Teacher Characteristics and Student Achievement: Evidence from Teacher Surveys

Christopher Jepsen
Public Policy Institute of California
500 Washington Street, Suite 800
San Francisco, 94111
jepsen@ppic.org

Abstract:

Teachers and peers are believed to have a strong influence on student achievement, but the specific characteristics that affect student achievement are hard to identify. This paper utilizes teacher survey data to investigate teacher characteristics that are not usually available in administrative data, as well as more readily available attributes such as experience and education. Classroom fixed effects explain a large portion of within-student variation in test score growth, suggesting a potentially important role for teachers and peers. Teacher characteristics are generally insignificant predictors of student achievement, especially for the lower grades.

Introduction

Teacher quality is an important concern of parents and policymakers, as illustrated by the provision in No Child Left Behind (NCLB) requiring a qualified teacher in every classroom. But the definition of qualified has come under much scrutiny. For example, there is intense debate over whether alternatively certified teachers should be included under the heading of qualified teachers.

The challenge in defining quality teachers in NCLB illustrates a much more general concern in education: the inability to define and measure teacher quality. Recent evidence using administrative data suggests that the easily observable teacher characteristics such as experience, education, and certification are not strongly associated with student achievement in elementary school.[1] Rivkin, Hanushek and Kain [19] find that teachers in their first or second year of teaching are associated with lower student test scores in Texas, but teacher education and certification have no systematic relationship with achievement. Jepsen and Rivkin [16] obtain similar results using grade-level data from California. Preliminary results from Clotfelter, Ladd, and Vigdor [6] suggest positive impacts of teacher experience and teacher licensure test scores on student achievement in North Carolina. Betts, Zau, and Rice [4] find mixed results for teacher characteristics using detailed individual-level data from elementary schools in the San Diego Unified School District.

The lack of significant effects for these teacher characteristics should not be interpreted as evidence that teachers have no impact on student achievement. Teacher quality, measured by

---

[1] Although not discussed in detail here, there is also a substantial literature of the effects of teacher characteristics on achievement in secondary school. For example, see Ehrenberg and Brewer [8], Ehrenberg, Goldhaber, and Brewer [9], Goldhaber and Brewer [10], [11], Betts, Zau, and Rice [4], and Aaronson, Barrow, and Sander [1].

teacher fixed effects, has an important impact on student achievement in Rockoff [20]. In addition, Hanushek [12], Murnane [18], and Armor et al [3] find significant impacts of classroom fixed effects (i.e. combined impact of teachers and peers). Rivkin, Hanushek, and Kain [19] find large effects for overall teacher effects measured at the grade level. In other words, teacher quality may be important, but it is not well captured by levels of teacher experience, certification, and education.

This paper extends the literature by utilizing an extensive teacher survey to look at the relationship between teacher characteristics and student achievement.[2] In addition to the teacher characteristics readily available, such as experience, the survey also contains information on teacher enthusiasm, classroom resources, computer use, and amount of homework assigned. The teacher data are linked with student panel data for over 10,000 elementary school students in each of two cohorts.

The first part of the analysis considers the potential impact for teachers. Fixed effects at the classroom level explain roughly 25 to 40 percent of the within-student variation in test score growth. These sizable effects – consistent with earlier findings in the literature – suggest an important role for peers and teachers. However, the results show few significant links between student achievement and observable teacher characteristics, especially for the first-grade cohort. For the fourth-grade cohort, there is weak evidence that teacher experience is positively related to student achievement. Student achievement is not adversely affected if a teacher is not fully

---

[2] In many ways, this paper is similar to the analysis at the secondary school level in Goldhaber and Brewer [10]. They use NELS data to look at an extensive set of teacher characteristics for secondary school students. However, the elementary school data (called *Prospects*) have several advantages over NELS. For example, *Prospects* data include yearly observations for up to four years, but NELS data are only available every two years. Further, elementary school students typically have only one or two teachers, whereas high school students usually have one teacher for each subject.

credentialed or if a teacher does not have post-graduate education. The results for more detailed teacher survey items vary by subject and grade level. For example, teacher enthusiasm is negatively associated with classroom effectiveness in mathematics, but the effects for reading are insignificant. Overall, the results demonstrate that the attributes of high-quality teachers are difficult to measure.

Data

The data for this paper come from *Prospects*, the Congressionally mandated national study of the Title 1 program (formerly known as Chapter 1), which provides additional resources for at-risk students. The *Prospects* database contains two cohorts of elementary school students: first-grade students in the fall of 1991 and third-grade students in the spring of 1991. The database for each cohort of *Prospects* students is quite extensive. Each cohort originally contains over 200 schools and nearly 10,000 students. Public school districts and schools were carefully chosen to provide a representative sample of the schools and students receiving Title 1 funding, rather than to provide a nationally representative sample of *all* primary schoolchildren.[3] Within each school, if grade enrollment is under 150, all students in the grade are included in the baseline survey in 1991. Otherwise, every student in four classrooms (chosen at random) is chosen. Follow-up data for each student were collected in each spring through 1994.[4]

Students, parents, teachers, school administrators, principals, and district Title 1 coordinators completed questionnaires in the springs of 1991 (third-grade cohort only), 1992, 1993, and 1994. Furthermore, students took the fourth edition of the Comprehensive Test of Basic Skills (CTBS/4) in reading and mathematics in the springs of 1992, 1993, and 1994.

---

[3] More information on the sampling technique for public schools in *Prospects* is available in Bryant [5].

Students in the first-grade cohort also took tests in the fall of 1991 (i.e. in the fall of first grade), whereas students in the third-grade cohort took tests in the spring of 1991 (i.e. in the spring of third grade). These tests are reported as scale scores, in order to allow for comparisons across cohorts and years. In other words, a score of 500 for a second-grade student in 1992 measures the same absolute level of academic achievement as a score of 500 for a different student in another grade or during another school year. Two interesting trends in the test scores– both nationally and in the *Prospects* data – are worth noting. First, the standard deviation in test scores drops as the grade level increases. Second, the average increase in mean test scores between grades also drops as the grade level increases.[5] Unique student identifiers allow researchers to link the various data sources together to create a comprehensive student-level database.

