Equality of Educational Opportunity Revisited

John F. Kain and Kraig Singleton

Henry Lee Professor of Economics and Professor of Afro-American Studies at Harvard University, and Research Associate, Department of Afro-American Studies at Harvard University, respectively. *The research reported in this paper was assisted by a grant from The Spencer Foundation. The data presented, the statements made, and the views expressed are solely the responsiblility of the authors.*

The title of this paper refers to *Equality of Educational Opportunity*, a pathbreaking study of the nation's public schools prepared for the U. S. Office of Education nearly 30 years ago (Coleman et al. 1966). The most obvious similarity between the study described in this paper and the Coleman Report, as the earlier study is popularly known, is the extent of their data collection efforts.

Mosteller and Moynihan (1972, p. 5) describe the Coleman Report as the "second largest social science research project in history," noting that the survey (EEOS) it relied upon tested "Some 570,000 school pupils" and "some 60,000 teachers" and gathered elaborate "information on the facilities available in some 4,000 schools." In comparison, the research on Texas elementary schools described in this paper is based on five years of micro panel data for more than 1.8 million children in five student cohorts attending more than 4,500 elementary schools during the years 1990 to 1994. We have obtained results of five statewide achievement tests for two cohorts and two tests for a third, as well as extensive micro data for more than 235,000 individuals teaching in grades pre-K through 8 during the 1990–94 period. The measures of school inputs used in this paper are constructed from individual data for 230,697 individuals employed by Texas public schools in grades pre-K through 8 at any time during the years 1990 to 1994.

In the manner of the Coleman Report, this paper examines (1) the extent of racial segregation both among and within districts; (2) differences in achievement on standardized tests among five racial/ethnic groups (African-Americans, Anglos (non-Hispanic Anglos), Asian-Americans, Hispanics, and Native Americans); and (3) the extent and nature of the variation in several school quality measures within individual school districts. The Coleman Report also estimated educational production functions, in an effort to quantify the relationship between student achievement and the quantity and quality of school inputs and other factors. This paper does not, although their estimation is a major goal of the larger study of which the research described here is a part.

Most of the analyses included in the Coleman Report were for the 6th grade and above. In contrast, our study is limited to grades pre-K through 6. This was a deliberate decision that reflects our beliefs that the educational disadvantage of minority children has its origins, and must be corrected, in the early grades, when schools all too often fail to provide them with the basic skills they need to succeed in school and subsequently in the workplace (Farkas et al. 1995). In addition, we anticipated that the greater homogeneity of elementary schools would make it more likely that we would be successful in our efforts to distinguish the effects of school inputs, peers, family background, and other factors on achievement.

Principal Findings of the Coleman Report

Coleman et al. (1966) found that children attending the nation's schools were highly segregated by race. They also expected to find large quality differences between predominantly minority and predominantly majority schools. In an interview published in the *Southern Education Report* (November–December 1965), Coleman, anticipating the results of the Coleman Report, stated:

> [T]he study will show the difference in the quality of schools that the average Negro child and the average white child are exposed to. You know yourself that the difference is going to be striking. And even though everybody knows there is a lot of difference between suburban and inner city schools, once the statistics are there in black and white, they will have a lot more impact (cited in Mosteller and Moynihan 1972, p. 8).

The expectations of Coleman and others were not realized. Mosteller and Moynihan (1972, p. 8), summarize Coleman et al.'s (1966) findings on differences in school inputs as follows: "[W]hile there are reported differences in those available to the majority as against the minority groups, they are surprisingly small differences. And while on balance the differences favor the majority, it is by no means the case that they consistently do so."

Coleman et al. (1966) did not stop with quantifying the variation in school inputs among schools with different racial and ethnic compositions. They included school input measures and other explanatory variables in educational production functions designed to explain variations in individual student achievement. Setting forth a view that was novel at the time but has recently become more widely accepted as a result of Hanushek's several survey articles (1981, 1986, and 1989), they found that variations in the school input measures included in their analysis had very little impact on student performance on standardized tests.

A fair amount of confusion arose in subsequent discussions of the Coleman Report because its authors did not classify teachers or peers as school inputs. They did, however, include these variables in their educational production functions. As the following summary of the report's findings by Mosteller and Moynihan (1972, p. 20) reveals, Coleman et al. (1966) also found that teachers had little effect on achievement.

> Teachers appeared to matter, at least for Negroes. A list of variables concerning such matters as teachers' scores on a vocabulary test, their own level of education, their years of experience, showed little relation to achievement of white students, but some for Negroes, and increasingly with higher grade levels. Even so, none of these effects was large; the between-school variance was so little to begin with, dividing it up, parceling it out between this factor and that, produced unimpressive results at best, and demoralizing at worst.

Turning to peer effects, Coleman et al. (p. 302) stated that "*Attributes of other students account for more variation in the achievement of minority group children than do attributes of staff*" [their italics], adding that "a pupil's achievement is strongly related to the educational backgrounds and aspirations of the other students in the school" and that "children from a given family background, when put in schools of different social composition, will achieve at quite different levels."

The Extent of Racial/Ethnic Segregation: Then and Now

Coleman et al. found that the nation's schools were highly segregated by race. Writing in 1966, they found that 65 percent of all Negro students in the 1st grade went to schools that were between 90 and 100 percent Negro. While it may not be true in some other parts of the country, public schools in Texas and in other parts of the South are now much less segregated by race than they were at the time the Coleman Report was completed (Welch et al. 1987). Table 1 shows the percentages of black, black and Hispanic, and Anglo students attending campuses with varying percentages of their enrollments of the same racial/ethnic mix. (Throughout this paper, blacks are defined as

Table 1

Percent of Black, Black plus Hispanic, and Anglo Students[a] by Campus Percent Black, Black plus Hispanic, and Anglo: State of Texas and Large Texas MSAs

(1) Campus Percentage	(2) Percent Black Students 1990	(3) Percent Black Students 1994	(4) Percent Black + Hispanic Students 1990	(5) Percent Black + Hispanic Students 1994	(6) Percent Anglo Students 1990	(7) Percent Anglo Students 1994
Entire State						
0–10	12.5	14.2	1.7	1.6	1.5	1.3
11–20	15.4	17.2	5.1	5.7	1.9	2.2
21–30	13.8	13.4	8.1	8.4	3.3	3.1
31–40	10.2	10.6	9.3	10.6	5.2	5.0
41–50	7.3	10.9	9.2	11.3	8.0	7.2
51–60	7.1	6.5	10.0	10.1	9.5	10.5
61–70	5.3	4.9	8.4	8.5	12.6	14.1
71–80	5.3	5.2	7.9	8.2	17.1	16.8
81–90	6.7	6.0	7.0	8.2	23.0	24.5
>90	16.5	11.1	33.4	27.4	17.0	15.3
All	100.0	100.0	100.0	100.0	100.0	100.0
>50	40.8	33.6	66.7	62.4	79.2	81.2
Large MSAs						
0–10	12.3	13.9	1.4	1.5	1.8	1.6
11–20	15.2	15.5	4.6	4.4	2.2	2.7
21–30	11.7	11.6	5.9	6.2	3.9	3.8
31–40	6.8	7.4	6.4	7.8	5.0	5.3
41–50	5.3	9.3	6.6	7.4	8.0	6.8
51–60	7.1	6.5	8.2	7.7	8.6	9.4
61–70	3.7	5.3	6.5	6.3	12.5	13.8
71–80	5.6	6.6	6.9	9.6	17.3	16.1
81–90	8.5	8.0	7.5	10.0	23.7	26.3
>90	23.7	16.0	46.0	39.1	17.0	14.2
All	100.0	100.0	100.0	100.0	100.0	100.0
>50	48.7	42.4	75.1	72.7	70.4	70.4

[a]Data for 1990 cover pre-K to grade 3 enrollments, 1994 data grades 3 to 7.

non-Hispanic blacks, Anglos as non-Hispanic whites, and all Hispanics as Hispanics.) As the data in the second column of Table 1 indicate, only 16.5 percent of black children enrolled in grades pre-K through 3 in 1990 attended schools that were 90 percent or more black. The same statistic for black students enrolled in grades 3 to 7 in 1994 was only 11.1 percent.

Some care should be taken in reading Table 1 because its interpretation depends on which column is being considered. Campus percentage (column (1)) refers to campus percent black for columns (2) and (3), to campus percent black plus Hispanic for columns (4) and (5), and to campus percent Anglo for columns (6) and (7). Coleman and his coauthors found that 80 percent of all Anglo 1st-grade students went to schools that were between 90 and 100 percent Anglo in 1967. The situation is very different in Texas today; only 17.0 percent of Anglo students in 1990 and 15.3 percent in 1994 attended schools that were more than 90 percent Anglo.

The bottom panel of Table 1 reveals that racial concentration is much greater in the seven largest Texas MSAs. For these areas, the percentage of black students in the respective grades who attended public schools that were more than 90 percent black was 23.7 in 1990 and 16.0 in 1994. While these shares are significantly higher than the statewide figures, they are much lower than they were nationwide in 1967.

With the rapid growth of Texas's low-income Hispanic population, the use of campus percent black as the sole indicator of school racial composition increasingly is seen as incomplete and possibly misleading. The two columns labeled "Percent Black + Hispanic Students" in Table 1 show the impact of including both blacks and Hispanics in the definition of the campus minority share. When this is done, the 1994 percent of black and Hispanic students (grades 3 to 7) who are enrolled at campuses in large MSAs that are greater than 90 percent black plus Hispanic is seen to be to 39.1 percent; the same statistic using a greater than 50 percent cut-off is 72.7 percent.

Texas is a large and heterogeneous state with great racial and ethnic diversity by region and metropolitan area. The racial/ethnic composition of its public schools cannot be meaningfully assessed without an understanding and appreciation of this diversity. Table 2 gives 1994 enrollment shares (grades 3 to 7) for the state's four largest racial/ethnic groups for the entire state, for large and small metropolitan areas, for non-metropolitan and rural areas, and for the central cities and suburbs of each of the state's seven largest metropolitan areas. Dallas appears twice in the MSA list. Greater Dallas combines districts in the Dallas MSA plus 13 districts (164 campuses) in the eastern part of the Ft. Worth MSA that many individuals employed in Dallas view as reasonable residential and school choices.

As these data reveal, in 1994 less than half of the state's grades 3 to 7 enrollment was Anglo. Hispanics, with 35.1 percent of total enrollment, were the largest minority, and Hispanics and African-Americans combined slightly outnumbered Anglos. Asian-Americans were only 2.1 percent of total state enrollment in 1994, but their numbers are growing rapidly. Native Amer-

Table 2
Shares of Total Grades 3 to 7 School Enrollments in Texas, by Race and Ethnicity, in 1994

Area	Black	Hispanic	Black + Hispanic	Anglo	Asian
Entire State	14.1	35.1	49.3	48.4	2.1
Large MSAs	16.3	34.4	50.7	46.1	3.0
Small MSAs	12.4	39.8	52.3	46.2	1.3
Non MSAs	9.6	31.6	41.2	58.2	.4
Rural	7.1	21.2	28.4	70.2	1.1
Individual Metropolitan Areas					
Austin	11.8	29.0	40.8	56.9	2.1
Corpus Christi	4.4	60.9	65.3	33.7	.7
Dallas	20.8	20.9	41.6	54.6	3.3
Greater Dallas	18.1	18.4	36.5	59.6	3.5
El Paso	3.3	80.4	83.7	15.5	.7
Fort Worth	14.5	16.1	30.7	66.0	3.0
Houston	21.9	30.2	52.0	43.4	4.5
San Antonio	8.0	59.9	67.8	30.9	1.1
Central City Districts					
Austin	18.7	37.9	56.6	41.3	1.9
Corpus Christi	6.0	66.7	72.7	26.3	.7
Dallas	43.6	40.1	83.6	14.2	1.7
El Paso	5.0	73.3	78.3	20.6	1.0
Fort Worth	33.8	33.7	67.5	30.2	2.2
Houston	35.6	49.6	85.2	12.2	2.6
San Antonio	11.3	82.6	93.8	5.7	.4
Suburban Districts					
Austin	5.8	21.3	27.1	70.5	2.2
Corpus Christi	2.7	54.3	57.0	42.0	.8
Dallas	11.8	13.3	25.2	70.5	3.9
Greater Dallas	10.9	12.2	23.1	72.5	4.0
El Paso	2.0	85.8	87.8	11.6	.4
Fort Worth	7.2	9.5	16.7	79.6	3.4
Houston	16.3	22.4	38.8	55.8	5.2
San Antonio	7.0	52.8	59.7	38.8	1.3

icans, who are not included in the table, made up only 0.2 percent of the state's population. These data also indicate that Anglos are disproportionately found outside of metropolitan areas and in rural areas, where they make up 58.2 and 70.2 percent respectively of the school population in grades 3 to 7.

