A working paper series

The Cecil and Ida Green Center for the Study of Science and Society

The University of Texas at Dallas

The Impact of Individual Teachers and Peers on Individual Student Achievement

John F. Kain

November 17, 1998 (Revised)

Draft Report / Do not cite without author's permission

Paper prepared for the Association for Public Policy Analysis and Management (APPAM) 20th Annual Research Conference

Grand Hyatt, New York

October 31, 1998

UTD Texas Schools Project

ACKNOWLEDGEMENTS

I am grateful to Ankur Gupta, Elim (John) Horton and Vivek Dhar for programming and reasearch assistance and especially to Jeong-Dai Kim, who devoted many evenings and weekends to completing the estimation for this paper. As usual, Dan O'Brien, made numerous critical and helpful comments on both the analysis and exposition. Finally, the descriptions of the Tennessee studies included in this paper benefited greatly from discussions with Kurt J. Beron about the methods used. The Spencer and Smith Richardson Foundations' support is gratefully acknowledged.

***

The Impact of Individual Teachers and Peers On Individual Student Achievement

Introduction

> A number of studies provide direct analyses of differential effectiveness of teachers by estimating differences in the average performance of each teacher's students (after allowing for differences in family backgrounds and initial achievement scores). The findings . . . are unequivocal: Teachers and schools differ dramatically in their effectiveness. The formal statistical tests employed in these studies confirm that there are striking differences in average gain in student achievement across teachers.
>
> Eric A. Hanushek. 1989. "The Impact of Differential Expenditures on School Performance," Educational Researcher (May), p. 48.

This paper is concerned with the impact of classroom teachers and peers on individual student achievement. It begins with a survey of previous research on teacher effects including a reanalysis, by the author, of data from a paper by Jordon, Mendro and Weerasinghe. The second part of the paper contains exploratory analyses of the effects of individual teachers and peers using data for more than 200,000 third and fourth grade students attending Texas public schools.

As the preceding quote by Hanushek demonstrates, the idea that individual teachers have a large impact on individual student achievement is hardly novel or new. [1] Close examination of the sources cited by Hanushek (1989), however, reveals that they consist of three small studies: (1) Hanusek's (1971) study of 1,000 third graders in one California school district, (2) Murnane's (1975) study of 900 second and third grade

[1] Hanushek (1989, p. 48) adds that while "an important element of skill is involved in successful teaching … it is currently impossible to measure with any precision any readily identifiable components or elements

Chapter I students attending 16 inner-city schools in New Haven, and (3) Murnane and Phillips' (1981) study of 800 third, fourth, fifth and sixth grade students enrolled in a federally funded welfare reform experiment in a large Midwestern city. The numbers of teachers and classrooms included in these studies are not provided, but they were less than150 for all three studies combined.

Quantitative evidence on the effects of individual teachers on individual student achievement has been growing. Using panel data developed by the Harvard/UTD Texas Schools Project, Rivkin, Hanushek and Kain (1998), for example, estimate that the contributions of individual teachers account for at least 7.5 percent of the total variation in student achievement and that the true percentage is likely to be significantly larger. This analysis uses 3rd, 4th, 5th and 6th grade reading and math scores for more than 750,000 Texas students in three cohorts who attended more than 3,000 Texas public schools. While the Texas Schools Microdata Panel (TSMP) data used for these analyses do not permit the identification of the particular classroom teacher(s) for each student, they do identify the campus, grade and program (regular, bilingual, ESL and special education) for both. By comparing student performance across grades for the same cohort and across cohorts/years for the same grade, Rivkin, Hanushek and Kain were able to make inferences about the contributions of individual classroom teachers to individual student achievement. While I find this evidence quite persuasive, it is nonetheless indirect. More direct evidence is available from studies by Tennessee and Dallas researchers. The following two sections examine these studies.

The Tennessee Value-Added Assessment System (TVAAS)

TVAAS is an ambitious effort to use a longitudinally merged database of test data for individual students to prepare quantitative estimates of the effectiveness of school districts, schools and teachers for the entire state of Tennessee (Sanders and Horn. 1994;Wright, Horn and L. Sanders, 1997). TVAAS, mandated by the Tennessee Educational Improvement Act which took effect on July 1, 1992, employs a variant of

---

of this skill." In addition, he observes that "it is unclear whether any form of teacher training could be organized to foster high levels of skill in teachers."

Henderson's mixed-model equations to estimate district, campus and teacher contributions to individual student achievement (Sanders, Saxton and Horn, 1998). The TVAAS database, which includes test data for Tennessee's entire grade 2-8 population from 1990 to the present, contains standardized test results for individual students in mathematics, reading, language arts, science and social studies. The test data used for TVAAS are the scaled scores for the norm-referenced portion of the Tennessee Comprehensive Assessment Program (TCAP). The scores for the five subjects listed above are included with each student's record (up to five years) as well as the student's campus and his/her teacher(s) in each grade/subject.

The most problematic feature of the analyses by Sanders, Saxton and Horn (1998) is their failure to include any student variables beyond prior test scores to explain achievement gains and to assess campus and teacher effectiveness. They argue that "by taking advantage of the longitudinal aspect of the data, each student serves as his or her own control." They add that "each child can be thought of as a 'blocking factor' that enables the estimation of school system, school, and teacher effects free of the socio-economic confoundings which have historically rendered unfair any attempt to compare districts and schools based upon the inappropriate comparison of group means." In spite of the fact that the authors claim that their research demonstrates the adequacy of this assumption, it is strongly contradicted by our value-added models of student achievement using TSMP data (Kain and O'Brien, 1998a and 1998b) and by results presented in this paper.

The mixed-model methodology used in TVAAS employs "shrinkage estimates" to provide conservative estimates of teacher effects. With this methodology all teachers are assumed to be at their system's mean, except when the gains of their students exceed or are less than their system's average by a significant amount and are based on significant numbers of students. While the TVAAS methodology employs up to five years of tests for individual students, it allows all students to be included in the teacher, school and district calculations, including students with a single year's scores.

A recent paper by Sanders and Rivers (1996) uses data for a cohort of students attending two Tennessee school districts during the four-year period 1992-1995 in an effort to quantify the cumulative effects of better or worse teachers on individual student achievement over a three-year period. They used Equation 1 to obtain teacher effects for each grade level and then assigned them to five teacher quality quintiles. The first included the least and the fifth the most effective teachers. Indexes of Teacher effects were not obtained for second grade teachers, but the second grade student scores were used in a second level analysis as a control for omitted student characteristics. The author's claim that a variety of analyses demonstrate that the prior test scores fully control for the effects of race, income and other omitted variables. I am not convinced.

(1) Current score = a + b(previous math score) + t(i) + error

Where a = constant to be estimated from the data

b = regression coefficient

t(i) = shrinkage estimate of the teacher effects

In the second stage the authors predict mean scores for students enrolled in each of 125 teacher quality by grade/year sequences defined by their assignment to five teacher quality (quintiles) over a three year period. The TVAAS procedure estimates teacher effects for each district separately. This has the advantage of reducing the variation in student characteristics that must be accounted for by prior test scores, but it also means that the teacher quality indexes cannot be compared across districts.

Table 1 presents Sanders and Rivers' (1996, p. 12) estimates of mean cumulative impacts for seven of the 125 teacher quality/grade sequences. The fifth grade scores in Table 1, which are percentiles and are adjusted for each district's and category's mean second grade score, are meant to demonstrate the cumulative impact of a particular teacher quality/grade sequence on student achievement. In comparing the high-high-high and low-low-low sequences, Sanders and Rivers' (1996, p. 3) state that "with an even start, the difference in these two extreme sequences resulted in a range of mean student percentiles in grade five of 52 to 54 points."

Dallas Estimates of Teacher Effects

Using data for five cohorts of students for the period 1993-1996, Jordon, Mendro and Weerasinghe (JMW) have replicated Sanders and Rivers' (1996) analysis of the cumulative impact of teachers on student achievement with equally stunning results. Using a somewhat different methodology, JMW (1997) assign individual teachers to five quality levels based on the mean difference in the actual and predicted performance of their students in both reading and math during a three year period 1994-1996. The student achievement equations that form the basis of their teacher effectiveness measures are limited to students with four years of complete testing data for the period 1993 through 1996. It also appears that the same teacher may be assigned to more than one teacher quality category (in different years). If so, this convention would increase the possibility that the teacher effects estimates obtained by JMW may include peer effects as teachers have a different group of students in each year (JMW, 1997, p. 2).

As in Sanders and Rivers (1996), the JMW analysis focused on the cumulative extent to which changes in individual student achievement were affected by their assignment to teachers of varying quality over a three year period. Specifically, they asked how the achievement of otherwise identical students would be affected over a three year period if they had three years of top rated teachers, three years of the lowest rated teachers or some intermediate combination. Following Sanders and Rivers (1996), JMW defined 125 categories of students for each of five cohorts and for reading and math. To examine the cumulative effects of these differential three-year assignments to teachers of different quality, the authors calculated mean NCE (Normal Curve Equivalent) scores for the base year (1993) and for each of the three years, 1994-1996 for each of the 125 teacher quality/year combinations.

Table 2 provides raw NCE mean reading scores in 1993, 1994, 1995 and 1996 for six groups of students and the number of students in each category as well as the absolute and percentage differences between 1993 (the base or control year) and 1996. The six groups in Table 2, which were used by JMW to illustrate their results, were again defined by teacher quality in each of three consecutive years. The first row gives estimates for students who were taught by the lowest quality teachers in all three years, while the first row in the lower panel (row 4) gives the same figures for students who were taught by the highest quality teacher in both years. JMW (1997: 6) selected the remaining four rows (rows 2 and 3 in the top panel and rows 5 and 6 in the lower panel) to provide "two groups in each cohort with 1993 means approximately equal to the 1993 111 group mean,"and "two groups in each cohort with means equal to the 1993 555 group mean."