As with all data, *Prospects* data have their strengths and weaknesses. The yearly panel of test scores from a low-stakes test in mathematics and reading is an advantage compared to state-specific high-stakes tests where teachers have incentive to "teach to the test."[6] The rich set of teacher characteristics extends far beyond the education, experience, and certification data that is the extent of most administrative data. For example, teachers also answered survey items on teacher attitudes, classroom materials, and teaching methods.[7]

The structure of the data is better suited to analysis at the classroom level rather than the teacher level. Nearly all teachers are observed in only one classroom, so the measures of teacher effectiveness are based on only one classroom. If each teacher were observed with several

---

[4] Due to financial constraints, only a subset of students who left the original schools completed follow-up surveys. Similarly, only a subset of students new to the original schools was surveyed. See Bryan [5] for more information.
[5] For more information on the test scores, see CTB [7].
[6] NELS data are available every two years, and only contain two observations in high school (grades 10 and 12).
[7] Unfortunately, the data on teacher major were not asked in all survey years.

different classrooms (as in Rockoff [20]), the effect of the teacher could be separated from the effect of the students (i.e. the peer effect). With only one classroom per teacher, such separation is not possible. Instead, the observed classroom effect is a combination of the teacher and the peer effects. Although the data contain a sizable number of classrooms, the sample size is far below what is available in state administrative data.

Table 1 contains the descriptive statistics for teacher characteristics.[8] Nearly all students in the first-grade cohort have the same teacher for mathematics and for reading, whereas almost forty percent of students in the third-grade cohort have different teachers for the two subjects. Average experience is around 14 years for both cohorts and subjects, but nearly 10 percent of teachers are in the first or second year of teaching. Over one quarter have only a bachelor's degree, and over one third have a master's degree or more. Very few teachers (6 to 8 percent) are not fully certified.

Method

Overall Classroom Effects

Much of the literature on teacher characteristics focuses on the relationship between specific teacher attributes (such as experience or education) and student achievement. In general, the relationship between teacher characteristics and student achievement is weak, with few consistent findings (see Hanushek [13]). The typical conclusion is that the attributes of a high-quality teacher are hard to measure. Implicit in this conclusion is that teachers are strong predictors of student achievement, but few papers other than Rockoff [20] and Rivkin, Hanushek, and Kain [19] have attempted to test this assumption.[9] Using detailed teacher and

---

[8] All descriptive statistics and student-level regressions are not weighted.
[9] Again, an exception at the high school level is Goldhaber and Brewer [10].

student data from two contiguous districts in New Jersey, Rockoff [20] finds that individual

teachers are strongly correlated with student achievement, but specific teacher attributes are

much weaker, and – aside from teacher experience – are insignificant predictors. Rivkin,

Hanushek, and Kain [19] cannot match teachers with students, but their analysis of grade-level

data in Texas reaches similar conclusions.

As mentioned above, separating the effects of teachers from those of peers is complicated

in *Prospects* data, so the analysis in this section considers the combined effect of teachers and

peers.[10] Specifically, the following equation is estimated for each cohort and subject:

(1) $Y_{ijst} = \alpha_i + C_{jst} + \varepsilon_{ijst}$

In this student-level equation, *Y* is the growth in test score, where growth has been standardized

to a mean of zero and a standard deviation of one. Consequently, the classroom fixed effects (*C*)

should be interpreted in terms of their impact on standard deviations of test score growth. The

model includes student fixed effects (*α*) to control for systematic (i.e. time invariant) student-

specific attributes such as race and ethnicity. In addition, *i* denotes students, *j* denotes

classrooms, *s* denotes schools, and *t* denotes years.

Figure 1 illustrates the distribution of the t-statistics for classroom fixed effects in

mathematics. The figure illustrates a well-dispersed distribution of classroom fixed effects, both

positive and negative. The distribution of reading classroom fixed effects is similar, as shown in

Appendix Figure 1. Table 2 contains descriptive statistics for the classroom fixed effects. The

table shows that 10.6 to 13.8 percent of the classroom effects are significantly different from

zero at the five-percent level. This number is small, but it is well above five percent, the

---

[10] Hanushek[12], Murnane [18], and Armor et al [3] face the same data constraints of observing teachers with only
one set of students.

expected percentage of significant fixed effects if these classroom effects were distributed randomly. Despite the low percentage of significant fixed effects, the classroom effects explain a substantial portion of the within-student variation in test score growth. The partial R-squared for the classroom fixed effects (not the student fixed effects) ranges from 0.269 in mathematics for the third-grade cohort to 0.414 in reading for the first-grade cohort. The R-squared for first-grade cohort is higher than for third-grade cohort, likely because a higher percentage of students in the first-grade cohort have the same teacher and classmates for the entire school day (as shown in Table 1).

Specific Teacher Characteristics

The above analysis suggests that teachers and peers are important factors in student achievement, but it says nothing about how teachers affect achievement. Several studies consider the relationship between specific teacher characteristics and student achievement. At the elementary school level, the previous studies rely on administrative data. The detail on the *Prospects* teacher survey permits an analysis of teacher characteristics generally not available in administrative data. Survey data on elementary school teachers are extremely rare, as most national surveys (such as High School and Beyond, National Educational Longitudinal Survey of 88, Adolescent Health, or the National Longitudinal Surveys) focus on individuals in middle and high school.

The approach used to study teacher effectiveness uses classroom level data, building on the classroom fixed effect specification in equation (1). In this specification, classroom effects are the dependent variable (i.e. the measure of teacher effectiveness), and the independent variables are teacher and student characteristics:

(2) $\quad C_{jst} = TC'_{jst}\beta + P'_{jst}\gamma + \rho_s + \mu_t + \eta_{jst}.$

C_{jst} is the classroom effect, TC_{jst} is a vector of teacher characteristics for each teacher, and P_{jst} is the effect of classmates, measured as the average lagged student achievement for the classroom. The equation contains school (ρ_s) and year (μ_t) fixed effects. As in equation (1), j denotes classrooms, s denotes schools, and t denotes years. For simplicity, I combine TC_{jst}, P_{jst}, ρ_s, and μ_t into one vector (Z_{jst}), so that the equation can be re-written as:

(3) $\qquad C_{jst} = Z'_{jst}\lambda + \eta_{jst}$.

      The true classroom effect (C_{jst}) is not observed, but it can be estimated from equation (1). The relationship between the true and estimated classroom effects is $C_{jst} = \hat{C}_{jst} + \pi_{jst}$, where $E(\pi_{jst}) = 0$ and $E(\pi_{jst}^2) = \sigma_{\pi j}^2$. Using these predict classroom effects, the following equation can be estimated:

(4) $\qquad \hat{C}_{jst} = Z'_{jst}\phi + \omega_{jst}$.