African-Americans disproportionately live in large metropolitan areas, where they accounted for 16.3 percent of total enrollments in grades 3 to 7 in 1994. Similarly, Hispanics accounted for 34.4 percent, and in combination with African-Americans made up 50.7 percent of the enrollment in grades 3 to 7 in these areas. Central city districts differ greatly in terms of the racial composition of their public schools. The Dallas Independent School District (ISD) has the highest concentration of blacks, 43.6 percent of total grades 3 to 7 enrollments, whereas El Paso, with only 5.0 percent black, has the smallest.

Anglo shares of central city enrollments in these grades in 1994 varied from a low of 5.7 percent for the San Antonio ISD to a high of 41.3 percent for the Austin ISD. Hispanics, the fastest-growing racial/ethnic group in Texas, comprised 33.7 percent of these students in the Ft. Worth ISD and 73.3 percent in the El Paso ISD in 1994. When the black and Hispanic populations are combined, the combined shares vary from 56.6 percent for the Austin ISD to 93.8 percent for the San Antonio ISD.

The data in Table 3 provide another way of looking at the racial composition of Texas elementary schools. In contrast to the previous discussion, which focused on the fractions of black, Hispanic, and Anglo students attending schools of varying racial compositions, these data show the distribution of schools according to racial composition. They are arguably more relevant to analyses of variations in school inputs among districts and campuses, the central research question considered in this paper.

Statistics on the fractions of schools with student bodies of varying racial composition reveal that about 7 percent of all Texas campuses have enrollments that are 50 percent or more black. For large MSAs, this fraction increases to nearly 10 percent, while the fraction for small MSAs (3.9 percent) is considerably below the statewide average. Nearly one-fourth of campuses in the seven largest central city districts, however, have enrollments that are more than 50 percent black. High levels of Hispanic concentration are much more common than high levels of black concentration. Using the same 50 percent cut-off, 26.4 percent of statewide campuses were more than 50 percent Hispanic in 1994, more than three times the rate for blacks. The figure is not much different for campuses located in large metropolitan areas. Nearly half (46.8 percent) of large central city campuses are more than 50 percent Hispanic and more than one-fourth are over 80 percent Hispanic.

The third panel in Table 3 shows the fraction of campuses in each area that have specified shares of black plus Hispanic students. These data reveal that 42 percent of campuses statewide, 48 percent of campuses in large metropolitan areas, and nearly 87 percent of campuses in the seven largest central city school districts have enrollments more than 50 percent black plus Hispanic. In the suburban rings of the

Table 3
Racial Composition of Texas Elementary School Campuses in 1994

	Entire State	Large MSAs	Large MSAs: Central Cities	Large MSAs: Suburbs	Small MSAs
Percent Black					
0	18.2	8.2	4.8	9.9	27.9
0–10	44.9	50.4	36.6	57.2	39.6
11–20	15.3	16.5	13.9	17.7	14.1
21–30	7.3	7.7	8.7	7.2	7.0
31–40	4.0	3.4	4.1	3.1	4.6
41–50	3.4	3.9	7.2	2.2	2.9
51–60	1.6	1.8	3.5	1.0	1.3
61–70	1.2	1.5	3.8	.4	1.0
71–80	1.1	1.5	4.0	.4	.6
81–90	1.2	1.8	5.2	.2	.5
91–100	1.9	3.2	8.3	.8	.5
All	100.0	100.0	100.0	100.0	100.0
>50	6.9	9.9	24.8	2.7	3.9
Percent Hispanic					
0	3.9	2.7	2.5	2.8	5.1
0–10	28.5	26.8	11.3	34.4	30.2
11–20	16.7	18.1	11.5	21.3	15.2
21–30	10.6	11.2	10.7	11.5	10.1
31–40	7.9	7.6	8.7	7.1	8.2
41–50	5.9	5.5	8.6	4.0	6.4
51–60	4.5	4.9	7.3	3.7	4.2
61–70	4.0	4.2	6.6	3.1	3.8
71–80	3.3	4.3	7.2	2.8	2.4
81–90	4.2	5.4	10.4	2.9	3.1
91–100	10.3	9.3	15.4	6.4	11.3
All	100.0	100.0	100.0	100.0	100.0
>50	26.4	28.0	46.8	18.9	24.9
Percent Black + Hispanic					
0	1.9	1.5	.5	2.0	2.3
0–10	12.4	12.0	.6	17.6	12.7
11–20	13.8	12.8	2.2	18.0	14.8
21–30	11.0	10.0	2.7	13.6	12.0
31–40	9.9	8.4	3.5	10.9	11.4
41–50	8.9	7.1	3.8	8.7	10.5
51–60	7.2	6.5	7.1	6.2	7.8
61–70	6.0	5.4	6.6	4.8	6.6
71–80	5.4	6.5	9.3	5.1	4.3
81–90	6.1	7.5	14.5	4.1	4.8
91–100	17.4	22.2	49.2	9.1	12.7
All	100.0	100.0	100.0	100.0	100.0
>50	42.1	48.1	86.6	29.3	36.2

seven largest metropolitan areas, enrollments in 29 percent of campuses are more than 50 percent black plus Hispanic.

Clearly, the levels of racial concentration in Texas elementary schools have declined by substantial amounts since Coleman and his coauthors completed their research. In addition, growing numbers of minority households residing in Texas metropolitan areas have moved to communities served by suburban school districts, particularly in the past decade. But even though increasing numbers of African-Americans have been moving to suburban communities in Texas, elementary school indexes of dissimilarity for Texas metropolitan areas are still quite high. (The index of dissimilarity compares the distributions of particular racial or income groups.) The black versus Anglo indexes vary from a high of 79 for Laredo to a low of 30 for Odessa, with mean values of 48 for the 21 small MSAs and 56 for the seven large MSAs. Black versus non-black and Hispanic versus non-Hispanic indexes of dissimilarity are somewhat lower.

Rivkin (1994) has shown that, as in other parts of the country, the continuing high levels of school segregation in Texas are largely due to even higher levels of residential racial segregation. It is therefore not particularly surprising to discover that the segregation indexes for Hispanic and Asian-American residents of Texas metropolitan areas, who are less highly segregated residentially than African-Americans, are lower as well (Farley 1993). San Antonio and El Paso, where Hispanic-Anglo segregation exceeds black-Anglo segregation by significant amounts, are exceptions. The explanation presumably is related to the small black and very large Hispanic population shares that characterize these areas.

Differences in Achievement by Race/Ethnicity and Income

As noted previously, the analyses presented in this paper are based in part on enrollment data for 1.8 million students who attended Texas elementary schools during the five-year period 1990 to 1994. Enrollments by year and grade for five student cohorts are shown in Table 4. We have obtained standardized test results for 12 tests/years/grades for students in cohorts 2, 3, and 4. While we did not acquire standardized test data for cohorts 1 and 5, the enrollment and attendance data for these cohorts will enable us to keep track of students who were retained in grade or double-promoted and who thus moved from one of the three central cohorts. When longitudinal samples are created by linking individual student records for two or more tests/years, as in Kain (1995) and Fergu-

Table 4
Snapshot Data for Enrollments by Year, Grade, and Cohort: Texas Elementary Schools, 1990 to 1994

	Cohort 5	Cohort 4	Cohort 3	Cohort 2	Cohort 1					
Year	Pre-K	Kindergarten	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Total
1990	71,185	250,887	297,318	272,845	269,123					1,161,358
1991		288,551	326,127	301,383	295,408	294,082				1,505,551
1992			303,173	282,566	279,476	278,636	276,444			1,420,295
1993				288,937	281,754	282,376	280,697	281,829		1,415,593
1994					289,073	284,179	284,120	285,186	286,350	1,428,908
Total	71,185	539,438	926,618	1,145,731	1,414,834	1,139,273	841,261	567,015	286,350	6,931,705

son and Ladd (1995), such students simply disappear from the analysis. This selective attrition may result in seriously misleading findings about student progress. In addition, in the absence of a student data base that is independent of the testing program, it is difficult to assess the apparently widespread practice of not testing some children who are expected to perform poorly on the tests (Orfield and Ashkinaze 1991).

The educational histories with multiple standardized test data that we are constructing for each student should be of great value to our efforts to determine the connections between school inputs (including teachers and peers) and individual student achievement. The mean reading scores by race/ethnicity and family income/poverty for the five statewide tests taken by cohort 2 students, shown in Table 5, provide a glimpse of the promise of these data. Cohort 2 students took five statewide achievement tests, including three different types of tests, between 1989 and 1994; the three tests were TEAMS (Texas Educational Assessment of Minimum Skills), TAAS (Texas Assessment of Academic Skills), and NAPT (Norm-referenced Assessment Program for Texas). NAPT is a norm-referenced test, while TEAMS and TAAS are criterion-referenced.

The achievement test data for cohort 2 students in Table 5 are mean scores for students taking each test in a particular year, rather than mean scores for a true panel of individual students.[1] Because significant numbers of students enter and leave the state, transfer to or from private schools, or are excused from taking these tests in one or more years, the composition of the samples used in calculating the mean scores may vary from one year to the next.

The top panel in Table 5 presents indexes of mean reading scores for each of the five race/ethnicity groups by grade/year, indexed to the statewide mean reading score for that year. Children participating in special education programs are omitted from these calculations, as are significant numbers of children who were excused from taking the examination for other reasons in each year. The first and third grade data also exclude the reading scores of Hispanic children who took a Spanish language version of these tests.

The indexes of mean reading scores in Table 5 reveal that Asian-Americans and Anglos have the highest reading scores, followed by the relatively small sample of Native Americans, who consistently score at about the statewide average. Blacks and Hispanics have the lowest scores, their relative scores are very similar in each year, and some evidence suggests that the achievement gaps between Asian-Americans/Anglos and African-Americans/Hispanics increase with years of school completed. While an increasing achievement gap of this kind between disadvantaged and advantaged groups is a widely reported finding, it is impossible in this instance, and very likely in many other studies that have reported this result, to be confident about the apparent deterioration of Hispanic scores, because of the widespread

[1] We are still completing the difficult, time-consuming, and frequently frustrating task of creating a multi-year linked student data base and adding the scores and other data from the tests to individual student records. At the time we prepared this paper, we had matched more than 97 percent of the 1994 test records and between 94.1 percent and 96.7 percent of the 1993 test records to the Public Education Information Management System enrollment/attendance records. We had less success with the five 1991 and 1992 tests; the rates of successful matches for them varied between 88.8 and 94.1 percent. A large fraction of the remaining non-matches, particularly for 1991 and 1992, are concentrated in a small number of districts. We are working with the Texas Education Agency in an effort to improve the match rates for these districts.

Table 5
Indexes of Student Mean Reading Scores, by Race/Ethnicity and Family Income/Poverty Level[a]: Cohort 2, Grades 1 and 3 through 6

Statewide Mean = 100	TEAMS Grade 1	TAAS Grade 3	NAPT Grade 4	NAPT Grade 5	TAAS Grade 6
By Race/Ethnicity					
Asian-American	107	105	111	108	110
Anglo	105	106	109	112	108
Black	94	93	87	88	89
Hispanic	94	93	88	86	91
Native American	99	101	99	100	100
All Race/Ethnicity	100	100	100	100	100
By Percent of Poverty Level					
More than 185 Percent	105	105	110	110	107
135–185 Percent	n.a.	n.a.	n.a.	97	98
Less than 135 Percent	n.a.	n.a.	n.a.	90	94
AFDC and like programs	n.a.	n.a.	n.a.	85	89
Less than 185 Percent	93	92	88	n.a.	n.a.
By Race/Ethnicity and Poverty Level					
Greater than 185% of Poverty Level					
Asian-American	109	108	114	115	114
Anglo	107	107	114	115	110
African-American	98	98	94	95	94
Hispanic	100	100	99	98	99
Native American	103	105	104	109	105
Less than 185% of Poverty Level					
Asian-American	100	98	95	92	101
Anglo	97	98	99	99	99
African-American	91	90	83	84	85
Hispanic	92	91	84	82	88
Native American	93	96	92	90	92

n.a. = not available.
[a]Family income/poverty level derived from eligibility for the school lunch program. See the text.

practice of excusing large numbers of low-achieving students, and particularly Hispanics, from achievement tests in the early grades.[2]

The family income/poverty level variables used in Table 5 are derived from information that indicates whether a particular child received a free or reduced-price lunch under the federal school lunch program. 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 belong to a family whose annual 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 refer to students belonging to high-income families (greater than 185 percent of the poverty level), middle-income families (135 to 185 percent of the poverty level), and low-income families (less than 135 percent of the poverty level).