Based on their review of 10 tables containing mean scores for 125 teacher quality combinations, JMW find that "the broad results Sanders and Rivers reported are confirmed by our data." They add, however, that "the effects have a strong additive component, they have a strong cumulative component, and they show little evidence of cumulative student outcomes" (JMW, 1997, p. 7). They also acknowledged that a search of the data in the ten appendix tables will uncover "some counterexamples to their illustrative groupings," but they nonetheless argue that they are generally consistent with the patterns shown in the example cases.

While I found it difficult to assimilate the results for the 1,250 teacher cohort/test/quality/year combinations provided by JMW, my inspection of them left me reasonably convinced by JMW's claims. Even so, I was uncomfortable with having to rely on such a casual analysis. Thus, I used the data from JMW's 10 appendix tables to estimate the equations reported in Tables 3 and 4.

Table 3 contains linear OLS estimates of the effect of base year (1993) test scores and indexes of teacher quality for 1994, 1995 and 1996 on 1996 reading and math scores for each of the five cohorts studied by JMW. The index of teacher quality was simply the values 1, 2, 3, 4 or 5 for each of the teacher quality categories. Starting with the reading regressions, the R^2 for the five equations varies from 0.67 to 0.88 (four are larger than 0.82) and the coefficient for mean 1993 reading scores varies in the narrow range 0.48 to 0.58. The relative magnitudes of the yearly teacher quality measures are also highly consistent across equations. For every cohort, the most recent teacher has the largest impact on 1996 reading scores, while the first teacher has the second largest impact for all but the fifth grade cohort

The overall explanatory power (R^2) of the five math equations are larger than for the comparable reading equations. The individual parameter estimates, moreover, are similar to those for reading, if somewhat less regular. The impact of 1993 test scores on 1996 math scores is more varied, from 0.39 to 0.76, and is greater for the older cohorts. Both these results and the previously discussed ones for reading illustrate the problem of relying on casual empiricism. From my earlier review of the mean 1993 test scores for the 125 teacher quality/year categories for each cohort and test, I had formed the distinct, though incorrect, impression that they were very similar and bore no consistent relationship to 1996 levels. As the coefficients for both Read 93 and Math 93 in Table 3 clearly demonstrate, however, that this view was incorrect as mean 1993 scores have a large impact on mean 1996 scores in all 10 equations.

Returning to the teacher quality variables, the most recent teacher once again has the largest effect on 1996 math scores, holding constant the effect of 1993 math scores and the effects of the two other teachers. In contrast to the reading results in which the first teacher had a larger impact on 1996 achievement than the second for all but one cohort, this result does not hold for the fourth, fifth and eighth grade math equations.

The log-log results for reading, which are shown in the top panel of Table 4, are qualitatively very much like the linear ones. All five equations explain more than 70 percent of the variance in the natural logarithm of mean 1996 reading scores for the 125 groups of students. The coefficient of the natural logarithm of mean 1993 reading scores varies between 0.44 for the fifth grade to 0.65 for the fourth. The most recent teacher once again has the largest impact on 1996 reading scores and there is again a tendency for the first teacher to have the second most influence. The math results are similarly broadly consistent with both the results for the log-log reading equation and the linear math equations. Taken in their entirety, the results generally support the findings of Sanders and Rivers (1996) and JMW (1997). They also provide further evidence of the importance of doing well in the early years and by implication the difficulty students face in gaining ground. In addition, they provide evidence that a high quality recent teacher can have considerable success in helping their students overcome the effects of previous bad teaching or teachers. At the same time the results indicate that over a three year period both the second and third teachers cumulatively affect student achievement. Finally, it should be emphasized that the results of this analysis assume that JMW have produced valid measures of teacher quality and that the values 1, 2, 3, 4 and 5 and their natural logarithms are meaningful indexes of this quality measure.

Using Texas Schools Microdata Panel Data (TSMP) to Quantify Teacher Effectiveness

As noted previously, the TSMP data used by Rivkin, Hanushek and Kain (1998) do not currently identify the classroom teachers of individual students. Student-teacher links are absent from TSMP because they are not collected by TEA for inclusion in its Public Education Information Management System (PEIMS). It turns out, however, that the NAPT (Norm-Referenced Assessment Program for Texas) tests included in TSMP have information that may be used to estimate these links for seven grade/year combinations, albeit with some error. As Table 5, which shows the five cohorts currently included in TSMP as well as the years and grades both NAPT and TAAS (Texas Assessment of Academic Skills) were given, reveals NAPT was given for only two years, 1992 and 1993. While we have been working with TEA since 1996 to create individual student-teacher links for the NAPT tests, we have thus far obtained them for only one, the NAPT administered to fourth grade students in 1992, i.e. NAPT492. Close inspection of Table 5 demonstrates that while we will be able to infer individual student teacher links from all seven NAPT tests, a prior test score will be available for only four of them.

While the exploratory analyses described in this paper, are quite informative, NAPT492 is less than ideal for studying teacher effects. The principal problem is that it was given in spring 1992 while TAAS391, which we use as the prior test in estimating value-added regressions, was given in fall 1991. This means that instead of quantifying the contributions of a single classroom teacher, the value-added equations based on NAPT492 and TAAS391 reflect the combined contributions of at least two separate classroom teachers. Before we consider the regression results, however, we provide a brief discussion of the student-teacher matches and various problems associated with them.

How Good Are the Student-Teacher Matches?

When NAPT was given, the persons who administered it were asked to write their name on a header sheet. TEA's contractor then scanned these header sheets and added these unformatted Test Administrator (TA) names to the individual test records. The NAPT tapes that TEA originally supplied to us did not include these names and we only learned of their existence in 1996. Since then we have been working with TEA to develop a mutually agreeable procedure for having TEA match the NAPT TA names to teacher names included in their PEIMS database. While this has been a lengthy and tedious process, we currently have teacher-student name matches for NAPT492 and TEA is working on the remaining six.

As Table 6 indicates there were 267,093 NAPT492 test booklets and 13,826 separate test administrations for NAPT492, although the number of useable names was somewhat smaller. The first round of the matching procedure produced 11,664 valid name matches. When 438 more cases, where a match could be inferred from a single unmatched TA name and a single unmatched fourth grade teacher at a single campus, were added the result is that 12,529 test administration groups were matched with a particular teacher. It should be emphasized that useful information may be obtained from this procedure even when a match to a teacher in the PEIMS database is not achieved. When unmatched TA groups are a single classroom, they provide highly useful information about within campus differences in classroom composition. In the analyses that follow, we calculate mean lagged reading and math scores by classroom and use them as explanatory variables in our exploratory regressions.

Even a brief glimpse at Table 6 makes it clear that not all of the test administration groups are individual classrooms. Of the 13,826 TA groups, 343 had more than 30 students and their average size exceeded 87. These groups are simply too large to be single classrooms and must be instances where a single individual administered the test to two or more classrooms. The number of students in these large TA groups totaled nearly 30,000 and more than 80 percent of these groups did not have a teacher match. Finally, 1,254 TA groups had fewer than 10 students, although because of their small size, 4.5 students per group, the number of students totaled just over 5,400.

As I will discuss in greater detail at a later point, significant numbers of Limited English Proficient (LEP) and special education students were exempted from NAPT or did not have valid scores. As Table 7 reveals, these students were disproportionately concentrated in very small and large TA groups. When students without valid scores are excluded, the number of test administrator groups decreases to 12,362 (11,336 for students with teacher matches) and the number of students declines to 249,696 (214,278 for students with teacher matches).

4 provide the same equations for students in all TA groups. The results are nearly identical to those obtained for students in TA groups with 10-30 students.

For each of the four categories defined by classroom size and type of student, we have estimated three types of equations, an OLS equation, a campus fixed effects equation and a classroom/TA group fixed effects equation. All three of these equation types include the same explanatory variables with three exceptions. The OLS and campus fixed equations include a peer variable and two ratios that measure the rates at which students in a particular classroom in 1992 are not tested. They are omitted from the classroom fixed effects equation.

The first ratio, Prop No Score, is the proportion of students in a particular classroom that take NAPT, but do not receive a score. The second, Prop No Attempt, is similarly the proportion in a particular classroom that did not take the test at all. Most of those whose tests were not scored or who did not take the test were either LEP or special education students. The classroom ratio variables are excluded from the classroom fixed effects because the use of classroom fixed effects make them redundant, any influence these variables have on individual student achievement are fully captured by the classroom fixed effects.

The peer variable (Peer Read 91 /Peer Math 91) is the third variable that is included in the OLS and campus fixed effects equations and excluded from the classroom fixed effects equation. These peer variables are the mean of prior (1991) reading or math scores for the "other" students in a particular classroom. In an effort to minimize the bias that might arise from calculating the peer scores for only those students who had both 1992 and 1991 scores, we estimated 1991 scores for all students who took NAPT in 1992 and did not have a valid 1991 score. Both the actual and predicted values of the Peer Read 91 and Peer Math 91 are used in calculating the classroom means. The equations used for this purpose are given in Appendix Table A-5.