Given the relationship between C_{jst} and $\hat{C}_{jst}$, the estimated equation becomes:

(5) $\qquad \hat{C}_{jst} = C_{jst} - \pi_{jst} = (Z'_{jst}\phi + \eta_{jst}) - \pi_{jst} = Z'_{jst}\phi + (\eta_{jst} - \pi_{jst}) = Z'_{jst}\phi + v_{jst}$.

Under the assumption that η_{jst} and π_{jst} are independent (for all j, s, and t), $E(v_{jst}) = 0$ and $E(v_{jst}^2) = \sigma_\eta^2 + \sigma_{\pi j}^2$. GLS is estimated to account for the heteroskedasticity in v_{jst}. Specifically, I use weighted least squares with weights equal to $\sqrt{\dfrac{1}{\hat{\sigma}_\eta^2} + \dfrac{1}{\hat{\sigma}_{\pi j}^2}}$. The predicted value for $\hat{\sigma}_{\pi j}^2$ is the square of the standard error of $\hat{C}_{jst}$, derived from the first-stage regression in equation (1). The predicted value for $\hat{\sigma}_\eta^2$ is $\dfrac{RSS_5 - \sum_j \hat{\sigma}_{\pi j}^2}{N_5 - K_5}$, where RSS_5 is the residual sum of squares and N_5-K_5 is the degrees of freedom, all obtained from OLS estimation of equation (5).

A total of four regressions are estimated: one for each subject (mathematics and reading) and cohort (first grade and third grade). The goal of equation (5) is to investigate the determinants of effective classrooms, where effectiveness is defined as the classroom fixed effect. The main advantage of this approach is that the dependent and independent variables all are defined at the classroom level. State accountability systems often define success in terms of achievement at the classroom or school level, rather than in terms of individual achievement.

Because the model contains school fixed effects (ρ_s in equation (2)), identification comes from within-school differences in teacher characteristics rather than from between-school differences. Hoxby [15] argues that using class size variables that do not control for between-school variation are biased given their likely correlation with unobserved changes in family background, and that logic applies to other teacher and classroom attributes. Year fixed effects control for time trends.

One concern with the analysis is that the assignment of teachers to classroom within a school is not necessarily random. Preliminary evidence from Clotfelter, Ladd, and Vigdor [6] suggests that the assignment of elementary school students within schools in North Carolina is based on student and teacher characteristics, whereas Rockoff [20] cannot reject the hypothesis of random assignment in his data from two contiguous New Jersey districts.

Appendix Tables 1 and 2 contain information on the within-school placement of teachers in *Prospects* data. The tables contain regression results from a series of regressions where teacher characteristics are the dependent variables, and student characteristics (measured at the classroom level) are the independent variables. Appendix Table 1 contains the results for the first-grade cohort, whereas Appendix Table 2 contains the results for the third-grade cohort. The findings from the tables show little systematic relationship between student attributes and teacher

placement. The only consistently significant result is the negative correlation between prior year test scores and the likelihood of an aide in the classroom. In other words, classrooms with high test scores in the previous year are less likely to have an aide. The results for teacher experience depend on the cohort: classrooms with high previous test scores have less-experienced teachers in the first-grade cohort, but they have more-experienced teachers in the third-grade cohort. Teacher education, certification, and computer use are not related to student characteristics. The results for teacher-reported enthusiasm and adequate classroom materials depend on the cohort and subject. There is no evidence on non-random placement of teachers with respect to student income or race/ethnicity.

Prospects data are not ideal for a classroom-level model, as the database was designed to follow students rather than classrooms or teachers. Entire classrooms are only surveyed in the first year of the analysis (spring of 1991 for the third-grade cohort; fall of 1991 for the first-grade cohort). Although the students in these classrooms are followed up in subsequent years, each classroom in these years contains Prospects students and other students. Because the estimated classroom effects are based only on the Prospects students, the observed classroom effects are estimates of the true classroom effects. In order to test the robustness of the results, I also re-estimate the second-stage equation where I exclude classroom effects that are based on fewer than six Prospects students. The results are generally similar to those reported in Tables 3 and 4; notable differences are discussed in the next section.

An alternative to the methods used here is hierarchical linear model (HLM), sometimes called a mixed model, which includes both fixed and random effects. These types of models are becoming increasingly popular in education because their error structure explicitly accounts for the nesting of observations in groups or hierarchies. Students are nested in classrooms, which

10

are in turn nested in schools. However, this structure generally does not control for systematic

variation. Instead, it parses the variation between levels, so that the researcher can estimate the

amount of variation that occurs within students, within classrooms (but between students), and

within schools (but between classrooms). In other words, the approach assumes that all the

variance within a classroom (but between students) is due to classroom-specific factors, whereas

a substantial portion is likely due to omitted student-level factors such as parental motivation.

Due to these concerns, an HLM approach is not utilized here.[11]

Results

Table 3 contains the results for the first-grade cohort from the classroom level model in

equation (5). The dependent variable is the classroom fixed effect in either mathematics (the

first three columns) or reading (the second three columns). For each test, the first column

contains controls typically available in administrative data: class size, teacher experience,

education, and certification. The second column adds teacher survey items, and the third also

adds the lagged test score (in that subject) of the students in that classroom. Additional student

attributes (measured at the classroom level), such as the percentage of nonwhite students, are too

highly correlated with lagged achievement to identify a separate effect. Coefficients for teacher

characteristics from a model including additional student attributes were nearly identical to those

that contain only one student attribute, the lagged test score.

The general finding from the second-stage model is that teacher characteristics are not

strongly associated with teacher effectiveness as measured by the classroom fixed effect. This

finding is especially strong for reading effects, where none of the teacher characteristics is

---

[11] Rowan, Correnti, and Miller [21] use HLM analysis to examine *Prospects* data, but it is unclear how to interpret
their results given the concerns with HLM models.

11

significantly different from zero at the ten-percent level. Teacher enthusiasm (measured on a three-point scale of how much the teacher looks forward to each day) is negatively associated with the mathematics fixed effect, with a modest effect of around 0.05 standard deviations. This effect could be the result of systematic assignment of enthusiastic teachers to "difficult" classrooms, but the results in Appendix Table 1 show no relationship between student characteristics and teacher enthusiasm for mathematics teachers in the first-grade cohort. However, this effect becomes insignificant when the sample is limited to classrooms with six or more *Prospects* students. Thus, this coefficient should be interpreted with caution.

