful public policy. *Educational Policy Analysis Archives*, 10 (37). <http://epaa.asu.edu/epaa/v10n37>

Monk, D.H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of Education Review*, 13(2), 125-145.

Murnane, R., Singer, J., Willett, J., Kemple, J., & Olsen, R. (1991). Who will teach? Policies that matter. Cambridge, MA: Harvard University Press.

National Commission on Teaching and America's Future. (NCTAF) (1996). What matters most: Teaching for America's future. New York: Author.

National Commission on Teaching and America's Future. (NCTAF) (2003). No dream denied. Washington, DC: Author.

National Reading Panel (2000). Teaching Children to Read: An Evidence-based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction. Washington, DC: National Institute of Child Health and Human Development.

Raymond, M. & Fletcher, S. H. (2002). The Teach for America evaluation. Education Next, 2 (1): 62-68.

Raymond, M., Fletcher, S.H., & Luque, J. (2001). Teach for America: An evaluation of teacher differences and student outcomes in Houston, Texas. Stanford, CA: The Hoover Institution, Center for Research on Education Outcomes.

Snyder, J. (2000). Knowing children -- understanding teaching: The developmental teacher education program at the University of California, Berkeley. In L. Darling-Hammond (Ed.), Studies of excellence in teacher education: Preparation at the graduate level (pp. 97-172). Washington, D.C.: American Association of Colleges for Teacher Education.

Strauss, R. P. and Sawyer, E.A. (1986). Some New Evidence on Teacher and Student Competencies. Economics of Education Review, 5 (1): 41-48.

U.S. Department of Education (2002). The Secretary's report on teacher quality. Washington, DC: U.S. Department of Education.

Wenglinsky, H. (2000). *How teaching matters: Bringing the classroom back into discussions of teacher quality.* Princeton, NJ: Educational Testing Service.

Wilson, S.M., Floden, R., & Ferrini-Mundy, J. (2001). Teacher preparation research: Current knowledge, gaps, and recommendations. A research report prepared for the U.S. Department of Education. Seattle: Center for the Study of Teaching and Policy, University of Washington.

Zeichner, K. M. (2000). Ability-based teacher education: elementary teacher education at Alverno College. In L. Darling-Hammond, (Ed.), *Studies of excellence in teacher education: Preparation in the undergraduate years* (pp. 1-66). Washington, D.C.: American Association of Colleges for Teacher Education Publications.

## Appendix

**Table A-1**  
Certification Status of 4<sup>th</sup> and 5<sup>th</sup> Grade Teachers, by Year

	1996-97*		1997-98*		1998-99		1999-00		2000-01		2001-02	
	Non-TFA	TFA	Non-TFA	TFA	Non-TFA	TFA	Non-TFA	TFA	Non-TFA	TFA	Non-TFA	TFA
<b>Standard</b>	56.5% (1028)	65.2% (15)	61.6% (1089)	68.6% (24)	65.4% (1041)	73.1% (19)	67.8% (1181)	48.3% (14)	68.7% (1216)	45.7% (16)	64.8% (1137)	45.2% (19)
<b>Alternative</b>	0.2% (4)	0.0% (0)	0.3% (5)	0.0% (0)	0.3% (5)	3.8% (1)	0.3% (5)	0.0% (0)	0.6% (11)	8.6% (3)	0.7% (12)	4.8% (2)
<b>Emergency/ Temporary</b>	9.5% (172)	0.0% (0)	8.7% (154)	0.0% (0)	5.8% (92)	0.0% (0)	3.9% (68)	0.0% (0)	2.7% (48)	0.0% (0)	2.0% (35)	0.0% (0)
<b>Certified Out- of-field</b>	0.1% (2)	0.0% (0)	0.2% (3)	0.0% (0)	0.3% (4)	0.0% (0)	0.3% (5)	0.0% (0)	1.0% (17)	0.0% (0)	2.2% (38)	0.0% (0)
<b>Certified, No test</b>	2.3% (42)	0.0% (0)	1.2% (22)	0.0% (0)	0.6% (9)	0.0% (0)	0.9% (16)	0.0% (0)	0.8% (15)	0.0% (0)	1.0% (18)	0.0% (0)
<b>Not yet certified</b>	20.4% (370)	34.8% (8)	20.1% (355)	31.4% (11)	23.0% (366)	23.1% (6)	23.2% (404)	51.7% (15)	22.1% (391)	45.7% (16)	25.8% (453)	50.0% (21)
<b>Certification missing</b>	11.0% (200)	0.0% (0)	7.9% (139)	0.0% (0)	4.7% (74)	0.0% (0)	3.6% (62)	0.0% (0)	4.1% (72)	0.0% (0)	3.5% (61)	0.0% (0)

\* In these years, the HISD data set included a number of candidates with certification codes but not certification dates. They are listed in this table in the category indicated by their certification code for each year that they appear in the data set.