The peer variable is not perfectly correlated with the classroom dummy because each individual's score is subtracted from the classroom mean in calculating the peer mean for that individual. It is, nonetheless, omitted from the classroom fixed effects equations because the correlation between it and the classroom dummies are too high to yield meaningful estimates of peer effects. The fixed effects thus include the effects of mean 1991 reading/math scores for peers as well as the effects of both 1991 and 1992 teachers and other factors that may have affected the achievement of all students in a particular classroom. If we can believe the results derived from JMW data, the impact of the 1992 teacher is probably greater than the impact of the 1991 teacher.

In addition to the peer variables, all regressions include both the number of correct reading/math answers and the square of these variables, 14 dummy variables for the household race/ethnicity and household income of each student, student age and gender and two mobility variables. The equations for all students also have dummy variables indicating whether a particular student is LEP or enrolled in ESL, bilingual or special education programs.

Results for Students in Classes with 10-30 Students

The six equations in Table 9 and 10 explain between 48 and 67 percent of the variance in individual student achievement. In addition, the R^2s for the "all students" equations are higher than for the "regular" students equations and the R^2s for the math equations are uniformly higher than for the comparable reading equations. Finally, and not surprisingly, the overall explanatory power of the campus and fixed effects equations are greater than the simple OLS equations. In both the all and regular students equations using classroom in place of campus fixed effects increases the overall explanatory power by 5-6 percentage points, even though the campus fixed effects equations already include classroom peer reading/math variables as well as the classroom No Attempt and No Score dummies. In contrast, campus fixed effects added only 2-3 percentage points to the R^2 obtained in the OLS equations.

Peer Read 91, the first variable in Table 9, indicates how much a student's peers affect his individual achievement holding his individual prior achievement score and a large number of other explanatory variables constant. The peer variables are highly significant in all equations in which they appear; the smallest t statistic for the eight peer variables in Tables 9 and 10 is 30. The obvious next question is how much impact do these peer variables have on individual achievement. Using standard deviations as the measure, the impact on individual achievement of being in a classroom that has an average peer score one standard deviation above the mean as compared to one standard deviation below varies from 1.7 points (All Students, OLS) to 2.1 points (Regular, Campus FE). As the mean 1992 reading scores are 27.4 for all students and 28.6 for regular students, this difference would result in an increase in the individual reading score of a student in the classroom with a higher mean score by between 6.2 and 7.3 percent of the average 1992 reading score.

Since some students took a Spanish language version of TAAS in the third grade all of the equations include both a lagged Spanish language and lagged English language score as explanatory variables. The value of the English language variable is zero for students taking the Spanish language test and similarly the value of the Spanish language test is zero for those taking the English language test. Finally, we test for nonlinearities by including a squared term for both tests. Figure 3 graphs these functions for the Spanish and English language tests using coefficients from the second equation in Table 9, the campus fixed effects equation for all students. The graphs are constructed so that they pass through both the third and fourth grade mean scores of each group, i.e. those taking the English language and those taking the Spanish language TAAS in the third grade. While the graph for the English test increases more rapidly than the one for the Spanish language test, the more important result is that prior achievement in reading has a large effect on current reading achievement whether the prior test is in Spanish or English. The graph for those taking the English language test suggest a catch up or reversion to the mean. Students with very low third grade scores have more rapid gains and tend to catch up to students with higher third grade reading achievement. These effects are much less pronounced for those who took the Spanish language test.

The next 14 variables are interactions that quantify the joint effects of race/ethnicity and household income/poverty on individual reading/math scores. The household income variables are based on the eligibility of individual students for free or reduced price lunches under the federal school lunch program. Eligibility is based on federal definitions of the poverty level and thus depends on both family income and family size. To receive a free lunch, a child must be a member of a family whose income is less than 135 percent of the poverty level for its size. Similarly, to receive a reduced-price lunch, family income must be between 135 and 185 percent of the poverty level. Students whose families receive AFDC benefits or who participate in a number of other poverty programs are also eligible for a free lunch. In subsequent discussions, we use the terminology high-, low- and very low-income households in referring to these categories.

The income/race/ethnicity dummy variables are obtained by multiplying the three household, zero-one, income variables times the zero-one race/ethnicity dummy variables. Even though we use three income categories in calculating the income/race interaction variables, the equations include only two income/race variables for Anglos. This is because one of the variables must be omitted in order to estimate the equations. In this and similar cases, the omitted dummy becomes the base case and the coefficients of the remaining categorical variables are interpreted relative to it. The base case for these variables is high-income Anglos.

The regularity of the coefficient estimates for the race/ethnicity by income variables is remarkable. Holding income and the other variables included in the equations constant, Asians have the highest reading scores followed by Anglos. Hispanics and Native Americans are next with similar coefficient values and African Americans have the lowest scores by a significant amount. With the exception of the estimates for Asians, moreover, the pattern of coefficients by income in all six reading equations are precisely as expected; students from high-income families have the highest scores and those from very low-income households have the lowest. Inspection of the t statistics in Appendix Table A-2 reveals that none of the t statistics for low- and very low income Asians exceed one, indicating that these differences should be ignored. The lack of precision of these estimates no doubt reflects the relatively small number of Asian students. Interestingly enough, the expected pattern is obtained for Native Americans even though there are fewer of them.

The male coefficient is highly consistent across the six reading equations ranging from -0.29 in the classroom FE equation for regular students to -0.37 for the OLS equation for all students. The effect is small but highly significant statistically, the t statistics exceed 10 for all six equations. The coefficient for age is also negative indicating reading scores decrease by 0.09 to 0.11 for each additional year of age. These estimates are also highly significant statistically.

The notion that mobility has an adverse effect on student achievement is widely held. The two mobility variables provide strong support for this view and further suggest that it makes little or no difference whether students move from one school to another within the same district or move to a different school in a different district. The estimates indicate that moving from one campus to another, whether between or within the same districts reduces reading scores by between 1.4 and 2.1 percent of the mean fourth grade reading score.[2] Again these estimates are highly significant statistically, the t statistics exceed seven for all 12 mobility coefficients.

The control variables for the classroom percentage of students whose tests were not scored or who did not take the test yielded inconsistent results. With one exception the No Score variable is negative and in all four reading equations where it appears and the No test variable is positive, indicating that reading scores tend to be higher for students in classrooms with a higher fraction of excused students. On reflection these variables are not well specified. The tests include codes that provide more information on the reasons for non testing, particularly whether students were excused because they were LEP or special education students. This is an area where some improvement may be possible.

The last four variables appear only in the all students equations identify LEP students and students enrolled in ESL, bilingual and special education programs. Since the equations for regular students mostly omit these groups they are excluded from these equations. Being LEP reduces a student's fourth grade reading score by between 3.0 and 3.5 percent and students participating in a bilingual program experience an additional 0.9 to 1.5 percent reduction. These results further show that students in special education

[2] As Kain and O'Brien (1998) note, there are actually two types of within district moves, which they term voluntary moves and transfers. Voluntary moves are when an individual student moves from one campus to another because of a residential move or because of differences between them or their parents school administrators. Transfers, in contrast, are when all or most of the students in a particular grade move to another campus. Graduation from elementary or middle school are the most common causes of transfers, although it should be understood that there is nearly every combination of grades imaginable among Texas public schools. Kain and O'Brien (1998) in value added regressions for grades 4-7 found, somewhat surprisingly, that transfers had a larger negative effect on individual achievement than voluntary moves.

programs, every thing else held constant, have fourth grade reading scores that are 5.8 to 6.0 percent lower than otherwise identical students. There is a large difference in the mean reading scores of special education and LEP students, a fact that explains the higher R^2s obtained for the All Student equations. It is perhaps worth noting that separate equations for regular, LEP and special education students do a better job of explaining these differences than a pooled equation even though the pooled equation has a higher R^2 than any of the individual ones. It should also be emphasized that the results for LEP and special education students include only those students who took the fourth grade NAPT and that significant numbers of students in these groups were not required to take the exam.

The reading and math coefficients cannot be directly compared as there are 67 questions on the math section of the fourth grade NAPT in 1992 as opposed to 44 reading questions. Nonetheless, as Table 10 indicates, the math results are very similar to those for reading. As I already noted, the R^2s for the math equations are uniformly higher than for the comparable reading equations. As in the reading equations, the coefficients of the peer variable are highly significant statistically. In addition, the difference in this variable in the OLS and campus fixed effects equations is greater than in the reading equation. As a result, the effects of peers on math achievement with campus FE are considerably larger than for reading.

The effects of Spanish and English third grade scores in math are more dissimilar than those obtained from the reading equations. Comparison of Figure 1 and Figure 2 illustrates important differences in the effect of third grade scores on fourth grade scores for reading and math. In contrast to the results obtained for reading using the same equation, higher scores on the third grade Spanish language test do not have much effect until the third grade score is 30 or above. The graph for the math test exhibits an even greater tendency for students with very low scores to catch up than was true for reading.

Obtaining Student-Teacher Links for Additional Grades/Years

The complete set of seven NAPT tests will enable us to learn a great deal more about the contributions of individual teachers to individual student achievement. To provide convincing results about teacher effects, however, multiple observations for the same teacher are needed. The full complement of NAPT tests will provide only four tests with both a current and prior test score. Worse yet, only two will provide both a current and prior score for the same teacher and both of them suffer from the disadvantage of having an interval between the current and prior test scores that reflects the effects of two teachers. The solution, of course, is to obtain test administrator, teacher name matches for the much larger number of TAAS tests.

By Christmas TSMP should include eight years of data for five cohorts of students and 32 test/grade/years of standardized test data grades shown in the upper left hand corner of Table 11. This portion of the table is identical to Table 5. For reasons that will become clearer subsequently, we hope to add three years of data, seven additional cohorts and 30 TAAS grades/years to TSMP. These cohorts and years comprise the rest of Table 11.