In contrast, lagged class test scores are extremely strong predictors of classroom effectiveness.[12] A one-standard deviation increase in classroom test scores (about 80 points) corresponds with a *decrease* in classroom effectiveness of over 0.5 standard deviations. This negative impact is consistent with a "mean reversion" explanation. Students who have large gains in one year have smaller gains in the following year (and vice versa). The measure of peer test score is based on the subset of the current classmates who took the test in the previous year, so it is only an estimate of the actual peer test score. Hanushek et al [14] express concerns about the exogeneity of peer effects measured by one-year lags in test scores, so the lagged peer test score coefficient in Table 3 should be interpreted with caution.

The findings for the third-grade cohort, reported in Table 4, show slightly stronger relationships between teachers and student achievement. When lagged peer test scores are included in the model (columns (3) and (6)), teacher experience is positively associated with classroom effectiveness, with an effect of 0.02 standard deviations for each year of experience. For mathematics, the returns to experience are decreasing, as the experience squared term is

---

[12] In contrast, Rockoff [20] finds insignificant effects of lagged test scores at the classroom level.

negative and significant. The overall effect[13] of a one-standard deviation increase in teacher experience (roughly nine years) corresponds with an increase in classroom effectiveness of 13 percentage points in mathematics and 8 percentage points in reading. Even though these results explicitly control for lagged test scores, they may overstate the causal effect of experience, as more experienced teachers may be assigned to classrooms with better students (Appendix Table 2). As with the first-grade cohort, lagged classroom level test scores are highly significant, negative predictors of classroom fixed effects.

The effects for the teacher survey items are different for mathematics and for reading. For mathematics, teachers who believe that they have adequate classroom materials are associated with more effective classrooms, with an effect of approximately 0.15 standard deviations.[14] The results reported in Appendix Table 2 show no evidence that teachers' perceptions of adequate materials are dependent on the characteristics of their students. The classroom materials variable includes materials provided by the teacher or by students in addition to those provided by the school or district. Thus, the effect could be a proxy for motivated teachers or parents who provide additional materials when needed.

Increased homework is associated with a modest increase in mathematics achievement, although the effect is only significant when controlling for lagged peer test scores. Also for mathematics, enthusiastic teachers are associated with slightly lower effectiveness in mathematics of approximately 0.05 standard deviations. However, this effect is marginally significant at the ten percent level once I control for lagged peer test scores. This effect is not

---

[13] The overall effect is the sum of the linear and quadratic terms for teacher experience.
[14] Although it is a common belief that classroom materials are even distributed across classrooms within a school, personal communication with teachers suggests that some classrooms are much more appealing physically than others.

the result of assigning more enthusiastic teachers to low-achieving classrooms, however, as more enthusiastic teachers are assigned to classrooms with better students (Appendix Table 2). Instead, the finding suggests that mathematics teachers are more content but less effective with high-ability classrooms.

As in the first-grade cohort, the effects of teacher survey items are weakly associated with reading test scores. The only significant effect is the negative impact of increased computer use (measured on a five-point scale) of 0.07 standard deviations. The effect of adequate materials is sizable and positive, but it is imprecisely estimated and therefore is not statistically significant from zero at the ten-percent level.

The specifications at the classroom level do not demonstrate a strong relationship between teacher characteristics and student achievement. The relationship is particularly weak for the first-grade cohort. For both grade levels, teacher enthusiasm is negatively associated with classroom effectiveness in mathematics. This effect does not merely represent the assignment of enthusiastic teachers to low-achieving classrooms, but it may represent other factors besides enthusiasm. Lagged classroom test scores also have a negative impact for both subjects and cohorts. One conclusion is that the determinants of classroom effectiveness vary across grade levels and subjects. An alternative explanation is that there is insufficient variation in teacher characteristics after controlling for student and school fixed effects.

Conclusion

There is a general belief that teachers are extremely important determinants of student achievement, but identifying specific teacher attributes that improve student performance is a difficult task. For years, data limitations hampered researchers' ability to identify exogenous variation in specific teacher characteristics. The recent availability of detailed administrative

data in Texas, San Diego, North Carolina, and elsewhere provides new opportunities to study the effects of observable teacher characteristics on student achievement in elementary school. This paper contributes to this literature by utilizing an extensive teacher survey – in addition to administrative data – for a national sample of more than 10,000 elementary school students.

The analysis begins by investigating the underlying assumption that teachers are important predictors of student achievement. Using a simple student- and classroom-fixed effects specification, the set of classroom fixed-effects explain between 25 and 40 percent of the within-student variation in test score growth. In other words, classrooms matter. This finding suggests a large role for teachers and peers in affecting student achievement. This finding is echoed in recent work by Rivkin, Hanushek, and Kain [19] and Rockoff [20].

The second portion of the analysis attempts to determine the specific attributes of teachers that lead to higher (or lower) achievement. Despite the availability of detailed teacher characteristics unique to *Prospects* data, few teacher characteristics have significant impacts on student achievement as measured by classroom fixed effects. The results are particularly weak for the first-grade cohort. For the third grade cohort, there is some evidence that teacher experience is positively associated with classroom effectiveness in both mathematics and reading, consistent with the findings in Rivkin, Hanushek, and Kain [19]; Clotfelter, Ladd, and Vigdor [6]; and Rockoff [20].

In addition, teachers who report that their classroom materials are adequate are associated with sizable gains in mathematics achievement of 0.15 standard deviations for the third-grade cohort. Although this finding suggests that classroom materials are related to achievement, there is the possibility that the teacher responses reflect a more general attitude concerning adequacy in the classroom than actual adequacy itself. For example, the correlation coefficient between

the question on classroom materials and questions on the relationship with staff and parents are above 0.5. However, the insignificant or even negative impact of teacher enthusiasm – clearly an attitudinal question – suggests that teacher perceptions themselves may not be driving the classroom materials coefficient. Hopefully, future work with even richer data sets can separate the effects of classroom materials from teacher perceptions and attitudes.

Teacher education and certification, on the other hand, do not have significant impacts on student achievement for either cohort. Betts, Zau, and Rice [4]; Rivkin, Hanushek, and Kain [19]; and Rockoff [20] find generally insignificant effects of teacher education and certification.

The finding that smaller class size is not associated with higher student achievement does contradict the general finding of modest gains for smaller classes.[15] Perhaps the within-student variation in class size, roughly two students per class, is simply too small to identify a causal effect. Possibly there is no effect of class size on a low-stakes test in this national sample. There could be benefits to smaller class sizes, however, when teachers "teach to the test" in preparation for exams that are part of state accountability systems.

The findings from the *Prospects* data illustrate the importance of selected teacher characteristics in measuring teacher quality as defined by student achievement. The insignificance of most characteristics, especially those available in administrative data, shows that further work is needed in order to determine how teachers affect student achievement.