Note: Includes all teachers who taught any 4<sup>th</sup> or 5<sup>th</sup> grade student



**Table A-2: TFA Status and Student Achievement**  
 4<sup>th</sup> and 5<sup>th</sup> Grade Students, All Available Years Pooled  
 Unstandardized Coefficients (T-Value)

	<b>TLI Math</b>	<b>TLI Reading</b>	<b>SAT-9 Math</b>	<b>SAT-9 Reading</b>	<b>Apranda Math</b>	<b>Apranda Reading</b>
Constant	30.39*** (77.92)	31.69*** (66.24)	15.41*** (27.67)	9.86*** (18.37)	25.33*** (6.60)	7.82* (2.05)
Previous year test score	.498*** (201.92)	.526*** (184.37)	.683*** (223.72)	.743*** (240.46)	.670*** (81.54)	.706*** (83.36)
Free / reduced price lunch status	-.920*** (-13.74)	-1.41*** (-17.19)	-1.70*** (-12.36)	-1.68*** (-12.76)	.693 (.878)	.978 (1.25)
American Indian	.136 (.160)	.129 (.123)	1.72 (.967)	-.278 (-.165)	...	...
Asian / Pacific Islander	.679*** (4.51)	.543** (2.95)	2.32*** (7.79)	.609* (2.14)	9.90 (1.39)	10.01 (1.42)
African American	-1.65*** (-17.16)	-2.23*** (-18.94)	-3.03*** (-15.39)	-2.48*** (-13.15)	-2.91 (-.545)	11.19~ (1.84)
Hispanic	-.666*** (-6.99)	-1.33*** (-11.34)	-1.21*** (-6.23)	-1.85*** (9.92)	-.942 (-.287)	4.40 (1.35)
Limited English Proficient	1.590*** (20.78)	.667*** (7.12)	.883*** (6.33)	.870*** (6.49)	2.01*** (3.59)	1.34* (2.41)
Total teaching experience	-.010*** (-3.95)	.001 (.286)	.037*** (7.24)	.057*** (11.71)	.136*** (7.66)	.035* (1.98)
No degree	.035 (.156)	-.575* (-2.12)	1.06** (2.89)	1.69*** (4.83)	3.19*** (6.88)	1.10* (2.39)
Masters degree	-.282*** (-4.79)	-.104 (-1.44)	-.290* (-2.45)	-.340** (-3.00)	-.644~ (-1.73)	-.089 (-.239)
Doctoral degree	.907*** (4.07)	-.053 (-.192)	-.042 (-.105)	-.328 (-.862)	1.03 (1.01)	3.36** (3.32)
Classroom number of students	.149*** (25.06)	.020** (2.68)	-.029** (-2.64)	.020~ (1.91)	.118*** (4.25)	.183*** (6.64)
Class average previous year test score	.116*** (25.72)	.149*** (28.32)	.100*** (17.22)	.124*** (21.33)	-.049** (-2.63)	-.061** (-3.20)
% African American students in school	.010*** (4.24)	-.003 (-1.07)	.003 (.587)	-.023*** (-5.12)	-.147*** (-6.79)	.012 (.565)
% Hispanic students in schools	.013*** (5.28)	-.002 (-.570)	-.010* (-2.08)	-.030*** (-6.59)	-.129*** (-6.41)	-.004 (-.209)
% free/ reduced price lunch students in school	-.003** (-2.68)	.001 (.761)	-.001 (-.677)	.005** (2.66)	.046*** (6.19)	.004 (.499)
TFA teacher status	.696*** (4.02)	-.056 (-.265)	-.840** (-2.60)	-.575~ (-1.86)	-2.39* (-2.39)	-2.37* (-2.39)
<b>R Squared</b>	.43	.39	.62	.68	.42	.43
<b>N of Students</b>	105511	103122	60488	60607	11437	11436

#####