TEA's TAAS records do not include test administrator names. As a result, we cannot use the procedures we have used in creating the NAPT student-teacher links to obtain TAAS student-teacher links. Fortunately, the test records maintained by TEA's contractor, National Computer Systems (NCS), do include this information and TEA has agreed, in principal, to allow us to contract with NCS to complete these name matches. This is likely to be a fairly expensive and complex task and we have just begun discussions with NCS about how it might be done and its likely cost. Nonetheless, I am quite hopeful that I will be able to develop procedures that both NCS and TEA can agree to for completing the work.[3]

As these data illustrate, large gains would come from adding years and cohorts to TSMP. By adding the cohorts and years shown in Table 11, the number of grades/years with student-teacher links increases from 28 to 57 and the numbers with both a prior test

[3] I worked with NCS in the past on a similar task that involved obtaining TECAT scores for the individual teachers included in TSMP. TEA had information on which of the teachers took TECAT and whether they passed or failed. They did not have the scores. NCS's contract with TEA, moreover, stipulated that TEA could not have the scores and NCS was concerned that they could be sued if TEA somehow obtained them. As a result, I had to develop a procedure that provided us with the TECAT scores in a form that enabled us to match them with our teacher files and which at the same time prevented TEA from accessing them.

score and student-teacher links increases from 23 to 44. A final caveat is in order. As grade levels increase, the assumptions that all or most academic subjects are taught by a single classroom teacher and that the test administrator is that teacher become increasingly tenuous. This means the student-teacher links obtained by matching test administrator and teacher names are likely to be most valid for grades 3, 4 and 5 and of declining utility for grades 6 and beyond. Nonetheless, the augmented TSMP would provide 13 test/years of data for the third grade, 11 for the fourth and 10 for the fifth. Even if we determine that it is not worth completing this procedure for grades 7-10, we will have more than enough data to complete path-breaking research on the determinants of teacher effectiveness.

In addition to working with TEA to create approximate student-teacher matches from NAPT and TAAS records, we have began an even more time consuming, but ultimately superior approach. This is persuading individual districts to provide us with data on student-teacher links. Our efforts to obtain these critical data are part of an extensive outreach effort to strengthen TSMP. In their current form TSMP data permit analyses of a sort hardly imagined heretofore. Nonetheless, if we are able to add specific student links, early test scores and other data that may be obtained from individual districts the analytical power of TSMP will be greatly increased. O'Brien's (1998) paper for this meeting illustrates the kinds of gains that are provided by augmenting TSMP in this way.

The work on teacher quality and teacher effects in Tennessee and Dallas provide good models for the kinds of analyses that would be possible with an augmented TSMP if individual student-teacher links can be obtained. My reanalysis of Dallas data on teacher effectiveness provides strong support for the view that individual teachers can have a huge effect on individual student achievement.

The exploratory analyses of the impact of peers and teachers on individual student achievement presented in this paper indicate that useful analyses should be possible from the full set of seven NAPT tests we are obtaining student-teacher matches for. They also, however, illustrate the importance of obtaining data that will provide multiple years of data for a large number of teachers. Augmenting TSMP and obtaining individual student teacher links for the large number of TAAS tests would satisfy that goal.

References

Hanushek, Eric A. 1981. “Throwing Money at Schools,” Journal of Policy Analysis and Management, Vol. 1, No. 1, 19-41.

Hanushek, Eric A. 1989. “The Impact of Differential Expenditures on School Performance,” Educational Researcher (May).

Jordon, Heather R., Robert Mendro and Dash Weerasinghe. 1997. “Teacher Effects on Longitudinal Student Achievement: A Report on Research in Progress,” Presented to the CREATE Annual Meeting (July).

Kain, John F. and Daniel M. O’Brien. 1998a. “How Much Has Moving to the Suburbs Increased African-American Educational Opportunities,” Paper prepared for the meetings of the American Economic Association, Chicago, Illinois, January 5, 1998.

Kain, John F. and Daniel M. O’Brien. 1998a. “A Longitudnal Assessment of Reading Achievement: Evidence for the Harvard/UTD Texas Schools Project,” Presented at the Green Center Conference, “Achieving Universal Literacy,” The University of Texas at Dallas, April 2-4, 1998.

Murnane, Richard. J. The Impact of School Resources on the Learning of Inner City Children. Cambridge, MA: Ballinger.

Murnane, Richard. J. and B. Phillips. 1981. “What Do Effective Teachers of Inner-City Children Have in Common?” Social Science Research, 10(1), 83-100.

Sanders, William L. and Sandra P. Horn. 1994. “The Tennessee Value-Added Assessment System (TVAAS): Mixed-Model Metholdogy in Educational Assessment,” Journal of Personnel Evaluation in Education 8:299-311.

Wright, S. Paul, Sandra P. Horn & William L. Sanders. 1997. “Teacher and Classroom Context Effects on Student Achievement: Implications for Teacher Evaluation,” Journal of Personnel Evaluation in Education 11:57-67.

Sanders, William L., A. M. Saxton and Sandra P. Horn. 1998. “The Tennessee Value-Added Assessment System (TVAAS): A Quantitative, Outcomes-Based Approach to

Educational Assessment," in Jerome Milliman (ed). Grading Teacher, Grading Schools: Is Student Achievement a Valid Evaluation Measure? Thousand Oaks, CA: Corwin Press.

Steven G. Rivkin, Eric A. Hanushek and John F. Kain, "Teachers, Schools and Academic Achievement," January 4, 1998 (Revised July 1998).

Table 1. Cumulative Effects of Teacher Sequences on Fifth Grade Math Scores for Two Metropolitan Systems Measured in Percentiles

Teacher Quality Sequence	System A	System B	A-B
Low-Low-Low	44	29	15
Low-Low-Avg	63	40	23
Low-Low-High	83	59	24
Avg-Avg-Low	61	39	22
Avg-Avg-Avg	80	50	30
Avg-Avg-High	92	70	22
High-High-High	96	83	13

Table 2. Fourth Grade Mean NCE Reading Scores by Year and 1993-1996 Differences for Six Teacher Quality Sequences

Teacher Quality Sequences	Students	Mean NCE Reading Scores by Year				1993-96 Change
		1993	1994	1995	1996	
111	24	57.2	40.8	33.3	33.4	-23.8
314	28	57.3	53.6	43.0	50.2	-7.1
535	30	56.6	60.5	52.4	59.7	3.1
555	9	53.1	58.9	54.0	55.0	2.0
324	31	53.1	53.7	47.4	50.0	-3.1
112	18	55.8	39.2	33.1	30.6	-25.2

Table 3. OLS Linear Regressions of 1996 Student Scores on 1993 Student Scores and Teacher Quality in Each Year (Based on JMW Data)

Reading Regressions

Explanatory Variables	4th Grade		5th Grade		6th Grade		7th Grade		8th Grade	
	Coef.	t stat	Coef.	t stat	Coef.	t stat	Coef.	t stat	Coef.	t stat
Read 93	0.58	7.5	0.48	6.0	0.49	8.2	0.57	10.9	0.53	5.4
Teacher 94	0.11	9.3	0.06	6.0	0.08	8.0	0.08	6.9	0.08	5.2
Teacher 95	0.09	6.4	0.07	4.9	0.07	6.9	0.03	2.5	0.04	2.5
Teacher 96	0.19	15.2	0.14	12.7	0.16	13.6	0.16	10.0	0.15	9.7
Constant	1.11	3.6	1.67	5.6	1.70	7.7	1.32	7.0	1.44	4.0
R Sq	0.83		0.83		0.87		0.88		0.67	

Math Regressions

Explanatory Variables	4th Grade		5th Grade		6th Grade		7th Grade		8th Grade	
	Coef.	t stat	Coef.	t stat	Coef.	t stat	Coef.	t stat	Coef.	t stat
Math 93	0.39	6.6	0.51	8.6	0.58	11.0	0.70	13.6	0.76	14.8
Teacher 94	0.07	7.9	0.06	6.8	0.10	11.3	0.07	8.1	0.05	5.6
Teacher 95	0.09	8.7	0.09	9.1	0.08	8.8	0.04	4.8	0.07	5.4
Teacher 96	0.20	19.8	0.19	19.3	0.16	17.3	0.15	11.8	0.12	11.8
Constant	2.13	9.4	1.60	6.8	1.39	6.7	0.92	4.8	0.68	3.6
R Sq	0.88		0.88		0.87		0.91		0.90	

Table 4. OLS Log-Log Regressions of 1996 Student Scores on 1993 Student Scores and Teacher Quality in Each Year (Based on JMW Data)

Reading Regressions

Explanatory Variables	4th Grade		5th Grade		6th Grade		7th Grade		8th Grade	
	Coef.	t stat	Coef.	t stat	Coef.	t stat	Coef.	t stat	Coef.	t stat
Read 93	0.65	8.2	0.51	6.2	0.53	8.6	0.48	8.8	0.44	4.7
Teacher 94	0.04	8.3	0.02	5.6	0.03	7.7	0.03	7.0	0.03	5.6
Teacher 95	0.03	5.4	0.03	4.4	0.03	5.8	0.01	2.9	0.02	3.4
Teacher 96	0.08	14.4	0.06	12.0	0.06	12.9	0.08	11.3	0.07	11.5
Constant	0.76	2.40	1.48	4.9	1.50	6.7	1.58	8.2	1.68	4.9
R Sq	0.81		0.82		0.86		0.89		0.72	