---

[15] See Krueger [17]; Angrist and Lavy [2]; Jepsen and Rivkin [16]; Rivkin, Hanushek, and Kain [19]; and Betts, Zau, and Rice [4]. However, Rockoff [20] finds insignificant effects of class size.

References

[1] D. Aaronson, L. Barrow, W. Sander, Teachers and student achievement in the Chicago public high schools, revision of Federal Reserve Bank of Chicago working paper WP2002-28, 2003.

[2] J.D. Angrist, V. Lavy, Using Maimonides' rule to estimate the effect of class size on scholastic achievement, Quarterly Journal of Economics 114 (2000) 533-576.

[3] D. Armor, et al, Analysis of the School Preferred Reading Program in Selected Los Angeles Minority Schools, RAND, Santa Monica, CA, 1976.

[4] J. Betts, A. Zau, L. Rice, Determinants of Student Achievement: New Evidences from San Diego, Public Policy Institute of California, San Francisco 2003.

[5] E. Bryant, *Prospects*: The Congressionally mandated study of educational growth and opportunity, technical report #1: Sampling procedures for the baseline and first follow-up surveys, Abt Associates, 1993, Unpublished manuscript.

[6] C.T. Clotfelter, H.F. Ladd, J.L. Vigdor, Teacher Sorting, Teacher Shopping, and the Assessment of Teacher Effectiveness, Duke University, 2003, unpublished manuscript.

[7] CTB Macmillan/McGraw-Hill, Comprehensive Test of Basic Skills, Fourth Edition: Technical Bulletin 1, CTB Macmillan/McGraw-Hill School Publishing Company, Monterey, CA 1989.

[8] R. Ehrenberg, D. Brewer, Do school and teacher characteristics matter? Evidence from High School and Beyond, Economics of Education Review 13 (1994) 1-17.

[9] R. Ehrenberg, D. Goldhaber, D. Brewer, Do teachers' race, gender, and ethnicity matter? Evidence from NELS88, Industrial and Labor Relations Review 48 (1995) 547-561.

[10] D.D. Goldhaber, D.J. Brewer, Why don't schools and teachers seem to matter? Journal of Human Resources 32 (1997) 505-523.

[11] D.D. Goldhaber, D.J. Brewer, Does teacher certification matter? High school teacher certification status and student achievement, Educational Evaluation and Policy Analysis 22 (2000) 129-145.

[12] E. Hanushek, Teacher characteristics and gains in student achievement: Estimation using micro data, American Economic Review 61 (1971) 280-288.

[13] E. Hanushek, The economics of schooling: Production and efficiency in public schools, Journal of Economic Literature 24 (1986) 1141-1177.

[14] E. Hanushek, J. Kain, J. Markman, S. Rivkin, Does peer ability affect student achievement? National Bureau of Economic Research Working paper 8502, 2001.

[15] C.M. Hoxby, The effects of class size on student achievement: New evidence from population variation, Quarterly Journal of Economics 115 (2000) 1239-1286.

[16] C. Jepsen, S. Rivkin (2002), What is the tradeoff between smaller classes and teacher quality? National Bureau of Economic Research Working paper 9205, 2002.

[17] A. Krueger, Experimental estimates of education production functions, Quarterly Journal of Economics 114 (1999) 497-532.

[18] R. Murnane, The Impact of School Resources on the Learning of Inner City Children, Ballinger, Cambridge, MA, 1975.

[19] S. Rivkin, E. Hanushek, J. Kain, Teachers, schools, and academic achievement, revision of National Bureau of Economic Research Working paper 6691, 2002.

[20] J.E. Rockoff, The impact of individual teachers on student achievement: Evidence from panel data, The American Economic Review 94 (2004) 247-252.

[21] B. Rowan, R. Correnti, R.J. Miller, What large-scale, survey research tells us about teacher effects on student achievement: Evidence from the *Prospects* study of elementary schools, Teachers College Record 104 (2002) 1525-1567.

Table 1: Descriptive Statistics for Teacher Characteristics

	First Grade Cohort				Third Grade Cohort			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Diff Teach for Math/Eng	0.086	0.280	0	1	0.378	0.485	0	1
Mathematics Teacher Characteristics								
Class Size	22.895	4.762	6	53	24.865	5.493	6	54
Total Experience	14.306	9.087	1	41	14.547	9.445	1	45
Bachelor's Degree Only	0.284	0.451	0	1	0.256	0.436	0	1
More than Bachelor's	0.339	0.473	0	1	0.302	0.459	0	1
Master's or More	0.376	0.484	0	1	0.442	0.497	0	1
Not Fully Certified	0.062	0.241	0	1	0.071	0.258	0	1
Presence of Aide	0.248	0.432	0	1	0.183	0.387	0	1
Use of Computers	3.292	1.262	1	5	3.151	1.260	1	5
Hours of Homework	1.116	1.249	0	12	1.977	1.350	0	14.5
Teacher Enthusiasm	4.828	1.170	1	6	4.723	1.221	1	6
Adequacy of Materials	0.808	0.394	0	1	0.776	0.417	0	1
English Teacher Characteristics								
Class Size	22.742	5.009	6	53	24.522	5.615	6	54
Total Experience	14.374	9.100	1	41	14.197	9.277	1	44
Bachelor's Degree Only	0.277	0.448	0	1	0.250	0.433	0	1
More than Bachelor's	0.341	0.474	0	1	0.313	0.464	0	1
Master's or More	0.382	0.486	0	1	0.437	0.496	0	1
Not Fully Certified	0.060	0.238	0	1	0.070	0.254	0	1
Presence of Aide	0.345	0.475	0	1	0.196	0.397	0	1
Use of Computers	3.292	1.258	1	5	3.138	1.248	1	5
Hours of Homework	1.371	1.273	0	12.5	1.834	1.415	0	12
Teacher Enthusiasm	4.837	1.151	1	6	4.674	1.267	1	6
Adequacy of Materials	0.805	0.396	0	1	0.769	0.422	0	1
Number of Students		18,837				18,639		

Table 2: Descriptive Statistics for Classroom Fixed Effects

Variable	Obs	Mean	Standard Deviation	R-squared	Percent Significant
First Grade Cohort					
Mathematics	1,099	-0.075	1.287	0.328	10.6%
English	1,255	-0.026	1.316	0.414	12.2%
Third Grade Cohort					
Mathematics	1,246	0.057	1.342	0.269	13.8%
English	1,397	-0.174	1.432	0.272	13.5%

NOTE: Percent significant measures the percent of fixed effects that are statistically significant from zero at a five-percent level (two-sided test).