The estimates in Table 5 are obtained from individual test records, and the definition of the school lunch variable differs somewhat among the several tests. Only two poverty/lunch categories are available for grades 1, 3 and 4, while three are available for grades 5 and 6. In the second panel, we use four family income/poverty categories for grades 5 and 6, but only two categories for grades 1, 3 and 4. Both 5th and 6th grade reading scores exhibit a consistent relationship with family income levels; the differences are especially large for the 5th grade scores. The reading scores of 5th grade students in the high-income category are 110 percent of the statewide average, while the reading scores of 5th grade students in the middle- and low-income categories are 97 and 90 percent of the statewide average, respectively. Fifth-grade students who qualified for a free lunch through their participation in AFDC or other welfare programs have a mean reading score of only 85 percent of the statewide average.

The bottom panel of Table 5 shows mean reading scores by race/ethnicity and two family-income categories, greater than and less than 185 percent of the poverty level. Stratification by family income level

[2] Of the 233,883 3rd-grade students with a reading score, 4.7 percent took a Spanish language test. A large fraction of these students are subsequently excused from taking the 4th grade test because of limited English proficiency; they are classified as LEP (Limited English Proficiency) in the 4th grade and thereby are excused. Including the raw reading score for these students in the calculations of mean reading scores changes only one number in Table 5, the mean reading score index for Hispanic 3rd graders from families whose incomes were less than 185 percent of the poverty level.

narrows the differences between Anglo reading scores and those of African-American and Hispanic students, but by no means eliminates them. Not too much should be made of this result; the two-category family income/poverty variable is only a crude index of socioeconomic differences.

Within- and Between-School Variations in Grade 6 Reading Scores

Mosteller and Moynihan (1972, p. 19) in their discussion of Coleman et al.'s findings place great emphasis on the fact that "90 percent of the variance in student achievement was found to lie *within*—not between—schools." Commenting further on this theme, they observe:

> EEO found that schools receive children who already differ widely in their levels of educational achievement. The schools therefore do not close the gaps between students aggregated into ethnic/racial groups. Things end much as they begin. . . . such findings might be interpreted to mean that "schools don't make any difference." This is absurd. Schools make a very great difference to children. . . . But given that schools have reached their present level of quality, the observed variation in schools was reported by EEO to have little effect upon school achievement. This actually means a large joint effect owing to both schools and home background (including region, degree of urbanization, socioeconomic status, and ethnic group), little that is unique to schools or home. They vary together.

Equation 1 in Table 6 presents the results of regressing individual TAAS reading scores for 228,051 sixth graders in this study on 16 explanatory variables that are measures of individual and family background characteristics of these students. These variables explain 18 percent of the variance in reading scores; they include 14 dummy variables for race/ethnicity by family income (Anglos eligible for free lunch is the omitted category in all three equations) plus the sex and age of each student. In Equation 2, we have augmented the individual student variables included in Equation 1 by three campus-level variables that measure parents' education and median household income. In contrast to the individual student variables, the percentages of each racial/ethnic group who were college graduates or did not complete high school and median family income for all groups are averages for the school campus zip code. While we plan to create more precise estimates of these family background and community variables by aggregating 1990 census block group data to individual attendance areas, these campus-level variables, following Ferguson and Ladd (1995), are based on 1990 Census tabulations for the zip code of each campus. Adding the three campus-level variables increases the R^2 to 0.19. In Equation 3, we replace the three campus/zip code variables with 1,986 campus dummies; the resulting campus fixed-effects specification increases the explained variance from 19 percent to 23 percent. When only campus dummies are included, the resulting equation explains 16 percent of the total variance in reading scores. In interpreting this result, it should be understood that "campus" measures both the effects of school inputs (facilities, teachers, and peers) and the effects of grouping children with similar racial/ethnic and other individual and family background characteristics.

The results for the fixed-effect equations indicate that holding the effects of campus, age, and sex constant, the representative low-income African-American student had 3.4 fewer right answers on the 6th grade reading test than a low-income Anglo student. This difference falls to 1.5 points for middle-income black students and to 1.2 points for high-income black students. High-income Anglo students have 2.7 more right answers than low-income Anglo students and 3.8 more right answers than high-income black students.

In considering the 6th-grade reading regressions, it should be understood that nearly 20 percent of all 6th-graders are not included in the analysis. Of this number 11.9 percent were excluded from our calculations because they were special education students and an additional 6.9 percent were excused from the test for a variety of other reasons, including absences on the day the test was given, LEP (Limited English Proficiency) exemption, ARD exception (for students with disabilities who were not already excluded from the sample because of their participation in special education classes), cheating, or illness. The fractions differ among ethnic/racial categories. More than 13 percent of Hispanics and 10.5 percent of Asian-American students were excused from taking the test, mostly because they were classified as LEP. African-Americans were somewhat more likely to be in special education classes (14.4 percent versus 11.9 percent for all students), but were less likely to be excused from taking the exam (4.3 percent versus 6.9 percent for all students). These issues, which deserve careful attention, will be examined after we have completed linking the data.

The most obvious missing variable from the 6th-

Table 6
Grade Six TAAS Reading Score Regressions

Explanatory Variables	Equation 1 Individual Characteristics		Equation 2 Individual plus Zip Code: Education & Income		Equation 3 Individual plus Campus Fixed-Effects	
	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
High Income						
Anglo	3.35	59.19	2.91	49.60	2.70	47.60
Asian-American	4.25	35.61	2.63	19.62	3.17	26.48
Native American	2.06	5.38	1.85	4.62	1.39	3.73
Black	−1.28	−16.43	−1.39	−17.21	−1.15	−14.07
Hispanic	.17	2.40	.27	3.34	.15	2.16
Middle Income						
Anglo	1.24	11.97	1.23	11.57	1.09	10.79
Asian-American	1.35	3.53	.48	1.23	.90	2.42
Black	−1.69	−11.10	−1.62	−10.37	−1.54	−10.25
Hispanic	−1.01	−10.06	−.72	−6.51	−.89	−8.80
Native American	−.36	−.35	−1.21	−1.09	−.10	−.10
Low Income						
Asian-American	.54	2.80	−.23	−1.13	.57	2.96
Black	−3.95	−54.76	−3.75	−50.74	−3.40	−44.53
Hispanic	−3.00	−49.55	−2.58	−33.26	−2.53	−38.50
Native American	−2.05	−3.91	−1.27	−2.30	−1.55	−3.03
Other Individual Characteristics						
Male	−.62	−22.70	−.63	−22.11	−.61	−22.93
Age	−1.38	−60.18	−1.37	−57.81	−1.28	−56.80
Campus Mean by Race/Ethnicity						
Percent College Grad			7.80	22.10		
Percent Less than High School			.18	1.05		
Campus Mean, All Households						
Median Income			.04	22.61		
Campus					F-stat. (1,986 225,914) =	9.07
Constant	45.73	164.11	44.08	150.92	44.63	163.57
Observations	228,051		212,601		227,917	
Variables	16		19		2,002	
R^2	.18		.19		.23	

Note: Anglos with low-income families is the omitted category for the race/ethnicity by family income/poverty level variables in the first three panels.

grade reading score regressions is any measure of each student's cognitive abilities or earlier educational experiences. Because we have not finished linking the test and enrollment data, we were unable to include these critical control variables. In an earlier exploratory analysis of TAAS reading scores for individual 4th-graders, however, Kain (1995, p. 44) found that the same student's 3rd-grade reading score by itself explained 44 percent of the variation in 4th-grade reading scores. Adding a large number of individual student and school input measures increased the explained variance by only an additional 5 percentage points to 50 percent, and these same variables without the lagged reading score explained about 28 percent of the total variance in individual 4th-grade reading scores.

Variations in School Quality Measures

Educational production function studies, including the Coleman Report, have had modest success at best in their efforts to quantify the relationship between school inputs and student achievement. At least two explanations are possible. First, in the spirit of the Coleman findings, public schools may not vary much in terms of those variables that affect student achievement on standardized tests. Or second, as Hanushek and Kain (1972, p. 117) argued in their critique of the Coleman Report, there may be important differences, but the crude measures that Coleman and his colleagues used in their analyses (and have been used in most subsequent educational production function studies) may not adequately account for these differences. This second possibility, which seems to us to be a very real one, was one of the reasons we decided to undertake this study.

Educational production function studies have had modest success at best in their efforts to quantify the relationship between school inputs and student achievement. This failure may be due to inadequate school-input quality measures.

The view that the failure of most studies of educational production functions to find significant school input effects may be due to inadequate school quality measures finds support in Hanushek's survey of 147 published "separately estimated educational production functions," which begins by observing that "*Teachers and schools differ dramatically in their effectiveness*" [his italics]. Hanushek (1986, p. 1159) then argues that the "very different impressions . . . left by the Coleman Report and indeed by a number of subsequent studies . . . have primarily resulted from . . . the difficulty in explicitly measuring components of effectiveness," adding that "existing measures of characteristics of teachers and schools are seriously flawed and thus they are poor indicators of the true effects of schools." Finally, he suggests that "when these measurement errors are corrected, schools are seen to have important effects on student performance."

When it comes to assessing the evidence concerning the relationship between particular school input measures and student achievement, however, Hanushek becomes much more negative. Commenting on the findings of his most recent survey of educational production function estimates, Hanushek concludes, based on his examination of 187 studies, that "the results are startlingly consistent in finding no strong evidence that teacher-student ratios, teacher education, or teacher experience have the expected positive effect on student achievement" (1989, p. 46).

Hanushek's negative conclusions about the effect of various school input variables on achievement have been strongly disputed by Ferguson (1991) in an influential paper that presents educational production function estimates obtained using aggregate (district-level) data for nearly 900 Texas school districts.[3] In both his Texas analysis and a more recent paper with Ladd based on both micro and aggregate campus-level data for Alabama, Ferguson suggests that differences in mean teacher test scores, average class size, the fraction of teachers with master's degrees, and per pupil expenditures account for a large part of the variation among districts and campuses in student achievement levels (Ferguson 1991; Ferguson and Ladd 1995). Ferguson's (1991) educational production function estimates for Texas schools obviously are highly relevant to our research.

Our review of earlier educational production function studies has made it clear to us that we should begin by making every effort to develop better school input measures than have been used in most earlier studies. Part of this effort entails a careful assessment of the extent of the variation among districts and campuses and, in particular, among campuses with different racial/ethnic and socioeconomic makeups. The selection of school input measures to be considered in this paper was determined by data processing considerations and by the findings of earlier educational production function studies. In the following sections, we examine the within-district variation of four types of school inputs by campuses of varying race/ethnic and income composition: (1) teacher scores on standardized tests, (2) the percentage of teachers with advanced degrees, (3) teacher experience, and (4) class size. We begin this analysis with composite reading and writing scores for Texas teachers, obtained by combining individual test results from seven different teacher certification exams.

[3] Ferguson's study used district-level aggregate data for all Texas school districts with complete data except the Dallas and Houston Independent School Districts, which he excluded because "the weighting scheme in the estimating procedure would have given these two cities too much influence over the results" (Ferguson 1991, p. 470).

Teacher Test Scores

Several educational production function studies have found that teacher verbal ability and high scores on other standardized tests had a significant effect on student achievement (Hanushek 1971 and 1972; Ferguson 1991; Ferguson and Ladd 1995; and Murnane 1975). Moreover, Hanushek (1986, p. 1164) has observed that "The closest thing to a consistent finding among the studies is that 'smarter' teachers, ones who perform well on verbal ability tests, do better in the classroom"; he adds, however, that "even for that, the evidence is not very strong."