Math Regressions

Explanatory Variables	4th Grade		5th Grade		6th Grade		7th Grade		8th Grade	
	Coef.	t stat	Coef.	t stat	Coef.	t stat	Coef.	t stat	Coef.	t stat
Math 93	0.45	7.6	0.51	8.5	0.58	11.6	0.67	12.8	0.66	11
Teacher 94	0.03	7.2	0.03	6.9	0.04	11.7	0.03	7.6	0.02	5.5
Teacher 95	0.03	7.6	0.04	9.2	0.03	9.4	0.02	5.0	0.03	6.0
Teacher 96	0.08	19.1	0.08	18.9	0.06	18.5	0.06	12.1	0.05	11.8
Constant	1.78	7.90	1.52	6.6	1.28	6.5	0.94	4.8	0.98	4.6
R Sq	0.87		0.87		0.88		0.91		0.90	

Table 5. Total Students and Standardized Tests Included in the Texas Schools Data Base by Cohort, Grade and Test

			Cohort 1		Cohort 2		Cohort 3		Cohort 4		Cohort 5	
Year	Total Students (Enrollment)	Sem	Gr	Test/Grade	Gr	Test/ Grade	Gr	Test/ Grade	Gr	Test/ Grade	Gr	Test/ Grade
89-90		F	3		2		1		K		PK	
89-90	1,161,358	S	3		2		1		K		PK	
90-91	1,505,551	F	4		3	T-3	2		1		K	
90-91	1,391,735	S	4		3		2		1		K	
91-92	1,420,295	F	5		4		3	T-3	2		1	
91-92		S	5	N-5	4	N-4	3	N-3	2		1	
92-93	1,415,593	F	6		5		4		3	T-3	2	
92-93		S	6	N-6	5	N-5	4	N-4 & T-4	3	N-3	2	
93-94	1,428,908	F	7		6		5		4		3	
93-94		S	7	T-7	6	T-6	5	T-5	4	T-4	3	T-3
94-95	1,438,632	F	8		7		6		5		4	
94-95		S	8	T-8	7	T-7	6	T-6	5	T-5	4	T-4
95-96	1,459,220	F	9		8		7		6		5	
95-96		S	9		8	T-8	7	T-7	6	T-6	5	T-5
96-97		F	10		9		8		7		6	
96-97		S	10	T-10	9		8	T-8	7	T-7	6	T-6

Table 6. Number of TA Groups (TAGs) and Number of Students by Class Size: All and TA Groups with Teacher Matches

TA Group	All TA Groups			TA Groups w Teacher Matches		
Size (Students)	Number of TAGs	Number of Students	Students per TAG	Number of TAGs	Number of Students	Students per TAG
1	207	207	1.0	143	143	1.0
2	199	398	2.0	148	296	2.0
3	162	486	3.0	124	372	3.0
4	144	576	4.0	107	428	4.0
5	136	680	5.0	105	525	5.0
6	95	570	6.0	75	450	6.0
7	94	658	7.0	72	504	7.0
8	107	856	8.0	89	712	8.0
9	110	990	9.0	96	864	9.0
10	120	1,200	10.0	100	1,000	10.0
11	148	1,628	11.0	128	1,408	11.0
12	157	1,884	12.0	142	1,704	12.0
13	211	2,743	13.0	200	2,600	13.0
14	277	3,878	14.0	251	3,514	14.0
15	428	6,420	15.0	405	6,075	15.0
16	706	11,296	16.0	654	10,464	16.0
17	1,096	18,632	17.0	1,017	17,289	17.0
18	1,547	27,846	18.0	1,463	26,334	18.0
19	1,838	34,922	19.0	1,730	32,870	19.0
20	2,007	40,140	20.0	1,903	38,060	20.0
21	1,678	35,238	21.0	1,584	33,264	21.0
22	1,125	24,750	22.0	1,065	23,430	22.0
23	507	11,661	23.0	480	11,040	23.0
24	217	5,208	24.0	205	4,920	24.0
25	81	2,025	25.0	76	1,900	25.0
26	53	1,378	26.0	50	1,300	26.0
27	20	540	27.0	20	540	27.0
28	7	196	28.0	6	168	28.0
29	4	116	29.0	3	87	29.0
30	2	60	30.0	2	60	30.0
31-50	72	2,893	40.2	33	1,259	38.2
51-100	162	12,094	74.7	33	2,298	69.6
101-150	89	10,850	121.9	17	2,004	117.9
151-200	12	2,059	171.6	2	370	185.0
GT 201	8	2,015	251.9	1	219	219.0
All	13,826	267,093	19.3	12,529	228,471	18.2

Table 7. Number of Test Administrator Groups (TAGs) and Number of Students by Class Size with Valid Reading Scores: All and TAGs with Teacher Matches

TA Group	All TA Groups			TA Groups w Teacher Matches		
Size (Students)	Number of TAGs	Number of Students	Students per TAG	Number of TAGs	Number of Students	Students per TAG
1	54	54	1.0	30	30	1.0
2	56	108	1.9	42	82	2.0
3	45	137	3.0	36	107	3.0
4	53	197	3.7	36	131	3.6
5	51	254	5.0	44	217	4.9
6	43	253	5.9	35	200	5.7
7	52	354	6.8	40	275	6.9
8	67	538	8.0	58	474	8.2
9	86	776	9.0	77	698	9.1
10	102	1,003	9.8	89	866	9.7
11	125	1,401	11.2	111	1,248	11.2
12	141	1,697	12.0	133	1,589	11.9
13	195	2,458	12.6	185	2,332	12.6
14	261	3,608	13.8	236	3,262	13.8
15	406	6,111	15.1	385	5,785	15.0
16	664	10,666	16.1	618	9,924	16.1
17	1,038	17,708	17.1	963	16,446	17.1
18	1,471	26,422	18.0	1,390	24,999	18.0
19	1,758	33,306	18.9	1,657	31,378	18.9
20	1,907	38,118	20.0	1,806	36,147	20.0
21	1,570	33,426	21.3	1,486	31,610	21.3
22	1,071	23,404	21.9	1,017	22,193	21.8
23	473	10,860	23.0	449	10,327	23.0
24	201	4,754	23.7	191	4,509	23.6
25	68	1,794	26.4	65	1,720	26.5
26	50	1,267	25.3	48	1,220	25.4
27	19	473	24.9	19	473	24.9
28	7	174	24.9	6	151	25.2
29	3	73	24.3	2	48	24.0
30	2	50	25.0	2	50	25.0
31-50	67	2,622	39.1	31	1,157	37.3
51-100	153	11,352	74.2	29	2,117	73.0
101-150	83	10,307	124.2	17	1,939	114.1
151-200	12	1,942	161.8	2	369	184.5
GT 201	8	1,949	243.6	1	205	205.0
All	12,362	249,616	20.2	11,336	214,278	18.9

Table 8. Number Teachers and Bilingual, ESL and Special Education Teachers by TA Group Size

Class Size (Students)	Number of Teachers				Percent	
	All	Bilingual	ESL	Special Education	Bilingual & ESL	Special Education
1	179	1	3	154	2.2%	86.0%
2	153	8	1	129	5.9%	84.3%
3	129	3	1	98	3.1%	76.0%
4	112	4	3	84	6.3%	75.0%
5	89	11	0	49	12.4%	55.1%
6	70	8	1	40	12.9%	57.1%
7	73	11	0	36	15.1%	49.3%
8	87	11	2	35	14.9%	40.2%
9	95	15	3	16	18.9%	16.8%
10	97	22	2	8	24.7%	8.2%
11	128	31	3	9	26.6%	7.0%
12	129	27	2	9	22.5%	7.0%
13	199	40	4	5	22.1%	2.5%
14	246	28	4	8	13.0%	3.3%
15	405	33	8	5	10.1%	1.2%
16	653	37	12	5	7.5%	0.8%
17	1,016	71	14	0	8.4%	0.0%
18	1,461	79	21	2	6.8%	0.1%
19	1,730	80	21	3	5.8%	0.2%
20	1,902	110	47	3	8.3%	0.2%
21	1,584	92	40	2	8.3	0.1%
22	1,065	58	32	3	8.5%	0.3%
23	480	32	16	1	10.0%	0.2%
24	205	10	4	1	6.8%	0.5%
25	75	3	5	0	10.7%	0.0%
26	50	4	2	0	12.0%	0.0%
27	20	2	0	1	10.0%	5.0%
28	6	1	1	0	33.3%	0.0%
29	3	1	0	0	33.3%	0.0%
30	2	0	0	0	0.0%	0.0%
31	33	1	1	1	6.1%	3.0%
51	33	1	1	0	6.1%	0.0%
101	19	0	0	0	0.0%	0.0%
201	1	0	0	0	0.0%	0.0%
All	12,529	835	254	707	8.7%	5.6%

Table 9. Coefficients for Regular OLS and Fixed Effects Regressions of Fourth Grade Reading Scores for Classes with 10-30 Students