Table 3: Effects of Teacher and Student Characteristics
Dependent Variable is Predicted Classroom Fixed Effect
First Grade Cohort

	Mathematics			Reading		
	(1)	(2)	(3)	(4)	(5)	(6)
Class Size	-0.008	-0.007	-0.004	-0.003	-0.004	0.003
	(1.26)	(1.12)	(0.67)	(0.58)	(0.67)	(0.57)
Experience (Years)	0.003	0.004	0.002	-0.0001	-0.001	-0.004
	(0.33)	(0.39)	(0.18)	(0.01)	(0.08)	(0.53)
Experience Squared	0.0001	0.00004	0.0001	0.0002	0.0002	0.0003
	(0.16)	(0.15)	(0.28)	(0.68)	(0.71)	(1.11)
BA or less	0.009	0.004	-0.004	0.003	-0.004	-0.002
	(0.13)	(0.06)	(0.06)	(0.04)	(0.06)	(0.03)
More than BA (no MA)	0.016	0.018	0.024	-0.006	-0.007	0.000
	(0.27)	(0.32)	(0.44)	(0.11)	(0.13)	(0.01)
Not Fully Certified	-0.036	-0.008	-0.031	0.129	0.128	0.057
	(0.36)	(0.08)	(0.32)	(1.39)	(1.36)	(0.65)
Aide in Class		-0.056	-0.096		0.015	-0.050
		(0.86)	(1.52)		(0.26)	(0.96)
Use of Computers		0.029	0.038		0.002	0.014
		(1.30)	(1.79)		(0.10)	(0.72)
Hours Homework per Week		-0.024	-0.005		0.011	0.024
		(1.04)	(0.21)		(0.60)	(1.43)
Teacher Enthusiasm		-0.050	-0.052		0.023	0.015
		(2.44)	(2.61)		(1.19)	(0.82)
Adequacy of Materials		0.071	0.045		0.011	-0.005
		(1.16)	(0.76)		(0.19)	(0.08)
Different Math/Eng Teacher		-0.023	-0.013		-0.104	-0.154
		(0.27)	(0.15)		(1.22)	(1.93)
Lagged Test Scores			-0.006			-0.008
			(8.38)			(12.93)
Observations	1,099	1,099	1,099	1,255	1,255	1,255

Notes: Absolute value of t-statistic in parentheses. All specifications also contain school and year fixed effects.

Table 4: Effects of Teacher and Student Characteristics
Dependent Variable is Predicted Classroom Fixed Effect
Third Grade Cohort

|	Mathematics			Reading		
	(1)	(2)	(3)	(4)	(5)	(6)
Class Size	-0.004	-0.003	0.005	-0.009	-0.010	0.004
	(0.62)	(0.49)	(0.79)	(1.53)	(1.70)	(0.69)
Experience (Years)	0.015	0.018	0.021	0.012	0.014	0.020
	(1.41)	(1.59)	(1.98)	(1.13)	(1.27)	(1.92)
Experience Squared	-0.0003	-0.0004	-0.0004	-0.0005	-0.001	-0.001
	(0.98)	(1.08)	(1.31)	(1.51)	(1.66)	(2.14)
BA or less	0.001	0.008	0.007	-0.001	-0.026	-0.003
	(0.02)	(0.11)	(0.10)	(0.01)	(0.36)	(0.04)
More than BA (no MA)	-0.044	-0.061	-0.047	0.023	0.001	0.009
	(0.66)	(0.92)	(0.73)	(0.36)	(0.02)	(0.15)
Not Fully Certified	-0.161	-0.151	-0.144	-0.172	-0.136	-0.112
	(1.58)	(1.49)	(1.46)	(1.58)	(1.25)	(1.07)
Aide in Class		0.063	0.032		-0.003	-0.060
		(0.76)	(0.40)		(0.04)	(0.75)
Use of Computers		-0.034	-0.035		-0.066	-0.072
		(1.28)	(1.37)		(2.59)	(2.95)
Hours Homework per Week		0.018	0.040		0.012	0.025
		(0.88)	(2.03)		(0.60)	(1.31)
Teacher Enthusiasm		-0.047	-0.038		0.026	0.027
		(1.97)	(1.64)		(1.17)	(1.23)
Adequacy of Materials		0.147	0.152		0.098	0.107
		(2.09)	(2.25)		(1.44)	(1.64)
Different Math/Eng Teacher		-0.001	-0.016		0.146	0.100
		(0.01)	(0.18)		(1.68)	(1.20)
Lagged Test Scores			-0.007			-0.009
			(8.44)			(9.88)
Observations	1,246	1,246	1,246	1,397	1,397	1,397

Notes: Absolute value of t-statistic in parentheses. All specifications also contain school and year fixed effects.