Teacher verbal ability and high scores on other standardized tests have been found to have a significant effect on student achievement.

Ferguson (1991, p. 475) in his study of Texas school districts similarly found that "Teachers' language skills as measured by the TECAT score is the most important school input for both math and reading," and that "After the first grade, teacher scores on TECAT account for about one-fifth to one-quarter of all variation across districts in students' average scores on the TEAMS exam."[4] In addition, Ferguson and Ladd (1995, p. 35), in their study of Alabama schools, found that "The skills of teachers as measured by their test scores exert consistently strong and positive effects on student learning despite the fact that the data are limited and test scores are an imperfect measure of teacher skills."[5]

The educational production function estimates published by Ferguson in his 1991 paper were for the 1985–86 school year. By using TEAMS data for that year, Ferguson was able to take advantage of a feature of reform legislation implemented two years earlier that required all Texas public school teachers to pass the Texas Examination of Current Administrators and Teachers (TECAT), which the Texas Education Agency began using in 1986 to recertify existing teachers. Since nearly 97 percent of those who took the exam in March 1986 passed, TECAT was obviously not a very difficult hurdle. Furthermore, those who failed were allowed to retake the exam as many times as they wished and nearly all passed eventually. Given TECAT's low level of difficulty, the small variance of district averages, and the fact that Ferguson used average scores for all teachers in each district, it is quite surprising that TECAT scores were such a powerful predictor of student achievement in his regressions.

The Texas Education Agency was never happy with TECAT, and at the first opportunity it implemented a new and more demanding teacher certification system. With few exceptions, the Agency now requires new teachers, or teachers seeking certificates to teach in various special areas, to take one or more ExCET (Examination for the Certification of Educators in Texas) exams. In addition, all persons applying to teacher preparation programs in Texas are required to take and pass TASP, a general skills test, before they are admitted to these programs. TASP replaced an earlier skills test, PPST (Pre-Professional Skills Test) that served a similar purpose.

While the Texas Education Agency's decision to replace TECAT with a more complex system may have been good policy, it greatly complicated our research task. When we first began this research, we thought we might be able to use the TECAT data Ferguson had obtained from National Computer System (NCS). We found, however, that the individual teachers on Ferguson's tape could not be assigned to individual campuses and that only 59 percent of the 206,780 individuals who taught in Texas schools (grades pre-K to 8) during the 1990–94 period had taken TECAT.

In order to develop comparable ability measures for the 41 percent of teachers in our teacher data base who did not take TECAT, we obtained individual scores from the Texas Education Agency for TECAT and nearly 70 other teacher certification tests. Using

[4] Ferguson notes that "Primary school teachers appear to be particularly important for establishing the reading foundation upon which students depend in later years," adding that their "passing rates on the TECAT have three times the impact of secondary school teachers' passing rates for predicting eleventh graders' passing rates on the TEAMS reading exam" (Ferguson 1991, p. 476). Ferguson used *district-wide passing rates* rather than *average scores* for this part of the analysis because the mean district TECAT scores he used in his analysis were for elementary and secondary school teachers combined. His analyses of TECAT's impact on student achievement for primary and secondary school teachers separately relied on district-wide TECAT *passing rates*.

[5] Summers and Wolfe (1977, p. 644–45) in their careful study of 627 6th-grade students attending 103 Philadelphia elementary schools found a "perverse (negative) relationship between the National Teacher Exam score and learning," but also found that "Teachers who received BAs from higher-rated colleges were associated with students whose learning rate was greater."

Table 7
Linear Specifications of Fixed-Effects Regressions of Campus Mean Teacher TECAT Reading Scores

	All Districts		All MSAs		Large MSAs		Greater Dallas		Houston		San Antonio	
Variable	Coefficient	t-stat.	Coefficient	t-stat.	Coefficient	t-stat.	Coefficient	t-stat.	Coefficient	t-stat.	Coefficient	t-stat.
Campus Percent[a]												
Black	−2.44	−15.6	−2.37	−14.6	−2.54	−13.5	−2.03	−8.2	−3.43	−7.9	−2.17	−2.7
Hispanic	−1.41	−9.4	−1.30	−8.3	−1.30	−7.1	−1.04	−3.7	−1.80	−4.5	−.91	−1.3
High-Income	−.13	−.9	.00	.0	.00	.0	.18	.8	−.18	−.5	.37	.6
Black*High-Income	1.15	4.2	1.12	4.1	1.38	4.4	.77	1.8	2.00	3.3	4.58	2.2
Hispanic*High-Income	1.30	6.4	1.29	6.2	1.14	4.6	2.16	4.6	1.05	1.8	.00	.0
District Fixed Effect (F-stat.)		3.4		5.5		5.9		4.5		8.3		1.0
Constant	52.68	458.3	52.56	410.4	52.63	348.5	52.53	276.6	52.80	148.8	52.42	87.3
R^2	.73		.73		.71		.74		.77		.49	
No. Observations	4,839		3,449		2,354		763		706		290	
No. Districts	1,046		382		203		72		41		19	
Mean	52.0		51.9		52.0		52.3		51.7		51.9	
Standard Deviation	1.1		1.1		1.1		.9		1.4		.9	

[a]Percent of students enrolled in grades 3 to 7.

data for seven subsamples of teachers who took both TECAT and one or more of the other tests, we estimated explanatory models that explained TECAT reading and writing scores.[6] These equations, which we used to predict TECAT reading and writing scores for all teachers who had not taken TECAT, included as explanatory variables teacher race, sex, scores on the "other" exam, and several dummy variables that served as controls for different test administration dates.

Table 7 contains six fixed-effects regressions of mean TECAT reading scores of teachers of grades 3 to 7 on several explanatory variables that describe the campus percentages of students enrolled in grades 3 to 7 who are African-American, Hispanic, or from high-income families. These equations are for six geographic areas: the entire state, all metropolitan areas, large metropolitan areas, and, individually, the Greater Dallas, Houston, and San Antonio metropolitan areas.[7] The means and standard deviations of all six dependent variables and six independent variables used in the analyses of within-district, campus variations in school inputs are shown in Table 8 for the six areas. Since sample membership differs slightly depending on which dependent variable is being considered, the means of the independent variables also differ slightly by the dependent variable being used. The means and standard deviations of the five explanatory variables in Table 8 are for the sample used in estimating the TECAT reading score regressions.

Not surprisingly, the 1,046 district dummy variables have a large impact on the explanatory power of the TECAT reading equations shown in Table 7. The R^2 for the Large MSA equation, for example, increased from 0.55 to 0.71 when the district dummies were added to the equation. The inclusion of dummy variables in these fixed-effect equations holds constant the influence of school district policies and other factors that produce differences in mean TECAT scores among districts. Campus-level variables in turn quantify the average effect of differences in race/ethnicity and family income on TECAT reading scores and

[6] Of the 217,481 persons teaching in 1994 at the campuses included in this analysis, 79,425 had not taken TECAT. All but 166 of them, however, had taken at least one of the seven other teacher certification tests with sufficiently large samples of persons taking both it and TECAT to permit estimation of TECAT prediction equations. These tests and the number of TECAT reading and writing scores that were predicted with each equation are: Excet2 (Professional Development—Elementary) 46,691; Excet3 (Professional Development—Secondary) 17,579; Excet1 (Professional Development—All Level) 7,073; TASP 3,753; Excet4 (Elementary Comprehensive) 3,753; PPST 648; and TOPT81 387. Mean TECAT scores by race/ethnicity and highest degree earned were used to predict TECAT scores for the 166 persons who took none of the above tests.

[7] We also prepared estimates of the six equations in Tables 9 to 13 with squared terms for percent black, percent Hispanic, and percent high-income. The squared terms were added to test for nonlinearities in the campus race/ethnicity composition and income variables. Interested readers may obtain copies of these tables from the authors.

Table 8
Means and Standard Deviations of Dependent and Independent Variables

	All Districts	All MSAs	Large MSAs	Greater Dallas	Houston	San Antonio
Dependent Variables						
TECAT Reading Score	52.0	51.9	52.0	52.3	51.7	51.9
	(1.1)	(1.1)	(1.1)	(.9)	(1.4)	(.9)
TECAT Writing Score	27.0	27.0	27.0	27.3	26.7	26.9
	(.9)	(.9)	(1.0)	(.8)	(1.2)	(.8)
% Advanced Degrees	24.5	25.6	27.8	32.4	27.5	27.4
	(15.1)	(14.8)	(14.8)	(15.4)	(13.2)	(15.3)
% 0–3 Years Experience	24.8	25.0	24.9	24.3	26.5	22.5
	(14.0)	(8.2)	(13.4)	(13.4)	(13.4)	(12.5)
% 20+ Years Experience	15.9	15.6	15.8	16.5	14.5	17.5
	(10.4)	(9.8)	(9.9)	(10.1)	(9.0)	(10.5)
Class Size		19.1	19.5			
		(4.2)	(4.1)			
Independent Variables						
Campus % Black	13.7	15.9	17.0	18.9	23.1	8.4
	(.2)	(.2)	(.2)	(.2)	(.3)	(.1)
Campus % Hispanic	33.3	35.3	34.4	17.8	30.8	63.4
	(.3)	(.3)	(.3)	(.2)	(.3)	(.3)
Campus % High-Income	50.0	50.4	52.0	61.5	53.9	33.9
	(.3)	(.3)	(.3)	(.3)	(.3)	(.3)
% Black*% High-Income	5.6	6.3	6.6	7.5	9.4	2.6
	(.1)	(.1)	(.1)	(.1)	(.1)	(.0)
% Hispanic*% High-Income	11.0	11.0	11.4	7.3	11.6	14.6
	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)

document the within-district allocation of teachers with different levels of verbal ability among campuses according to the fractions of students at each campus who are black, Hispanic, and from high-income families. The estimates in Table 7 document a clear sorting of teacher ability by schools of differing racial/ethnic and income composition. There is also a sorting of teachers among districts, but the nature of this sorting is not quantified in this analysis.[8]

The tendency of teachers' mean TECAT reading scores to decrease as the campus shares of black and Hispanic students increase is clearly evident in Table 7. All 12 coefficients (campus percent black and campus percent Hispanic times six areas) are negative. The coefficient for campus percentage black in the Large MSA equation, which is −2.54, is highly significant statistically and quantitatively important (as it is for all districts and all MSAs). Since the standard deviation of the mean TECAT reading score is only 1.1, this result implies that the mean TECAT reading score for a 100 percent black school would, holding the effect of all other variables constant, be roughly 2.5 standard deviations less than the same score for a school that is zero percent black. The coefficient for campus percent Hispanic is also highly significant statistically, but it is only about half as large as the coefficient for campus percent black. The coefficient for campus percent high-income is essentially zero, but the coefficients for the two interaction variables are positive and highly significant statistically. The impacts of these interaction effects will be examined further in the concluding section.

The results obtained for the TECAT writing regressions, shown in Table 9, are very similar to those obtained for the TECAT reading scores. They also indicate that persons teaching at schools with higher fractions of black and Hispanic students and fewer students from high-income families tend to have lower TECAT scores, in this case on the writing portion.

While the analyses of variations in mean TECAT reading and writing scores presented in this section are based on estimated campus means for all grades (3 to 7) in 1994, we also estimated these equations for grades 3 and 6. The two most obvious changes, relative to the results discussed above, are a significant reduction in overall explanatory power and a decrease in sample size. Using the Large MSA equation as an example, the

[8] In an earlier version of this paper, given at Harvard's Urban Economics seminar, we presented equations that included both campus and district shares of total enrollment by race/ethnicity and family income in an effort to quantify the variations of school inputs among both districts and campuses with varying percentages of black, Hispanic, and high-income students. High correlations between the district and campus-level variables defeated this effort. The specifications used in this paper finesse this problem by limiting the assessment to the effects of within-district variations in campus racial/ethnic and income composition. We would like to acknowledge the very helpful suggestions on this issue by several seminar participants.