Explanatory Variables	All Students			Regular Students		
		Fixed Effects			Fixed Effects	
	OLS	Campus	Classroom	OLS	Campus	Classroom
Peer Read 91	0.27	0.31		0.33	0.39	
Eng Read 91	-0.67	-0.66	-0.60	-0.76	-0.75	-0.70
Eng Read 91 Sq	0.03	0.03	0.03	0.03	0.03	0.03
Span Read 91	-0.53	-0.54	-0.45	-0.47	-0.55	-0.41
Span Read 91 Sq	0.02	0.02	0.02	0.02	0.02	0.02
NA High Income	-0.37	-0.30	-0.37	-0.54	-0.50	-0.55
NA Low Income	-1.20	-1.01	-0.60	-0.78	-0.59	-0.32
NA Very Low Inc	-2.53	-2.50	-2.38	-2.64	-2.62	-2.42
Asian High Inc	1.21	0.84	0.70	1.21	0.85	0.72
Asian Low Inc	-0.47	-0.39	-0.34	-0.39	-0.40	-0.24
Asian Very Low Inc	-0.31	-0.10	0.05	-0.14	0.00	0.24
Black High Inc	-2.47	-2.44	-2.33	-2.44	-2.43	-2.30
Black Low Inc	-2.72	-2.59	-2.47	-2.60	-2.48	-2.40
Black Very Low Inc	-3.74	-3.51	-3.34	-3.64	-3.44	-3.29
Hisp High Inc	-1.35	-1.42	-1.36	-1.29	-1.40	-1.34
Hisp Low Inc	-1.78	-1.86	-1.75	-1.69	-1.81	-1.70
Hisp Very Low Inc	-2.37	-2.34	-2.23	-2.26	-2.31	-2.21
Anglo Low	-1.17	-0.96	-0.86	-1.15	-0.98	-0.88
Anglo Very Low Inc	-1.76	-1.49	-1.42	-1.69	-1.44	-1.39
Male	-0.37	-0.36	-0.34	-0.31	-0.31	-0.29
Age	-0.11	-0.10	-0.09	-0.11	-0.10	-0.09
District Move	-0.53	-0.44	-0.39	-0.56	-0.47	-0.42
W/i District Move	-0.57	-0.43	-0.43	-0.58	-0.48	-0.47
Prop No Score	-0.72	0.51		-1.08	-0.03	
Prop No Attempt	1.03	1.16		1.50	2.29	
LEP dummy	-0.84	-0.82	-0.96			
ESL dummy	0.13	0.02	-0.13			
Bilingual dummy	-0.25	-0.28	-0.42			
Special Ed dummy	-1.58	-1.64	-1.64			
Constant	16.59	15.10	23.46	16.05	14.05	24.70
Campus/ Classroom		2,662	12,068		2659	11,583
Obs	181,086	181,086	181,086	154,471	154,471	154,471
R-sq	0.54	0.56	0.61	0.48	0.51	0.56

Figure 1. Predicted Fouth Grade Reading Scores by Third Grade Reading Scores on the English and Spanish TAAS for Classrooms with 10-30 Students: Campus Fixed Effects for All Students

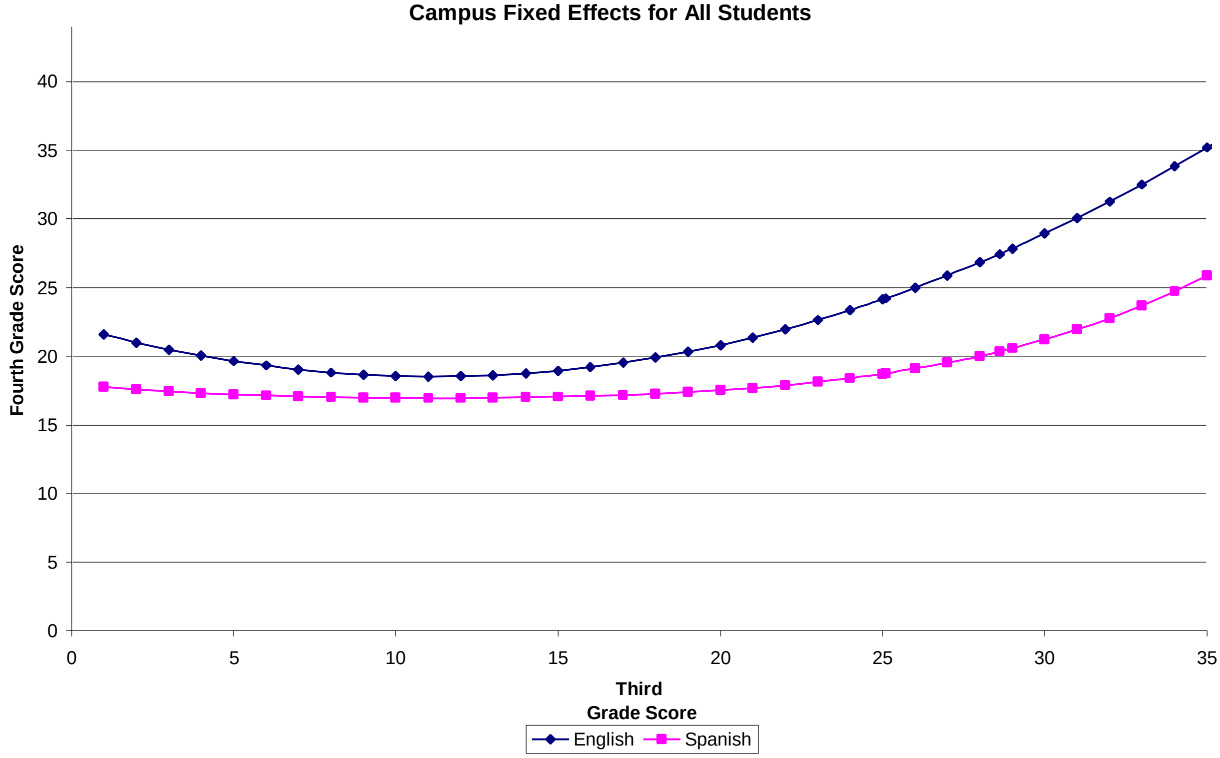

Table 10. Coefficients for Regular OLS and Fixed Effects Regressions of Fourth Grade Math Scores for Classes with 10-30 Students

Explantory Variables	All Students			Regular Students		
		Fixed Effects			Fixed Effects	
	OLS	Campus	Classroom	OLS	Campus	Classroom
Peer Math 91	0.19	0.43		0.25	0.51	
Eng Math 91	-1.30	-1.24	-1.18	-1.40	-1.32	-1.27
Eng Math 91 Sq	0.04	0.04	0.04	0.04	0.04	0.04
Span Math 91	-1.09	-1.10	-1.03	-1.03	-1.06	-0.94
Span Math 91 Sq	0.03	0.03	0.03	0.03	0.03	0.03
NA High Income	-1.14	-1.27	-1.32	-1.22	-1.30	-1.35
NA Low Income	-1.77	-1.92	-1.59	-1.78	-2.03	-1.46
NA Very Low Inc	-2.26	-1.81	-1.56	-2.08	-1.72	-1.25
Asian High Inc	2.26	1.84	1.77	2.06	1.67	1.63
Asian Low Inc	1.13	1.35	1.17	0.92	1.04	0.99
Asian Very Low Inc	0.72	1.13	1.23	0.78	1.21	1.26
Black High Inc	-3.36	-3.10	-3.05	-3.34	-3.14	-3.05
Black Low Inc	-3.64	-3.20	-2.98	-3.57	-3.19	-2.98
Black Very Low Inc	-5.15	-4.32	-4.07	-5.05	-4.31	-4.07
Hisp High Inc	-1.52	-1.45	-1.38	-1.42	-1.47	-1.41
Hisp Low Inc	-1.98	-1.77	-1.71	-1.90	-1.82	-1.71
Hisp Very Low Inc	-2.95	-2.44	-2.31	-2.66	-2.36	-2.19
Anglo Low	-1.65	-1.25	-1.10	-1.62	-1.25	-1.09
Anglo Very Low Inc	-2.41	-1.86	-1.73	-2.30	-1.79	-1.66
Male	0.24	0.26	0.31	0.29	0.32	0.36
Age	-0.28	-0.27	-0.26	-0.28	-0.27	-0.26
District Move	-0.73	-0.50	-0.41	-0.72	-0.51	-0.43
W/i District Move	-0.69	-0.51	-0.51	-0.67	-0.55	-0.57
Prop No Score	-3.12	-0.32		-3.11	-0.80	
Prop No Attempt	-0.43	0.88		-1.14	1.34	
LEP dummy	-1.73	-1.43	-1.38			
ESL dummy	0.42	0.02	-0.30			
Bilingual dummy	-0.69	-0.32	-0.74			
Special Ed dummy	-3.32	-3.33	-3.29			
Constant	36.04	25.78	40.37	35.81	24.15	41.92
Campus/ Classroom		2,662	12,064		2,658	11,576
Obs	182,210	182,210	182,210	154,316	154,316	154,316
R-sq	0.57	0.62	0.67	0.53	0.59	0.64

Figure 2. Predicted Fourth Grade Math Scores by Third Grade Math Scores on the English and Spanish TAAS for Classrooms with 10-30 Students: Campus Fixed Effects for All Students