Figure 1: Distribution of T-statistics for Classroom Fixed Effects in Mathematics
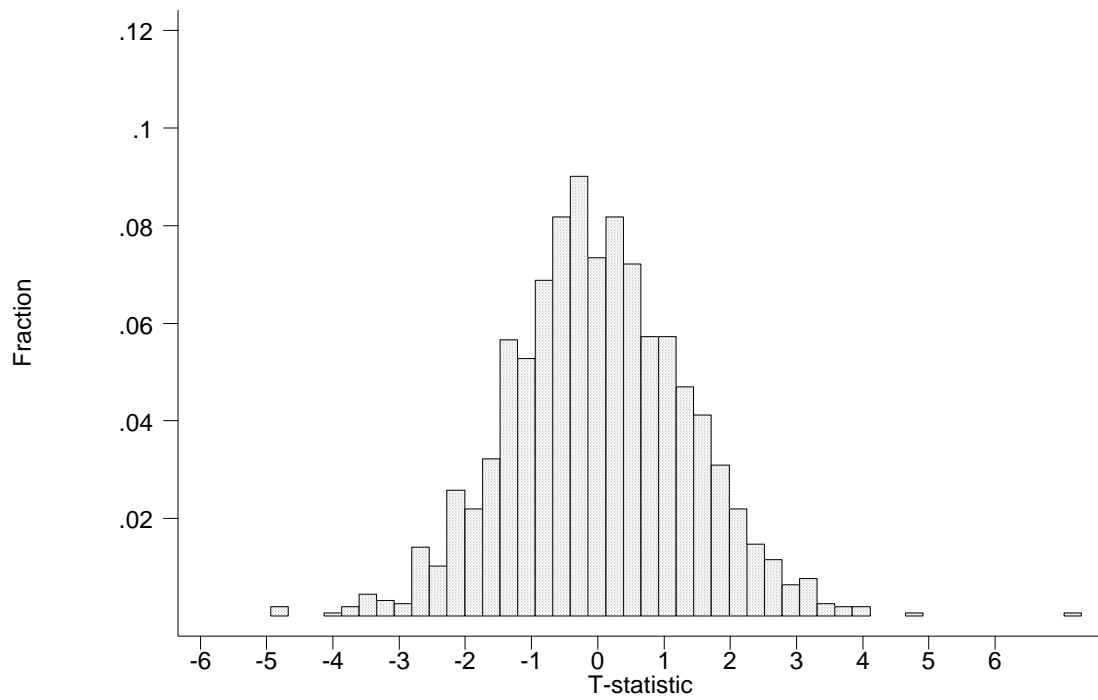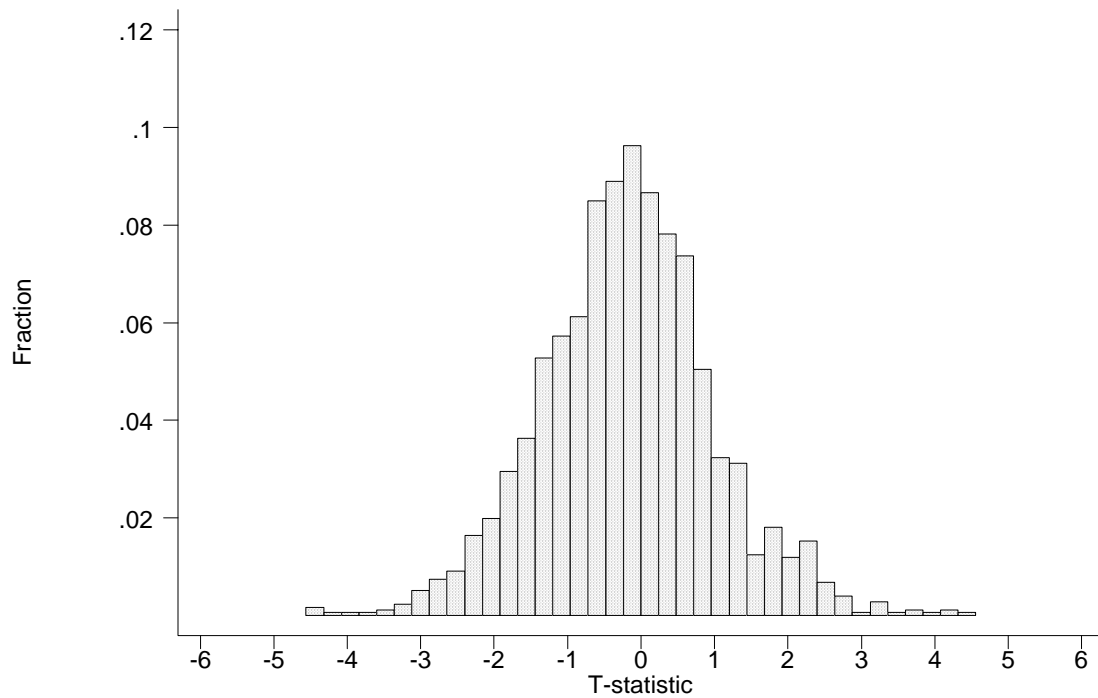
Grade 1 Cohort
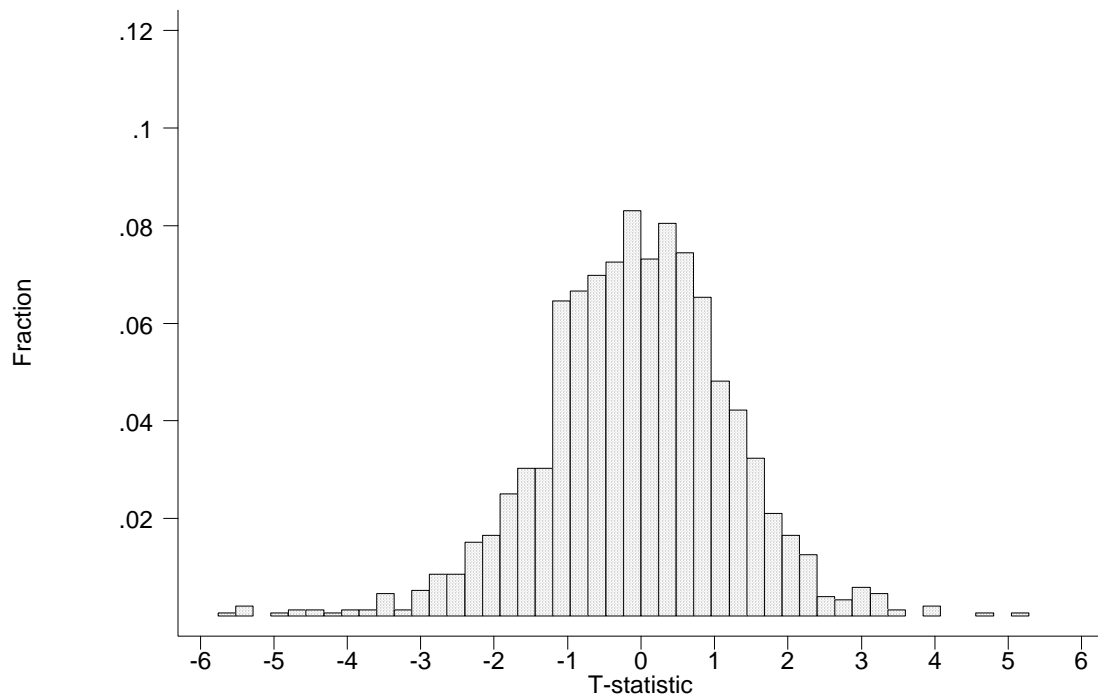
Grade 3 Cohort
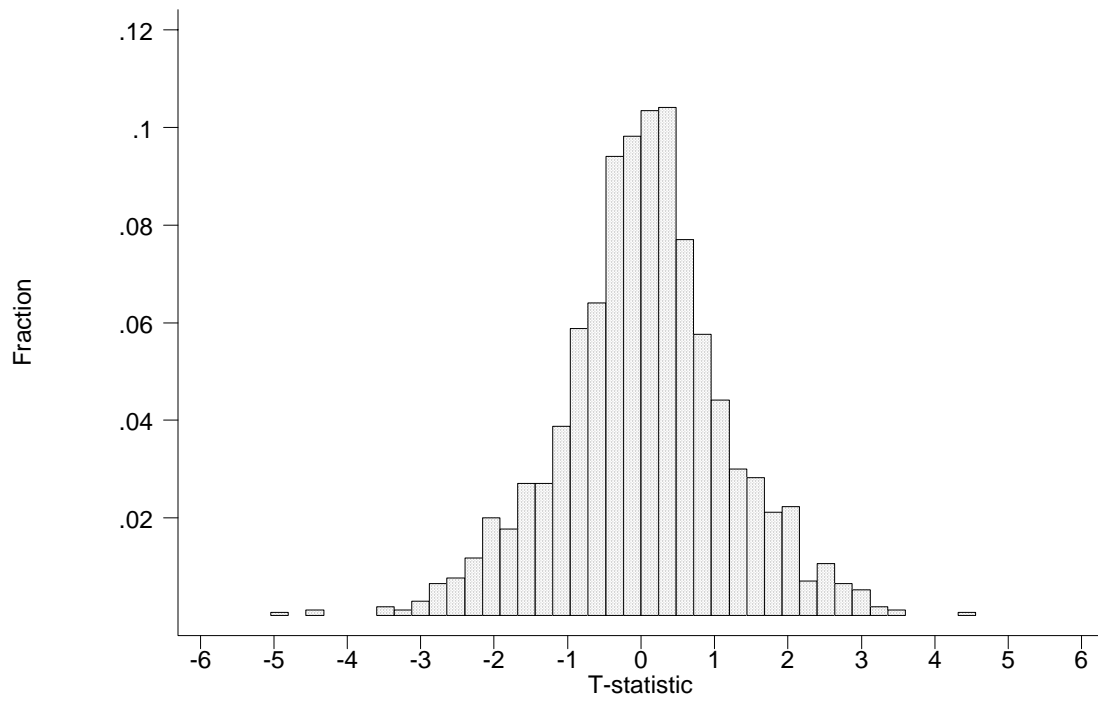
Appendix Figure 1: Distribution of T-statistics for Classroom Fixed Effects in Reading