Table 9
Linear Specifications of Fixed-Effects Regressions of Campus Mean Teacher TECAT Writing Scores

	All Districts		All MSAs		Large MSAs		Greater Dallas		Houston		San Antonio	
Variable	Coefficient	t-stat.	Coefficient	t-stat.	Coefficient	t-stat.	Coefficient	t-stat.	Coefficient	t-stat.	Coefficient	t-stat.
Campus Percent[a]												
Black	−2.30	−16.3	−2.15	−14.9	−2.20	−13.3	−2.02	−8.7	−2.33	−6.2	−2.28	−3.5
Hispanic	−1.64	−12.1	−1.46	−10.4	−1.35	−8.4	−1.25	−4.8	−1.39	−4.0	−1.81	−3.2
High-Income	−.40	−3.2	−.19	−1.4	−.06	−.4	−.01	.0	.09	.3	−.54	−1.0
Black*High-Income	.66	2.7	.67	2.7	.87	3.1	1.10	2.7	.54	1.0	2.65	1.6
Hispanic*High-Income	1.03	5.6	1.03	5.6	.77	3.6	1.47	3.4	.68	1.3	.68	1.0
District Fixed Effect (F-stat.)		2.9		4.6		5.4		3.79		7.88		1.41
Constant	27.95	268.9	27.77	243.3	27.69	208.7	27.70	156.2	27.46	89.3	28.29	57.3
R^2	.68		.69		.69		.69		.75		.51	
No. Observations	4,839		3,449		2,354		763		706		290	
No. Districts	1,046		382		203		72		41		19	
Mean	27.04		26.98		26.97		27.30		26.70		26.90	
Standard Deviation	.91		.94		.96		.82		1.18		.78	

[a]Percent of students enrolled in grades 3 to 7.

percent of variance explained is 43 percent for grade 3 and 46 percent for grade 6, as contrasted with 69 percent for all grades. The 6th grade Large MSA sample has 891 observations, the 3rd grade one has 1,612, and the all-grades sample has 2,354 observations. These differences in sample size result from the fact that many campuses do not have both 3rd and 6th grades and schools serving only the lower grades tend to be smaller (fewer students per grade).

Percent of Teachers with Advanced Degrees

In contrast to most other educational production function studies that have found no relationship between teachers' years of education and the performance of their students on standardized tests, Ferguson (1991, p. 477) in his study of Texas schools found that "*Master's degrees produce moderately higher scores in grades one through seven*" [his italics] and that "The percentage of teachers who have master's degrees accounts for about five percent of the variation in student scores across districts for grades one through seven." This finding also finds support in Ferguson and Ladd (1995, p. 35) who found that "additional education for teachers, as measured by the proportion of teachers with master's degrees, also appears to increase student learning, but by a lesser amount" (relative to teachers' test scores).

Table 10 contains fixed-effects estimates for the campus percentage of teachers with advanced degrees for the same six geographic areas used in the analysis of mean TECAT scores.[9] The sign patterns for the All Districts, All MSAs, Large MSAs, and Greater Dallas equations are identical. In all four equations, the percentage of teachers with advanced degrees declines as the campus percentages of black, Hispanic, and high-income students increases; it rises with increases in both the race/ethnicity-income interaction variables. At first glance the results for Houston and San Antonio appear to be very different, but the individual coefficient estimates have very low levels of statistical significance, a result that appears to be due to high levels of multicollinearity among a number of the explanatory variables.

In contrast to the TECAT regressions, where the campus percent black coefficient was much larger (in absolute value) than the campus percent Hispanic coefficient, these two coefficients are very similar in the first four equations (those with statistically signif-

[9] Including district dummies has a much larger impact on the overall explanatory power of these equations than was true for the TECAT reading or writing regressions. For the six equations without district dummies, the highest fraction of explained variance is only 16 percent (the San Antonio equation). Adding the district dummy variables to the San Antonio equation increases the explained variance to 26 percent. The largest increase in overall explanatory power from adding the district dummies is obtained for the All-District equation; the fixed-effects equation boosts the total explained variance from 6 percent to 50 percent.

Table 10
Linear Specifications of Fixed-Effects Regressions of Campus Fraction of Teachers with Advanced Degrees

	All Districts		All MSAs		Large MSAs		Greater Dallas		Houston		San Antonio	
Variable	Coefficient	t-stat.	Coefficient	t-stat.	Coefficient	t-stat.	Coefficient	t-stat.	Coefficient	t-stat.	Coefficient	t-stat.
Campus Percent[a]												
Black	−.14	−5.2	−.12	−4.4	−.12	−3.5	−.16	−3.0	.13	2.1	−.09	−.7
Hispanic	−.15	−5.9	−.13	−5.1	−.12	−3.8	−.17	−2.8	−.03	−.5	.13	1.1
High-Income	−.13	−5.7	−.11	−4.5	−.12	−3.9	−.10	−2.2	.04	.7	−.01	−.1
Black*High-Income	.20	4.5	.19	4.3	.19	3.6	.28	2.9	−.06	−.7	−.21	−.9
Hispanic*High-Income	.04	1.2	.01	.3	−.02	−.5	.17	1.6	.01	.1	−.22	−1.7
District Fixed Effect (F-stat.)		3.4		5.7		5.6		7.3		4.1		4.2
Constant	.36	19.6	.36	17.4	.39	15.3	.41	10.4	.24	5.0	.25	2.4
R^2	.50		.44		.35		.44		.26		.26	
No. Observations	5,015		3,592		2,464		791		720		318	
No. Districts	1,046		382		203		72		41		19	
Mean	.25		.26		.28		.32		.28		.27	
Standard Deviation	.15		.15		.15		.15		.13		.15	

[a]Percent of students enrolled in grades 3 to 7.

icant coefficients). Somewhat more surprising, the campus percent of teachers with advanced degrees declines as the fraction of high-income students increases, in all four equations. At the same time, the coefficient of the percent black and percent high-income interaction term is positive, indicating that increases in the high-income share reduce the negative impact of higher fractions of black students on the percentage of teachers with advanced degrees. Of the three individual metropolitan area equations, only the Greater Dallas estimates closely conform to the estimates obtained for the more comprehensive samples; the overall explanatory power of this equation is also considerably larger than those of the other two metropolitan areas. The within-district sorting of teachers with advanced degrees, moreover, appears to be more pronounced in Greater Dallas than in the other two individual metropolitan areas or in either All MSAs or Large MSAs.

Overall, the results in Table 10 provide strong evidence that, in Texas at least, teachers employed in schools with large fractions of Hispanic and African-American children, and particularly the latter, and also in schools with high fractions of children from low-income families, have fewer years of education. This evidence relating to systematic within-district variations in teachers' years of schooling by campus racial and income composition may not matter much, if Hanushek and others are right about the unimportance of years of teacher education as a determinant of student achievement. However, if Ferguson (1991) and Ferguson and Ladd (1995) are correct in their opposing view, this evidence could be quite important.

Teacher Experience

Ferguson (1991, p. 475–76) found that after teachers' language skill, as measured by the TECAT score, "the next most important school characteristic is teacher experience," and that teacher "experience accounts for a bit more than ten percent of the inter-district variation in student test scores." Even Hanushek relents a bit when it comes to teacher experience. Referring to his 1989 survey of educational production function estimates, Hanushek (1989, p. 47) observes that "Teacher experience is possibly different," since "At least a clear majority of estimated coefficients point in the expected direction, and almost 30% of the estimated coefficients are statistically significant by conventional standards." He then returns to form by adding the following qualifications:

> But these results are hardly overwhelming; they appear strong only relative to the other school inputs. Moreover, because of possible selection effects, they are subject to additional interpretive questions. In particular, these positive correlations may result from senior teachers being permitted to select schools and classrooms with better students. In other words, causation may run

from achievement to experience and not the other way around.

A few educational production function studies have suggested that an inverted U-shaped relationship may be present between teacher experience and student achievement. In this regard, the evidence that "inexperienced" teachers are less effective and that students taught by them do less well on standardized tests is more extensive than the evidence suggesting that teacher performance deteriorates beyond a certain point. Nonetheless, some basis exists for believing that there may be too much of a good thing when it comes to teacher experience, or perhaps age, and that the teachers with the most experience, generally older ones, are less effective than younger teachers with somewhat less experience. Some support for this proposition is provided by exploratory educational production functions for Texas elementary schools completed by Kain (1995).

Some basis exists for believing that there may be too much of a good thing when it comes to teacher experience, or perhaps age, and that the teachers with the most experience, generally older ones, are less effective than younger teachers with somewhat less experience.

Because the relationship between teacher experience and student achievement may be an inverted U shape, we have used two dependent variables to quantify experience, the campus proportions of teachers with zero to three years of experience and those with 20 or more. Like the analyses of TECAT scores and advanced degrees, the dependent variables in these equations are for the entire campus, although, as was true of the TECAT analysis, we estimated experience equations for 3rd and 6th grade teachers as well. The results for the 3rd and 6th grade are very similar to those for the entire campus except that the equations generally explained a smaller fraction of the variance in the several dependent variables.

The teacher experience regressions, shown in Tables 11 and 12, include the same explanatory variables and have the same structure as those described above for the regressions for TECAT scores and for percent of teachers with higher degrees in Tables 7, 9, and 10. The fractions of inexperienced and very experienced teachers in particular districts presumably are strongly affected by district demographics. Rapidly growing districts are likely to have proportionately more inexperienced teachers, although this tendency may be offset by policies that favor the recruitment and hiring of experienced teachers. Districts with declining enrollments similarly are likely to have large numbers of very experienced teachers and these district-wide tendencies will be felt at the campus level. The district dummy variables included in the fixed-effects equations account for district-to-district differences of this kind.[10]

Three of the five coefficients (excluding the constant term) for the Large MSA fixed-effects equation for inexperienced teachers have t-statistics of 2.8 or greater. The t-statistics for the remaining two are −1.4 for the coefficient for the percent black-percent high-income interaction and a minuscule 0.3 for the campus percent of high-income students. The coefficients for percent black and percent Hispanic are very similar in magnitude and indicate that increases in either are associated with a higher fraction of inexperienced teachers. The coefficient for the campus percent high-income is zero, suggesting that campus income has little effect on the mix of teachers.

The regression equations for the campus percentage of inexperienced teachers (those with 0–3 years of teaching experience) provide strong evidence that schools with higher percentages of black or Hispanic students have disproportionate numbers of inexperienced teachers. According to the estimates in Table 11, differences in the campus percentage of high-income students affect the allocation of inexperienced teachers only through an interaction with either campus percent black or campus percent Hispanic. In addition, these interaction effects are opposite in sign. Holding the campus percent black constant, increases in the campus percentage of high-income students reduces

[10] The large increase in overall explained variance with the addition of the district dummies suggests these district-level effects have major impacts on the campus fractions of inexperienced and very experienced teachers. When district dummies are not included in the six equations shown in Table 11, the fraction of explained variance varies from a low of 3 to a high of 14 percent; when district dummies are included, the R^2s of these equations vary from 28 to 40 percent.

Table 11
Linear Specifications of Fixed-Effects Regressions of Campus Fraction of Teachers with Zero to Three Years of Experience

	All Districts		All MSAs		Large MSAs		Greater Dallas		Houston		San Antonio	
Variable	Coefficient	t-stat.	Coefficient	t-stat.	Coefficient	t-stat.	Coefficient	t-stat.	Coefficient	t-stat.	Coefficient	t-stat.
Campus Percent[a]												
Black	.20	6.8	.20	6.4	.19	5.4	.27	4.8	.18	2.5	−.28	−2.3
Hispanic	.21	7.4	.20	6.7	.20	5.9	.24	3.7	.32	4.9	−.32	−3.0
High-Income	.02	.7	.01	.3	.01	.3	.05	1.0	.05	.8	−.42	−4.2
Black*High-Income	−.06	−1.1	−.11	−2.0	−.08	−1.4	−.21	−2.1	.00	.0	.65	2.1
Hispanic*High-Income	.05	1.3	.07	1.6	.13	2.8	.37	3.5	−.11	−1.1	.48	4.0
District Fixed Effect (F-stat.)		2.3		3.1		3.8		3.6		3.6		5.2
Constant	.14	6.3	.14	5.7	.13	4.7	.11	2.4	.11	1.9	.51	5.5
R^2	.40		.30		.28		.30		.30		.33	
No. Observations	4,842		3,451		2,355		761		706		291	
No. Districts	1,046		382		203		72		41		19	
Mean	.25		.25		.25		.24		.27		.23	
Standard Deviation	.14		.08		.13		.13		.13		.13	

[a]Percent of students enrolled in grades 3 to 7.

the fraction of inexperienced teachers. In the case of Hispanic students, the opposite result is seen.