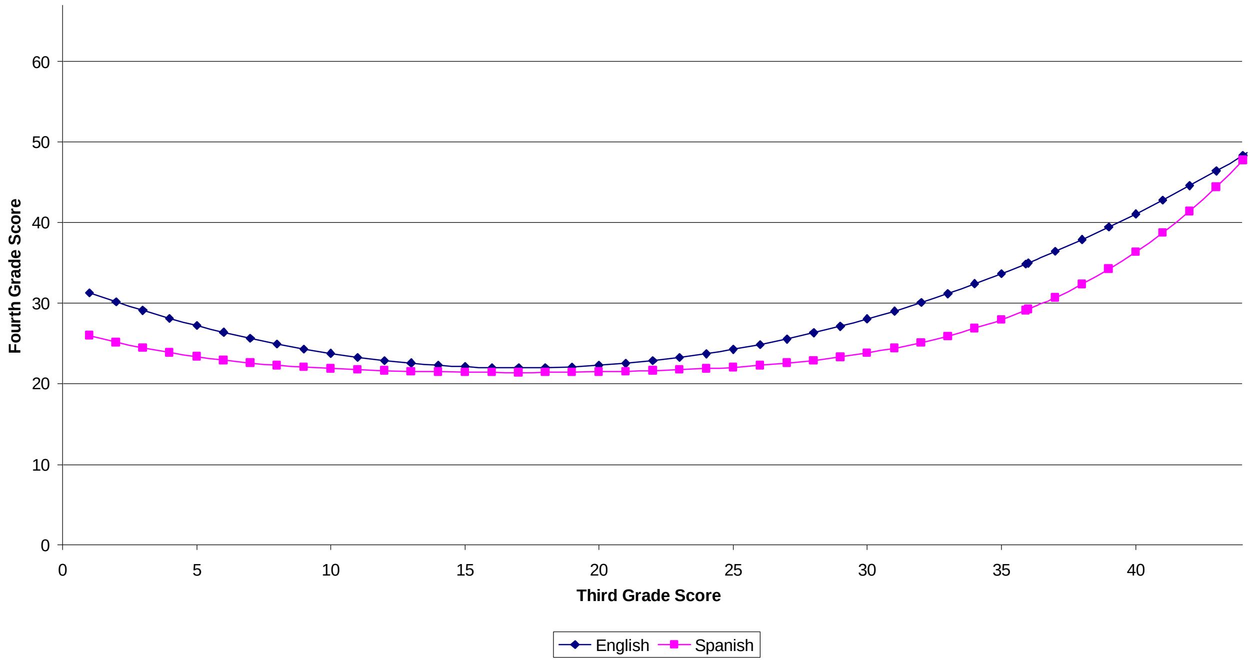

Table 11. Standardized Tests Included in the Texas Schools Microdata Panel (TSMP) by Cohort, Grade and Type of Test

		Cohort 1		Cohort 2		Cohort 3		Cohort 4		Cohort 5		Cohort 6		Cohort 7		Cohort 8		Cohort 9		Cohort 10		Cohort 11		Cohort 12	
Year	Sem	Gr	Test/ Grade	Gr	Test/ Grade	Gr	Test/ Grade	Gr	Test/ Grade	Gr	Test/ Grade	Gr	Test/ Grade	Gr	Test/ Grade	Gr	Test/ Grade	Gr	Test/ Grade	Gr	Test/ Grade	Gr	Test/ Grade	Gr	Test/ Grade
89-90	F	3		2		1		K		Pk															
89-90	S	3		2		1		K		Pk															
90-91	F	4		3	T-3	2		1		K		Pk													
90-91	S	4		3		2		1		K		Pk													
91-92	F	5		4		3	T-3	2		1		K		Pk											
91-92	S	5	N-5	4	N-4	3	N-3	2		1		K		Pk											
92-93	F	6		5		4		3	T-3	2		1		K		Pk									
92-93	S	6	N-6	5	N-5	4	N-4 & T-4	3	N-3	2		1		K		Pk									
93-94	S	7	T-7	6	T-6	5	T-5	4	T-4	3	T-3	2		1		K		Pk							
94-95	S	8	T-8	7	T-7	6	T-6	5	T-5	4	T-4	3	T-3	2		1		K		Pk					
95-96	S	9		8	T-8	7	T-7	6	T-6	5	T-5	4	T-4	3	T-3	2		1		K		Pk			
96-97	S	10	T-10	9		8	T-8	7	T-7	6	T-6	5	T-5	4	T-4	3	T-3	2		1		K		Pk	
97-98	S	11		10	T-10	9		8	T-8	7	T-7	6	T-6	5	T-5	4	T-4	3	T-3	2		1		K	
98-99	S	12		11		10	T-10	9		8	T-8	7	T-7	6	T-6	5	T-5	4	T-4	3	T-3	2		1	
99-00	S			12		11		10		9		8	T-8	7	T-7	6	T-6	5	T-5	4	T-4	3	T-3	2	
00-01	S					12		11		10	T-10	9		8	T-8	7	T-7	6	T-6	5	T-5	4	T-4	3	T-3

Note: In addition to the TAAS and NAPT tests listed above, we have also obtained the TEAMS first grade test which was given in Spring 1989. Regrettably, these individual test records include neither student names nor student ID numbers. As a result, they cannot be linked to other individual test records, although they are nonetheless valuable.

Table A-1. OLS and Fixed Effects Regression of Fourth Grade Reading Scores for Classes with 10-30 Students

Explanatory Variables	All Students						Regular Students					
	OLS		Fixed Effects				OLS		Fixed Effects			
			Campus		Classroom				Campus		Classroom	
	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t
Peer Read 91	0.27	46.4	0.31	41.8			0.33	48.2	0.39	43.1		
Eng Read 91	-0.67	-47.5	-0.66	-46.7	-0.60	-41.5	-0.76	-40.7	-0.75	-40.4	-0.70	-36.5
Eng Read 91 Sq	0.03	102.6	0.03	102.4	0.03	94.8	0.03	85.2	0.03	85.3	0.03	79.3
Span Read 91	-0.53	-24.3	-0.54	-24.8	-0.45	-19.7	-0.47	-3.5	-0.55	-4.1	-0.41	-2.8
Span Read 91 Sq	0.02	32.0	0.02	33.0	0.02	28.6	0.02	4.4	0.02	5.2	0.02	3.9
NA High Income	-0.37	-0.9	-0.30	-0.7	-0.37	-0.9	-0.54	-1.2	-0.50	-1.2	-0.55	-1.3
NA Low Income	-1.20	-1.2	-1.01	-1.0	-0.60	-0.6	-0.78	-0.7	-0.59	-0.6	-0.32	-0.3
NA Very Low Inc	-2.53	-4.9	-2.50	-4.9	-2.38	-4.6	-2.64	-4.6	-2.62	-4.6	-2.42	-4.2
Asian High Inc	1.21	10.3	0.84	7.0	0.70	5.8	1.21	9.7	0.85	6.7	0.72	5.7
Asian Low Inc	-0.47	-1.2	-0.39	-1.0	-0.34	-0.9	-0.39	-0.9	-0.40	-0.9	-0.24	-0.6
Asian Very Low Inc	-0.31	-1.5	-0.10	-0.5	0.05	0.2	-0.14	-0.6	0.00	0.0	0.24	1.0
Black High Inc	-2.47	-37.1	-2.44	-34.3	-2.33	-32.6	-2.44	-35.4	-2.43	-32.8	-2.30	-30.7
Black Low Inc	-2.72	-19.3	-2.59	-18.2	-2.47	-17.3	-2.60	-17.8	-2.48	-16.8	-2.40	-16.1
Black Very Low Inc	-3.74	-68.6	-3.51	-55.1	-3.34	-51.7	-3.64	-63.6	-3.44	-51.0	-3.29	-48.1
Hisp High Inc	-1.35	-26.0	-1.42	-25.5	-1.36	-24.2	-1.29	-23.3	-1.40	-23.5	-1.34	-22.4
Hisp Low Inc	-1.78	-20.7	-1.86	-20.8	-1.75	-19.5	-1.69	-17.8	-1.81	-18.4	-1.70	-17.0
Hisp Very Low Inc	-2.37	-56.4	-2.34	-45.7	-2.23	-42.9	-2.26	-50.0	-2.31	-41.4	-2.21	-39.0
Anglo Low	-1.17	-12.7	-0.96	-10.5	-0.86	-9.3	-1.15	-12.0	-0.98	-10.2	-0.88	-9.1
Anglo Very Low Inc	-1.76	-31.5	-1.49	-26.2	-1.42	-24.9	-1.69	-28.8	-1.44	-24.1	-1.39	-23.0
Male	-0.37	-13.6	-0.36	-13.5	-0.34	-12.9	-0.31	-10.6	-0.31	-10.6	-0.29	-10.1
Age	-0.11	-11.1	-0.10	-10.6	-0.09	-9.6	-0.11	-9.9	-0.10	-9.5	-0.09	-8.3
District Move	-0.53	-10.1	-0.44	-8.2	-0.39	-7.3	-0.56	-10.0	-0.47	-8.4	-0.42	-7.3
W/i District Move	-0.57	-15.7	-0.43	-8.5	-0.43	-8.4	-0.58	-14.4	-0.48	-8.5	-0.47	-8.3
Prop No Score	-0.72	-1.7	0.51	1.0			-1.08	-2.3	-0.03	-0.1		
Prop No Attempt	1.03	4.6	1.16	3.9			1.50	5.3	2.29	5.8		
LEP dummy	-0.84	-6.9	-0.82	-6.6	-0.96	-7.5						
ESL dummy	0.13	0.9	0.02	0.1	-0.13	-0.8						
Bilingual dummy	-0.25	-1.9	-0.28	-2.0	-0.42	-2.6						
Special Ed dummy	-1.58	-27.6	-1.64	-28.8	-1.64	-28.5						
Constant	16.59	64.4	15.10	52.1	23.46	116.2	16.05	49.5	14.05	38.3	24.70	93.0
Campus / Classroom			2,662	4.19	12,064	4.22			2,659	3.72	11,576	3.76
Obs	181,086		181,086		181,086		154,471		154,471		154,471	
R-sq	0.54		0.56		0.61		0.48		0.51		0.56	

Table A-2. OLS and Fixed Effects Regression of Fourth Grade Math Scores for Classes with 10-30 Students