Grade 1 Cohort

Grade 3 Cohort

Appendix Table 1: Within-School Assignment of Teachers to Classrooms
Effect of Student Characteristics for First Grade Cohort

Teacher Characteristics

	Average Experience	BA Only	BA or More	Not Fully Certified	Presence of Aide	Hours of Homework	Enthusiasm	Adequate Materials
Results for Mathematics Teachers								
Lagged Math Score	-1.290	0.0002	0.026	-0.012	-0.057	0.289	-0.028	-0.058
	(1.72)	(0.01)	(0.67)	(0.57)	(1.88)	(3.30)	(0.29)	(1.82)
Percent White	1.969	-0.029	-0.072	0.086	0.052	-0.263	0.351	0.019
	(0.88)	(0.28)	(0.63)	(1.40)	(0.58)	(1.00)	(1.22)	(0.20)
Percent Free Lunch	-2.690	0.072	-0.008	0.011	0.079	0.258	0.230	-0.008
	(1.47)	(0.83)	(0.09)	(0.22)	(1.07)	(1.21)	(0.98)	(0.10)
Obseravtions	1,099	1,099	1,099	1,099	1,099	1,099	1,099	1,099
Results for English Teachers								
Lagged Reading Score	-0.422	-0.0003	0.032	-0.033	-0.072	0.122	-0.078	-0.024
	(0.70)	(0.01)	(1.01)	(1.90)	(2.62)	(1.46)	(0.99)	(0.92)
Percent White	0.616	0.016	-0.156	0.086	0.025	0.098	0.367	-0.034
	(0.31)	(0.17)	(1.52)	(1.53)	(0.28)	(0.36)	(1.43)	(0.40)
Percent Free Lunch	-0.958	0.022	0.029	-0.008	0.071	0.255	-0.047	0.029
	(0.57)	(0.27)	(0.33)	(0.16)	(0.94)	(1.11)	(0.22)	(0.40)
Obseravtions	1,255	1,255	1,255	1,255	1,255	1,255	1,255	1,255

NOTES: Absolute values of t-statistics are in parentheses. Each column in each panel represents a separate regression. All regressions also contain controls for percentage female, a dummy variable for missing percent free lunch, and year and school fixed effects.

Appendix Table 2: Within-School Assignment of Teachers to Classrooms
Effect of Student Characteristics for Third Grade Cohort

Teacher Characteristics

	Average Experience	BA Only	BA or More	Not Fully Certified	Presence of Aide	Hours of Homework	Enthusiasm	Adequate Materials
Results for Mathematics Teachers								
Lagged Math Score	0.915	-0.026	0.019	-0.010	-0.020	0.265	0.115	-0.006
	(1.98)	(1.18)	(0.79)	(0.68)	(1.14)	(3.75)	(1.86)	(0.31)
Percent White	-1.768	0.004	-0.051	0.082	-0.029	0.005	-0.117	0.074
	(1.04)	(0.05)	(0.59)	(1.44)	(0.46)	(0.02)	(0.51)	(0.97)
Percent Free Lunch	0.612	-0.026	-0.030	0.028	0.057	-0.086	0.121	-0.030
	(0.41)	(0.37)	(0.39)	(0.55)	(1.02)	(0.37)	(0.60)	(0.45)
Obseravtions	1,246	1,246	1,246	1,246	1,246	1,246	1,246	1,246
Results for English Teachers								
Lagged Reading Score	0.805	-0.006	0.014	-0.004	-0.035	0.147	0.018	0.003
	(1.92)	(0.32)	(0.65)	(0.28)	(2.25)	(2.34)	(0.32)	(0.17)
Percent White	1.445	0.097	-0.025	-0.010	-0.039	0.103	0.334	0.020
	(0.88)	(1.24)	(0.29)	(0.21)	(0.64)	(0.42)	(1.50)	(0.27)
Percent Free Lunch	0.912	-0.032	0.056	0.009	0.079	-0.104	0.195	0.056
	(0.64)	(0.47)	(0.75)	(0.21)	(1.50)	(0.49)	(1.02)	(0.86)
Obseravtions	1,397	1,397	1,397	1,397	1,397	1,397	1,397	1,397

NOTES: Absolute values of t-statistics are in parentheses. Each column in each panel represents a separate regression. All regressions also contain controls for percentage female, a dummy variable for missing percent free lunch, and year and school fixed effects.