The within-district allocation of inexperienced teachers among schools with varying racial, ethnic, and income composition documented by the inexperienced teachers equations is no doubt due to the well-known "MFK" effect.[11] This effect refers to school district policies that permit teachers with more seniority to choose more desirable schools and the absence of any incentives or rewards for teaching in less desirable schools. The campus percentages of black, Hispanic, and high-income students are proxies for a larger number of school characteristics that determine the attractiveness of individual campuses to teachers. Since we know the teaching assignments of all teachers during the five-year period 1990–94, we should in the future be able to explicitly model this process.

The coefficient estimates obtained for the Greater Dallas and Houston inexperienced teachers equations are fairly similar to those obtained for All Districts, All MSAs, and Large MSAs. The results for San Antonio could not be more different. All five coefficients have t-statistics of more than 2.0 and the sign pattern, which is generally the opposite of that obtained for the other fixed-effects equations, indicates that the percentage of inexperienced teachers declines as the percentages of black, Hispanic, and high-income students increase and rises with increases in both black and Hispanic high-income interaction variables.

The results for teachers with 20 or more years of experience, shown in Table 12, exhibit a pattern consistent with the sorting mechanism described above for inexperienced teachers. Teachers with the most tenure tend to be underrepresented in campuses with

[11] This reference to the MFK effect refers to the experience of the senior author's wife, Mary Fan Kain, during the first year of their marriage when she took a job teaching in a overwhelmingly black junior high school in Oakland, California. Mary Fan did not come to teach in this school because of a commitment to teach disadvantaged children and she had no special preparation (she did her practice teaching in a small rural school in Ohio near Denison University, where she was a student). She took the Hoover Junior High School job because when she arrived in late August in Berkeley, where John was to attend graduate school, only two jobs were left in the East Bay in her areas of specialization (junior high school social studies and English). Both were in inner-city, overwhelmingly black schools, and she took the job closest to Berkeley. Hoover Junior High School was not a bad school by today's standards, and she found the kids were, for the most part "good kids." However, she was totally unequipped to deal with the problems she encountered, which included 7th graders who couldn't read and high rates of turnover. There were three kinds of teachers at Hoover Junior High School. About a third were completely unprepared first-year teachers like Mary Fan who came to teach at Hoover Junior High for the same reason she did—jobs in inner-city schools were the only ones available to beginning teachers. Another third were somewhat older, but still fairly young, dedicated black teachers, and the last third were older white teachers, who, with few exceptions, had lost interest in teaching and prided themselves on maintaining order and discipline and quiet classrooms. It is unclear whether any more learning went on in Mary Fan's classroom than in other classrooms, but we do know that her students had more fun and that her classroom was much less quiet.

Table 12
Linear Specifications of Fixed-Effects Regressions of Campus Fraction of Teachers with 20 or More Years of Experience

	All Districts		All MSAs		Large MSAs		Greater Dallas		Houston		San Antonio	
Variable	Coefficient	t-stat.	Coefficient	t-stat.	Coefficient	t-stat.	Coefficient	t-stat.	Coefficient	t-stat.	Coefficient	t-stat.
Campus Percent[a]												
Black	−.10	−4.8	−.11	−5.0	−.10	−4.1	−.18	−4.3	−.02	−.4	−.15	−1.4
Hispanic	−.14	−6.9	−.14	−6.7	−.15	−6.3	−.24	−5.0	−.17	−3.7	−.13	−1.3
High-Income	−.06	−3.2	−.06	−3.2	−.07	−2.9	−.09	−2.5	−.08	−1.8	−.09	−1.1
Black*High-Income	.11	2.9	.13	3.4	.09	2.1	.21	2.9	−.08	−1.1	−.17	−.6
Hispanic*High-Income	.07	2.3	.06	2.1	.05	1.4	.12	1.5	.15	2.2	.10	.9
District Fixed Effect (F-stat.)		2.5		3.9		4.3		3.5		4.1		3.7
Constant	.24	14.9	.24	14.0	.25	12.4	.27	8.4	.23	5.8	.29	3.5
R^2	.42		.35		.32		.33		.28		.26	
No. Observations	4,842		3,451		2,355		763		706		291	
No. Districts	1,046		382		203		72		41		19	
Mean	.16		.16		.16		.17		.15		.18	
Standard Deviation	.10		.10		.10		.10		.09		.11	

[a]Percent of students enrolled in grades 3 to 7.

large fractions of minorities and overrepresented in campuses with higher fractions of students from more well-to-do families. If the previously mentioned inverted U-shaped function between teacher experience and student achievement exists, the finding that teachers with more than 20 years are underrepresented among the faculties of campuses with large high percentages of disadvantaged minorities would work to these students' advantage. In contrast to the results for the inexperienced teachers equations, the San Antonio sign pattern is generally the same as the pattern obtained for the other five equations. None of the five coefficients in the San Antonio equation are significantly different from zero, however.

Class Size/Student Teacher Ratios

Ferguson (1991, p. 477) found that class size had a measurable impact on reading scores and placed great emphasis on the role of thresholds, indicating that "reducing the number of 'students per teacher' is important only when it exceeds eighteen. . . . Each additional student over eighteen causes the district average score to fall by between one-tenth and one-fifth of a standard deviation in the inter-district distribution of test scores for grades one through seven." Once again, Hanushek's survey articles offer little support for the notion that smaller classes have a significant effect on student achievement. Of the 187 educational production function studies that Hanushek reviewed for his 1989 survey article, 152 included the teacher-student ratio as an explanatory variable. Only 14 of these 152 studies obtained positive and statistically significant coefficients for the teacher-student ratio and nearly as many (13) found a negative and statistically significant relationship with student achievement (Hanushek 1989, p. 47).

The precision and certainty of Ferguson's (1991) conclusions about the effects of class size and threshold effects, particularly given the crudeness of his data, are stunning. Moreover, his findings about the effects of class size on achievement for Texas schools are further supported by his and Ladd's recent study of Alabama schools. In discussing their results, Ferguson and Ladd (1995, p. 35) observe that "the basic conclusion that class size matters for student learning emerges clearly and consistently, especially for math."[12]

Because Ferguson's (1991, p. 472) Texas study

[12] Summers and Wolfe (1977, p. 645) provide some support for the view that students in smaller classes have larger achievement gains, indicating that their analyses provide fairly strong evidence that smaller classes tended to increase the achievement gains for both low-achieving and high-achieving students, but had no effect on average students. In addition, they briefly describe the results of a survey of 85 earlier studies of the effect of class size on achievement, contained in an unpublished Ph. D. dissertation by Blake (1954), who found that 35 studies determined smaller classes were more effective, 18 determined that larger classes were more effective, and 32 were inconclusive.

relied on aggregate (district-level) data, he had to use students-per-teacher for the entire district as his measure of class size "because a direct measure of average class size is unavailable." Elaborating on this point, he observes:

> This study (and most others) lacks a direct measure of average class size. It does, however, have a measure of the number of students per teacher in the district. Average class size will be larger than "students per teacher" because some teachers are specialists who do not teach regular classes and because most teachers get periods off during the day. The results here show that reducing the number of "students per teacher" is important only when it exceeds eighteen. (Tests show that at least in this data set, the threshold is indeed at eighteen and not at seventeen or nineteen.) . . . This is among the stronger effects for any variable in the study. However, it is an effect that is clearly restricted to the primary grades.

In contrast to Ferguson's and most other studies, we have been able to use data for individual teachers to construct a "direct measure of average class size" for each campus and grade. The class size data come from what we have termed "teacher time cards." These time cards provide a detailed description of each teacher's workweek by grade and subject taught, days of the week, the fraction of total time spent on each class, and the number and type (regular, special education, gifted and talented, and so on) of students who are enrolled in each of their classes.

Analysis of the teacher time cards revealed that most elementary school teachers offer instruction in only one grade and to one student population (Regular, ESL, Gifted and Talented, Compensatory/Remedial, Bilingual, or Special Education) and that most persons teaching in kindergarten through 5th grade have only one teaching assignment, presumably a stand-alone classroom.[13] This analysis further indicates that the mean number of assignments (time cards) increases from 1.3 cards per teacher in kindergarten to 1.9 cards per teacher in the 5th grade. Starting in 6th grade, the instructional technology clearly changes, as shown by a sharp increase in the number of teaching assignments from 1.9 per teacher in the 5th grade to 4.0 in the 6th; the mean numbers are even higher for those teaching regular 7th- and 8th-grade students (4.8 and 4.6 per teacher). Special-education teachers, who average 6.5 to 9.8 time cards, have the most assignments.

Texas' school reform legislation also required that all public elementary schools have 22 or fewer students per classroom through the 4th grade. Districts were given four years to fully implement the rule. The normalized frequency distributions of mean class size by campus for all grades (3 to 7) and for the 3rd and 6th grades in 1994, shown in Figure 1, make it clear this regulation has had a significant impact on class sizes.[14]

Because mean classroom size varies substantially by grade, we present equations for all grades (3 to 7), grade 3, and grade 6 for all MSAs and for large MSAs in Table 13. In contrast to the results obtained for other school inputs, the class size all-grades equation does not have a higher R^2 than either the 3rd- or 6th-grade equations. Instead, no doubt reflecting the 22-student cap, the 3rd-grade equation has much less variance and a higher R^2 than the all-grades equation. The R^2 for the 6th-grade equation is also larger than that for the all-grades equation.[15]

Only two of the five coefficients in the all-grades, all MSAs regressions are significantly different from zero. The coefficients for campus percent high-income indicate that average classroom size increases with percent high-income, while the sign for the percent black and high-income interaction variable indicates that average classroom size declines as both the campus percent black and the campus percent high-income increase.

The 3rd-grade class size regression has only one coefficient that is significantly different from zero at the 5 percent level, a negative coefficient for the campus percent black, suggesting that average class size decreases as the percent black increases. In the 6th-grade regression, average class size increases as

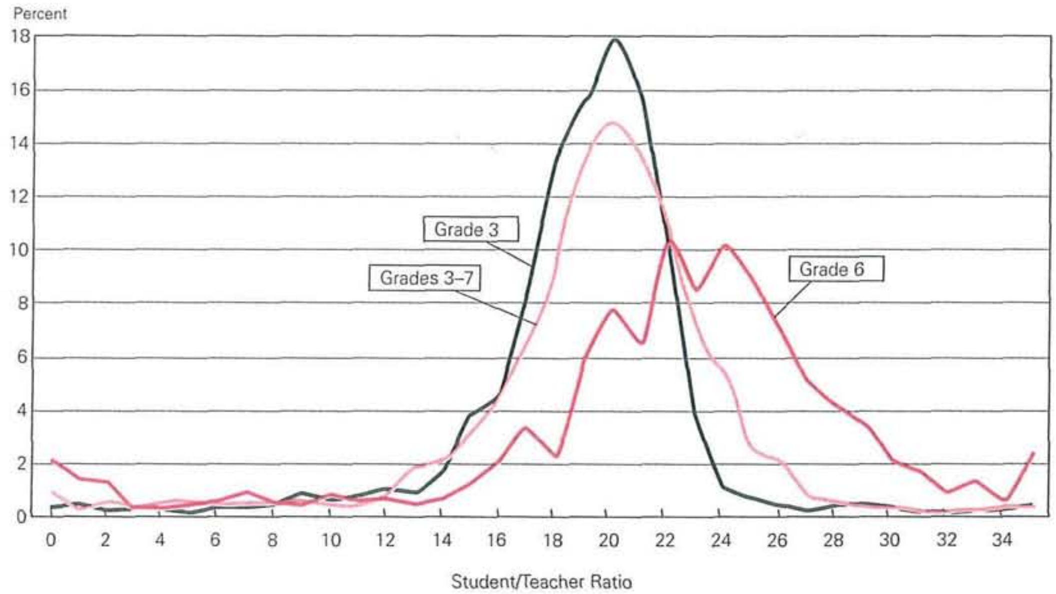
[13] This analysis is based on data for 139,565 classroom teachers in 1994, excluding only those teaching physical education or fine arts. These teachers reported a total of 389,491 different teaching assignments for an average of 2.8 assignments (cards) per teacher.