Explanatory Variables	All Students						Regular Students					
	OLS		Fixed Effects				OLS		Fixed Effects			
			Campus		Classroom				Campus		Classroom	
	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t
Peer Read 91	0.19	26.3	0.43	44.2			0.25	30.0	0.51	45.9		
Eng Read 91	-1.30	-56.0	-1.24	-55.4	-1.18	-52.9	-1.40	-47.8	-1.32	-46.6	-1.27	-44.6
Eng Read 91 Sq	0.04	104.8	0.04	107.2	0.04	103.4	0.04	88.2	0.04	89.4	0.04	86.1
Span Read 91	-1.09	-37.0	-1.10	-38.4	-1.03	-35.1	-1.03	-6.9	-1.06	-7.4	-0.94	-6.0
Span Read 91 Sq	0.03	48.1	0.03	51.2	0.03	47.8	0.03	7.0	0.03	7.8	0.03	6.5
NA High Income	-1.14	-2.1	-1.27	-2.4	-1.32	-2.5	-1.22	-2.1	-1.30	-2.4	-1.35	-2.5
NA Low Income	-1.77	-1.3	-1.92	-1.5	-1.59	-1.3	-1.78	-1.3	-2.03	-1.5	-1.46	-1.1
NA Very Low Inc	-2.26	-3.4	-1.81	-2.8	-1.56	-2.4	-2.08	-2.8	-1.72	-2.4	-1.25	-1.7
Asian High Inc	2.26	14.6	1.84	12.2	1.77	11.8	2.06	12.7	1.67	10.5	1.63	10.3
Asian Low Inc	1.13	2.2	1.35	2.8	1.17	2.4	0.92	1.6	1.04	1.9	0.99	1.9
Asian Very Low Inc	0.72	2.7	1.13	4.3	1.23	4.7	0.78	2.5	1.21	3.9	1.26	4.1
Black High Inc	-3.36	-38.4	-3.10	-34.3	-3.05	-34.0	-3.34	-36.9	-3.14	-33.3	-3.05	-32.5
Black Low Inc	-3.64	-19.7	-3.20	-17.8	-2.98	-16.7	-3.57	-18.7	-3.19	-17.1	-2.98	-16.0
Black Very Low Inc	-5.15	-72.1	-4.32	-53.2	-4.07	-50.3	-5.05	-67.5	-4.31	-50.2	-4.07	-47.4
Hisp High Inc	-1.52	-22.3	-1.45	-20.5	-1.38	-19.7	-1.42	-19.6	-1.47	-19.4	-1.41	-18.7
Hisp Low Inc	-1.98	-17.4	-1.77	-15.6	-1.71	-15.2	-1.90	-15.2	-1.82	-14.5	-1.71	-13.7
Hisp Very Low Inc	-2.95	-53.7	-2.44	-37.6	-2.31	-35.6	-2.66	-45.2	-2.36	-33.3	-2.19	-31.0
Anglo Low	-1.65	-13.7	-1.25	-10.8	-1.10	-9.6	-1.62	-12.8	-1.25	-10.2	-1.09	-9.0
Anglo Very Low Inc	-2.41	-33.0	-1.86	-25.9	-1.73	-24.3	-2.30	-29.9	-1.79	-23.7	-1.66	-22.1
Male	0.24	6.8	0.26	7.8	0.31	9.5	0.29	7.4	0.32	8.7	0.36	10.0
Age	-0.28	-22.0	-0.27	-22.3	-0.26	-21.4	-0.28	-19.5	-0.27	-19.9	-0.26	-18.6
District Move	-0.73	-10.6	-0.50	-7.4	-0.41	-6.1	-0.72	-9.8	-0.51	-7.1	-0.43	-6.0
W/i District Move	-0.69	-14.4	-0.51	-7.9	-0.51	-8.0	-0.67	-12.7	-0.55	-7.7	-0.57	-8.1
Prop No Score	-3.12	-5.5	-0.32	-0.5			-3.11	-5.0	-0.80	-1.1		
Prop No Attempt	-0.43	-1.4	0.88	2.3			-1.14	-3.0	1.34	2.7		
LEP dummy	-1.73	-10.9	-1.43	-9.1	-1.38	-8.7						
ESL dummy	0.42	2.1	0.02	0.1	-0.30	-1.5						
Bilingual dummy	-0.69	-3.9	-0.32	-1.8	-0.74	-3.6						
Special Ed dummy	-3.32	-46.7	-3.33	-48.8	-3.29	-48.4						
Constant	36.04	75.1	25.78	49.4	40.37	104.4	35.81	60.6	24.15	37.6	41.92	84.6
Campus			2,662	9.14	12,064	4.22			2,658	7.87	11,576	3.76
Obs	182,210		182,210		182,210		154,316		154,316		154,316	
R-sq	0.57		0.62		0.67		0.53		0.59		0.64	

Table A-3. Regular OLS and Fixed Effects Regression of Fourth Grade Reading Scores for the Entire Sample

Explanatory Variables	All Students						Regular Students					
	OLS		Fixed Effects				OLS		Fixed Effects			
			Campus		Classroom				Campus		Classroom	
	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t
Peer Math 91	0.27	47.7	0.31	42.4			0.32	49.3	0.38	43.5		
Eng Math 91	-0.69	-51.5	-0.67	-50.5	-0.61	-45.6	-0.78	-44.4	-0.76	-43.8	-0.71	-40.0
Eng Math 91 Sq	0.03	110.7	0.03	110.4	0.03	103.0	0.03	92.4	0.03	92.2	0.03	86.4
Span Math 91	-0.54	-26.2	-0.56	-26.7	-0.47	-21.3	-0.48	-3.6	-0.55	-4.2	-0.41	-2.8
Span Math 91 Sq	0.02	33.9	0.02	34.9	0.02	30.2	0.02	4.4	0.02	5.2	0.02	3.9
NA High Income	-0.40	-1.0	-0.34	-0.9	-0.41	-1.0	-0.59	-1.4	-0.54	-1.3	-0.60	-1.4
NA Low Income	-1.84	-1.9	-1.65	-1.7	-1.31	-1.4	-1.73	-1.7	-1.55	-1.5	-1.35	-1.3
NA Very Low Inc	-2.50	-5.2	-2.43	-5.0	-2.32	-4.8	-2.59	-4.9	-2.51	-4.7	-2.33	-4.3
Asian High Inc	1.13	10.4	0.78	7.1	0.66	6.0	1.10	9.7	0.76	6.6	0.65	5.6
Asian Low Inc	-0.62	-1.7	-0.49	-1.4	-0.43	-1.2	-0.46	-1.1	-0.40	-1.0	-0.24	-0.6
Asian Very Low Inc	-0.40	-2.0	-0.20	-1.0	-0.08	-0.4	-0.26	-1.1	-0.11	-0.5	0.07	0.3
Black High Inc	-2.48	-39.5	-2.45	-36.7	-2.36	-35.1	-2.45	-37.7	-2.44	-35.0	-2.34	-33.2
Black Low Inc	-2.79	-20.7	-2.69	-19.9	-2.60	-19.1	-2.67	-19.2	-2.59	-18.5	-2.53	-17.8
Black Very Low Inc	-3.73	-72.0	-3.52	-58.5	-3.37	-55.3	-3.64	-66.9	-3.45	-54.3	-3.33	-51.7
Hisp High Inc	-1.36	-28.1	-1.43	-27.5	-1.39	-26.5	-1.31	-25.3	-1.41	-25.4	-1.37	-24.5
Hisp Low Inc	-1.80	-22.1	-1.88	-22.4	-1.79	-21.1	-1.70	-19.0	-1.82	-19.6	-1.72	-18.3
Hisp Very Low Inc	-2.38	-60.2	-2.37	-49.6	-2.27	-47.0	-2.27	-53.5	-2.34	-45.1	-2.26	-42.9
Anglo Low	-1.07	-12.5	-0.88	-10.3	-0.78	-9.2	-1.05	-11.7	-0.88	-9.8	-0.79	-8.8
Anglo Very Low Inc	-1.76	-34.1	-1.50	-28.6	-1.45	-27.4	-1.71	-31.3	-1.46	-26.4	-1.42	-25.4
Male	-0.36	-14.3	-0.35	-14.1	-0.34	-13.6	-0.32	-11.4	-0.31	-11.3	-0.30	-10.9
Age	-0.12	-12.6	-0.11	-12.0	-0.10	-11.0	-0.12	-11.3	-0.11	-10.8	-0.10	-9.7
District Move	-0.53	-10.9	-0.45	-9.0	-0.41	-8.3	-0.56	-10.9	-0.50	-9.4	-0.45	-8.4
W/i District Move	-0.55	-16.1	-0.42	-8.6	-0.41	-8.4	-0.55	-14.9	-0.47	-8.7	-0.45	-8.3
Prop No Score	-0.69	-1.7	0.46	0.9			-0.99	-2.2	-0.05	-0.1		
Prop No Attempt	1.12	5.3	1.07	3.8			1.61	6.1	2.27	5.9		
LEP dummy	-0.79	-6.8	-0.73	-6.1	-0.85	-6.9						
ESL dummy	0.11	0.8	-0.03	-0.2	-0.18	-1.1						
Bilingual dummy	-0.26	-2.0	-0.32	-2.3	-0.41	-2.6						
Special Ed dummy	-1.56	-29.1	-1.61	-30.3	-1.62	-30.1						
Constant	16.94	69.8	15.45	56.0	23.67	123.7	16.51	54.2	14.42	41.2	24.88	99.8
Campus			3,065	4.08	12,482	2.65			3,062	3.63	12,186	2.45
Obs	206,155		206,155		206,181		176,804		176,804		176,812	
R-sq	0.54		0.56		0.60		0.48		0.51		0.55	