[14] There are also large differences in class size by population served. Mean class sizes by population served in the 3rd grade are: Bilingual (16.8 students per class), Compensatory/Remedial (13.1 students per class), ESL (16.6 students per class), Gifted and Talented (17.3 students per class), Regular (19.3 students per class), and Special Education (3.8 students per class). If students remained in the same classroom for all of their classes and had the same number of classmates for every class, the use of these data would be straightforward. We know, however, that students can be enrolled in more than one of these programs and it is possible that some, or even most, of these special classes are "pull-outs" from regular classrooms. In those cases, students enrolled in these programs may have taken the larger part of their course work in regular classrooms. Worse yet, these practices presumably differ from one district (or campus) to another.

[15] Using the Large MSA regressions as an example, the fraction of explained variance varies from 2 to 4 percent for the six equations without district dummies. When district dummies are added to the equations, the level increases to between 29 and 47 percent of the total variance.

Figure 1

Distribution of Mean Class Size by Grade, 1994



the campus percent black increases and as the campus percentage of high-income students increases, and decreases as the product of campus percent black and high-income increases.

Predicted Variations in School Inputs by Campus Race/Ethnicity and Income

To make it somewhat easier to understand the way in which each of the six school inputs examined in this paper is affected by differences in campus racial, ethnic, and income composition, we have used the Greater Dallas equations to carry out a parametric analysis of the way in which the several school inputs vary with representative levels of these campus-level variables. We use the Greater Dallas equation rather than the All District or the All MSAs because a single metropolitan area provides a more meaningful indication of the residential and public schooling choices available to households.

The predictions in Table 14 show how the estimated levels of each school input vary with specified changes in the percentages of black and Hispanic and high-income students for hypothetical Greater Dallas campuses.[16] These estimates are obtained using the input coefficients shown in previous tables, an arbitrary value for campus percent Hispanic of 10 percent for all cases, and quintile means of campus percentages black (column 3), and percentages high-income (shown in parentheses at the top of the table).[17]

While the analyses use the actual mean percentages of black and high-income students for each quintile, the quintiles themselves are defined by equal intervals of percent black and percent high-income from zero to one hundred. As column 2 (number of campuses) reveals, Greater Dallas has many more campuses in the interval zero to 20 percent black than in any of the remaining quintiles. Indeed, nearly three-fourths (73.8 percent) of all Greater Dallas campuses belong to this interval; the fourth quintile,

[16] An analogous analysis of the way in which the levels of these school inputs vary with changes in campus percent Hispanic and campus percent high-income is available from the authors.

[17] The actual mean Hispanic shares vary from a low of 6.5 percent for the fifth quintile to a high of 25.8 percent (column four in Table 14). We considered using the actual mean percent Hispanic for each quintile for these simulations, rather than a constant value of 10 percent Hispanic, but ultimately decided that using actual percent Hispanic confuses the respective contributions of campus percent black and campus percent Hispanic.

Table 13
Linear Specifications of Fixed-Effects Regressions of Campus Mean Class Size, All Grades, Grade 3, and Grade 6

	All MSAs						Large MSAs					
	All Grades (3 to 7)		Grade 3		Grade 6		All Grades (3 to 7)		Grade 3		Grade 6	
Variable	Coefficient	t-stat.	Coefficient	t-stat.	Coefficient	t-stat.	Coefficient	t-stat.	Coefficient	t-stat.	Coefficient	t-stat.
Campus Percent[a]												
Black	−.69	−.7	−1.78	−2.0	5.63	2.3	.77	.7	−1.59	−1.7	5.99	2.1
Hispanic	−1.79	−2.0	−1.70	−1.9	3.07	1.3	−.58	−.6	−1.16	−1.2	3.27	1.2
High-Income	2.59	3.1	.26	.3	7.90	3.4	3.95	4.0	1.07	1.2	9.34	3.5
Black*High-Income	−2.89	−1.9	2.21	1.4	−10.65	−2.7	−5.73	−3.2	2.01	1.2	−12.70	−3.0
Hispanic*High-Income	2.17	1.8	1.12	1.0	−2.78	−.9	.42	.3	−.41	−.3	−4.92	−1.3
District Fixed Effect (F-stat.)		3.9		3.7		2.1		3.9		4.1		2.4
Constant	18.48	25.0	19.25	26.6	16.54	8.2	17.88	20.5	19.13	24.1	16.58	7.1
R^2	.36		.43		.47		.29		.38		.43	
No. Observations	3,449		2,350		1,344		2,353		1,621		908	
No. Districts	382		374		379		203		202		203	
Mean	19.10		18.73		21.54		19.53		19.09		22.16	
Standard Deviation	4.15		3.40		6.55		4.08		3.09		6.65	

[a]Percent of students enrolled in grades 3 to 7.

which is 61 to 80 percent black, by contrast includes only 1.9 percent of all campuses and the quintile 81 to 100 percent black includes only 6.6 percent of all Greater Dallas campuses. The bottom two rows of Table 14, which give the number and percentage of campuses in each income quintile, reveal that Greater Dallas campuses tend to be concentrated in the upper end of the income distribution. The 81 to 100 percent high-income quintile contains nearly a third of all Greater Dallas campuses, while the zero to 20 percent category contains only 13 percent.

The hypothetical calculations by campus percentage black and by campus percentage high-income shown in Table 14 raise the question of how well these categories represent the actual distribution of elementary schools in Greater Dallas by racial, ethnic, and income composition. As just noted, most (73.8 percent) Greater Dallas campuses are less than 20 percent black. At same time, the 26 percent of campuses that are more than 20 percent black served 69 percent of Greater Dallas African-American students enrolled in grades 3 to 7 in 1994. It is also the case that very few (five) campuses are more than 50 percent black *and* more than 50 percent high-income. While only 10.6 percent of Greater Dallas campuses are more than 50 percent black, 94 percent of these campuses have low-income percentages in excess of 50 percent. While campuses with very high percentages of black students are disproportionately concentrated in the cells defined by low shares of high-income students, the distribution of campuses by racial composition alone (percent black) is surprisingly uniform. In particular, only 49 (6.2 percent) of the 791 elementary schools in Greater Dallas in 1994 had no African-American students enrolled in grades 3 to 7 in 1994.

The predicted values of TECAT teacher reading and writing scores in Table 14 decline as the campus percentage of black students rises, and they increase as the campus percentage of high-income students increases. To give an example, for campuses with only 9.5 percent high-income students, the predicted difference in TECAT reading scores is 1.6 points between campuses that are 6.9 percent black and 92.2 percent black. Reading the table the other way, for campuses that are 6.9 percent black, predicted TECAT reading scores are 0.4 points less for campuses with only 9.5 percent high-income students than for campuses that are 91.2 percent high-income. Similarly, for campuses that are mostly black (92.2 percent black), the predicted teacher TECAT reading scores are 0.9 points higher for those in the highest income category (91.2 percent high-income) than for those in the lowest (9.5 percent high-income). No Greater Dallas campuses are 90 to 100 percent black *and* 90 to 100 percent

Table 14

Estimated Relationship of School Inputs to Percentages of Black and Hispanic Students and Percentages of High-Income Students, on Greater Dallas Campuses[a]

		Actual Mean			Campus % High-Income				
Campus % Black	Number of Campuses	% Black	% Hispanic	Assumed % Hispanic	0–20% (9.5%)	21–40% (29.9%)	41–60% (50.9%)	61–80% (70.9%)	81–100% (91.2%)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					Predicted Teachers' TECAT Reading Score				
0–20	584	6.9	16.9	10.0	52.3	52.4	52.5	52.6	52.7
21–40	101	28.4	22.8	10.0	51.9	52.0	52.2	52.3	52.4
41–60	39	49.0	25.8	10.0	51.5	51.7	51.8	52.0	52.1
61–80	15	70.0	22.1	10.0	51.1	51.3	51.5	51.7	51.9
81–100	52	92.2	6.5	10.0	50.7	50.9	51.1	51.3	51.6
					Predicted Teachers' TECAT Writing Score				
0–20	584	6.9	16.9	10.0	27.5	27.5	27.5	27.6	27.6
21–40	101	28.4	22.8	10.0	27.0	27.1	27.2	27.3	27.4
41–60	39	49.0	25.8	10.0	26.6	26.8	26.9	27.1	27.2
61–80	15	70.0	22.1	10.0	26.2	26.4	26.6	26.8	27.0
81–100	52	92.2	6.5	10.0	25.8	26.1	26.3	26.5	26.8
					Predicted Class Size				
0–20	584	6.9	16.9	10.0	18.5	19.2	19.9	20.5	21.2
21–40	101	28.4	22.8	10.0	18.8	19.1	19.5	19.8	20.1
41–60	39	49.0	25.8	10.0	19.1	19.1	19.1	19.1	19.1
60–80	15	70.0	22.1	10.0	19.3	19.0	18.6	18.3	18.0
81–100	52	92.2	6.5	10.0	19.6	18.9	18.2	17.5	16.9
					Predicted % Teachers with Advanced Degree				
0–20	584	6.9	16.9	10.0	37.7	36.4	35.1	33.9	32.6
21–40	101	28.4	22.8	10.0	34.9	34.8	34.7	34.7	34.6
41–60	39	49.0	25.8	10.0	32.1	33.2	34.4	35.4	36.5
61–80	15	70.0	22.1	10.0	29.3	31.6	34.0	36.2	38.5
81–100	52	92.2	6.5	10.0	26.4	29.9	33.6	37.0	40.6
					Predicted % Teachers with 0 to 3 Years' Experience				
0–20	584	6.9	16.9	10.0	15.7	17.2	18.7	20.2	21.6
21–40	101	28.4	22.8	10.0	21.1	21.7	22.3	22.9	23.5
41–60	39	49.0	25.8	10.0	26.4	26.1	25.8	25.6	25.3
61–80	15	70.0	22.1	10.0	31.7	30.6	29.4	28.3	27.1
81–100	52	92.2	6.5	10.0	37.4	35.3	33.2	31.1	29.0
					Predicted % Teachers with 20 or More Years' Experience				
0–20	584	6.9	16.9	10.0	22.9	21.6	20.3	19.0	17.7
21–40	101	28.4	22.8	10.0	19.5	19.1	18.7	18.4	18.0
41–60	39	49.0	25.8	10.0	16.1	16.7	17.3	17.8	18.4
61–80	15	70.0	22.1	10.0	12.8	14.3	15.8	17.2	18.7
81–100	52	92.2	6.5	10.0	9.2	11.7	14.2	16.6	19.0
Number of Campuses in Income Quintile					103	85	130	213	260
Percent of Campuses in Income Quintile					13.0%	10.7%	16.4%	26.9%	32.9%

[a]Percent of students enrolled in grades 3 to 7.

high-income. The largest predicted difference in teacher reading scores, 2.0 points, is between a campus that is 92.2 percent black and 9.5 percent high-income and a campus that is 6.9 percent black and 91.2 percent high-income. The Greater Dallas area has a fair number of both of these types of campuses. In assessing the predicted TECAT scores and subsequent predictions of school inputs for campuses of varying percent black and high income, it should be remembered that all of the predictions assume a uniform campus 10 percent Hispanic.

The results obtained for predicted TECAT teacher writing scores are very similar to those obtained for TECAT reading scores. In contrast, the class size results are more complex. For campuses with few high-income students (9.5 percent), class size increases as campus percent black increases, from 18.5 students per teacher (campus percent black 6.9 percent) to 19.6 students per teacher (campus percent black 92.2 percent). For campuses with 91.2 percent high-income students (the top quintile), exactly the opposite result occurs; class size declines from 21.2 students per teacher when the percentage of black students is 6.9 to 18.9 students per teacher when 92.2 percent of the students are black.

The results for Greater Dallas provide strong evidence of systematic and large differences in the fraction of teachers with advanced degrees between low-income, minority and high-income, majority campuses in the same district.

The same kind of twist appears in the predicted fractions of teachers with advanced degrees. For the lowest income category, the percentage of teachers with advanced degrees declines from 37.7 percent for campuses that are 6.9 percent black to 26.4 percent for campuses that are 92.2 percent black. These results provide strong evidence of systematic and large differences in the fraction of teachers with advanced degrees between low-income, minority and high-income, majority campuses in the same district. As noted earlier, the importance of this result depends on whether teachers with advanced degrees are more effective teachers, something about which there is considerable disagreement.

The results for inexperienced (0 to 3 years' experience) teachers indicate very large differences in the fractions of inexperienced teachers between very low-income schools with relatively few black students and very low-income schools that are predominantly black. For campuses with the fewest (9.5 percent) high-income students, the fraction of teachers with limited experience increases from 15.7 percent for schools that are 6.9 percent black to 37.4 percent for schools that are 92.2 percent black. Smaller differences by campus percent black appear for schools with larger fractions of high-income students.

The results for teachers with 20 or more years of experience provide strong evidence that teachers with greater seniority avoid schools with high fractions of low-income black students. This result disappears for campuses in the top income quintile, however, where the predicted fraction of very experienced teachers in schools with 92.2 percent black students, 19 percent, is slightly higher than the percent for schools with only 6.9 percent black students, 17.7 percent.

Summary and Conclusions

Significant changes have taken place in the educational landscape since the Coleman Report was published nearly 30 years ago. As this paper demonstrates, the most obvious change in Texas has been substantial reductions in the extent of racial/ethnic segregation in the public schools. While Coleman and his colleagues (1966) found high levels of school segregation, data for Texas elementary schools presented in this paper show that in 1994 fewer than 16 percent of Anglo students attended schools that were greater than 90 percent Anglo. Additional measures of campus-level concentration for African-American, Hispanic, and Anglo students reveal relatively few campuses throughout the state where students attend schools composed solely of their own ethnic/racial group. Racial concentration continues to be higher in Texas's largest metropolitan areas. Nonetheless, the levels of racial/ethnic concentration today are much lower than the levels found 30 years ago.

In spite of significant declines in racial/ethnic concentrations, the large gaps in mean achievement identified by Coleman persist; analyses of mean reading scores for a synthetic cohort of students attending Texas elementary schools during the period 1989 (grade 1) through 1994 (grade 6) reveal that mean reading scores of African-American and Hispanic children in grade 1 are only 94 percent of the statewide average, while mean reading scores for Asian-American and Anglo 1st graders are 107 and 106 percent of the statewide average. Moreover, the reading scores for Hispanics exclude significant numbers of children who take the reading exam in Spanish or are excused from taking the exam because of limited English proficiency. Crude adjustments for differences in family income levels narrow, but do not eliminate, differences in the mean reading scores of Texas's major racial/ethnic groups.

Analyses presented in this paper also provide some support for the widely reported finding that racial/ethnic gaps in student achievement tend to increase as years of schooling increase. A final judgment on this finding should be reserved, however, until we have completed the linking of test and student records, repeated the same analyses for true cohorts, and more carefully evaluated the role of the 1st- and 3rd-grade Spanish language tests and non-test-taking by low-achieving students.

While the findings summarized above are important, this paper has been principally concerned with quantifying within-district variations in selected school inputs by campus racial/ethnic and family income composition. In contrast to Coleman et al.'s (1966) finding of no consistent differences in the quantity and quality of school inputs for predominantly majority and minority schools, the analyses presented in this paper reveal substantial within-district variations in four types of school inputs: teacher test scores, years of education, and experience, and class size (student-teacher ratios). The statistical models presented in this paper document a sorting of school inputs based on campus racial/ethnic and socioeconomic composition. In particular, the models suggest that teacher ability, as measured by verbal and written proficiency scores, decreases as the campus percentage of black and Hispanic students increases; measured teacher ability increases with the campus percentage of high-income students. Estimates of the variations in other school input measures provide strong evidence that, in Texas, teachers employed in schools with high fractions of disadvantaged minority students have fewer years of education and less experience; they also have more students in their classes.

While the findings presented in this paper are important in their own right, they also have important implications for the larger study of which this paper is a part, and particularly for the careful estimation of the determinants of educational achievement, a major goal of the larger study. In the past, educational production function studies have had only modest success in quantifying the relationship between school inputs and student achievement. As Hanushek and Kain (1971) argued a quarter of a century ago, the failure of earlier educational production function studies to obtain more consistent results may be attributable to imprecise measurement of school inputs. The results in this paper are a first step toward the goal of obtaining more reliable estimates of the relationship between school inputs and student achievement. The analyses reported in this paper of within-district variation in school input measures reveal that schools differ significantly in the level of inputs they provide and in the instructional technology that they employ. By creating a linked sample of student achievement scores combined with these and other precise measures of school inputs, we hope to determine how the variations in school input measures affect student achievement and the gaps that persist between disadvantaged minorities and more prosperous members of other racial/ethnic groups.

References

Blake, Howard. 1954. "Class Size: A Summary of Selected Studies in Elementary and Secondary Public Schools." Unpublished Ph.D. dissertation, Columbia University.

Campbell, Ernest Q. 1969. "Negroes, Education and the Southern States." *Social Forces*, vol. 47, no. 3 (March).

Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfeld, and Robert L. York. 1966. *Equality of Educational Opportunity*. Washington, D.C.: Government Printing Office.

Farkas, George et al. 1995. "Can All Children Learn to Read at Grade-Level by the End of Third Grade?" Working Paper, The Center for Education and Social Policy, University of Texas at Dallas, September 20.

Farley, Reynolds. 1993. "Changes in the Segregation of Anglos from Blacks during the 1980s: Small Steps Toward a More Integrated Society." Population Studies Center, University of Michigan, July.

Ferguson, Ronald F. 1991. "Paying for Public Education: New Evidence on How and Why Money Matters." *Harvard Journal on Legislation*, vol. 28, no. 2 (Summer), pp. 465–98.

Ferguson, Ronald F. and Helen F. Ladd. 1995. "Additional Evidence on How and Why Money Matters: A Production Function Analysis of Alabama Schools." Harvard and Duke Universities, June 29.

Hanushek, Eric A. 1971. "Teacher Characteristics and Gains in Student Achievement: Estimation Using Micro-Data." *The American Economic Review*, vol. 61, pp. 280–88.

———. 1972. *Education and Race: An Analysis of the Educational Production Process*. Cambridge, MA: Heath-Lexington.

———. 1981. "Throwing Money at Schools." *Journal of Policy Analysis and Management*, vol. 1, pp. 19–41.

———. 1986. "The Economics of Schooling: Production and Efficiency in Public Schools." *Journal of Economic Literature*, vol. XXIV (September), pp. 1141–77.
———. 1989. "Impact of Differential Expenditures on School Performance." *Educational Researcher*, vol. 18, no. 4 (May), pp. 45–51.
———. 1991. "When School Finance 'Reform' May Not Be a Good Policy." *Harvard Journal on Legislation*, vol. 28, pp. 423–56.
Hanushek, Eric A. and John F. Kain. 1972. "On the Value of Equality of Educational Opportunity as a Guide to Public Policy." In *On Equality of Educational Opportunity*. Frederick Mosteller and Daniel P. Moynihan, eds. New York: Random House, Inc.
Hedges, Larry, Richard Lane, and Rob Greenwald. 1994. "Does Money Matter? A Meta-Analysis of Studies of the Effects of Differential School Inputs on Student Outcomes." *Educational Researcher*, vol. 23, no. 3, pp. 5–14.
Jencks, Christopher. 1972. "The *Coleman Report* and the Conventional Wisdom." In *On Equality of Educational Opportunity*, Frederick Mosteller and Daniel P. Moynihan, eds. New York: Random House, Inc.
Kain, John F. 1968. "Housing Segregation, Negro Employment, and Metropolitan Decentralization." *Quarterly Journal of Economics*, vol. LXXXII, no. 2 (May), pp. 175–97.
———. 1979. "Failure in Diagnosis: A Critique of the National Urban Policy." *The Urban Lawyer*, vol. II, no. 2 (Spring).
———. 1985. "Black Suburbanization in the Eighties: A New Beginning or a False Hope?" In *American Domestic Priorities: An Economic Appraisal*, John M. Quigley and Daniel L. Rubinfeld, eds. Berkeley, CA: University of California Press.
———. 1992. "The Spatial Mismatch Hypothesis: Three Decades Later." *Housing Policy Debate*, vol. 3, issue 2, pp. 371–460.
———. 1995. "Impact of Minority Suburbanization on the School Attendance and Achievement of Minority Children: Final Report of Work Completed Under Spencer Foundation Small Grant." January 28.
Kain, John F. and Joseph J. Persky. 1969. "Alternatives to the Gilded Ghetto." *The Public Interest*, Winter, pp. 74–87.
Massey, Douglas S. and Nancy A. Denton. 1993. *American Apartheid: Segregation and the Making of the Underclass*. Cambridge, MA: Harvard University Press.
Mosteller, Frederick and Daniel P. Moynihan, eds. 1972. *On Equality of Educational Opportunity*. New York: Random House, Inc.
Murnane, Richard. 1975. *The Impact of School Resources on the Learning of Inner City Children*. Cambridge, MA: Ballinger.
———. 1981. "What Do Effective Teachers of Inner-City Children Have in Common?" *Social Science Research*, vol. 10, pp. 83–100.
Orfield, Gary and Carole Ashkinaze. 1991. *The Closing Door: Conservative Policy and Black Opportunity*. Chicago, IL: University of Chicago Press.
Orfield, Gary and Franklin Monfort. 1992. "Status of School Desegregation: The Next Generation." Report to the National School Boards Association, Metropolitan Opportunity Project.
Owen, John D. 1972. "The Distribution of Educational Resources in Large American Cities." *The Journal of Human Resources*, vol. VII, no. 1 (Winter), pp. 26–38.
Rivkin, Steven G. 1994. "Residential Segregation and School Integration." *Sociology of Education*, vol. 67 (October), pp. 270–92.
Summers, Anita and Barbara Wolfe. 1977. "Do Schools Make a Difference?" *The American Economic Review*, vol. 67, no. 4 (December), pp. 639–52.
Texas Education Agency. 1990. "The Texas State Plan and Guidelines for the Education of the Gifted-and-Talented." Austin, Texas.
U. S. Commission on Civil Rights. 1967a. *Racial Isolation in the Public Schools*. Washington, D. C.: Government Printing Office.
———. 1967b. *Racial Isolation in the Public Schools: Appendices*. Washington, D. C.: Government Printing Office.
Welch, Finis, et al. 1987. *New Evidence on School Desegregation*. Los Angeles, CA: Unicon Research Corp.

Discussion

Eric A. Hanushek, Professor of Economics and of Public Policy at the University of Rochester.

This is the right line of research at the right time. The fortieth anniversary of *Brown v. Board of Education* just passed and the thirtieth anniversary of the Coleman Report (Coleman et al. 1966) is arriving; the time has come for a serious reconsideration of race, segregation, and schooling. Over the past decades, a wide variety of desegregation and compensatory programs have been introduced, so that their effects should now be evident. Additionally, there is a new willingness (perhaps overwillingness) to consider major restructuring and even elimination of programs. Thus, it would be nice to have evidence about what is and is not important in student achievement. Toward this end, John F. Kain and Kraig Singleton are creating a truly unique data set that will permit investigation of some of the key questions that have almost completely eluded educational researchers. And of course Kain, an early interpreter of the Coleman Report and one of the nation's premier researchers into the nexus of race and space, is uniquely prepared to undertake this investigation.

Given the local basis of education and the patterns of local control of educational decisions, a discussion of education is inherently a discussion of the spatial distribution of opportunities. In terms of this conference, the spatial structure of schooling provides clear linkages between today and the future. So it is of some importance to understand how schooling opportunities interact with school attendance patterns and racial disparities in educational quality.

Race and Schooling

The motivation for the Coleman Report, a study mandated in the Civil Rights Act of 1964, was to investigate the "lack of availability of equality of educational opportunity for individuals by reason of race, color, religion, or national origin." This report and the follow-on by the U.S. Commission on Civil Rights (1967), entitled *Racial Isolation in the Public Schools*, focused attention on one of the most obvious characteristics of the schools of the mid-1960s, their separation by race of the students. While not in their direct charge, these studies also began to provide information that could be used to evaluate the achievement effects of what is one of the largest and most long-running social programs in our nation's history—the effort to desegregate the schools of both

The time has come for a serious reconsideration of race, segregation, and schooling.

the Old Confederacy and the rest of the Union. Given this backdrop, it is useful to begin with a quick summary of what we know about race and schooling from these original studies and intervening studies.

My overall summary is as follows:

1. Large disparities by race exist in school performance (measured, say, by the National Assessment of Educational Progress), although the gaps have closed some over the past 10 to 15 years.

2. The racial composition of schools has changed in fairly complicated ways related to the imposition of desegregation policies (voluntary or otherwise), to the development of housing patterns within cities, and to the general decentralization of the population. Nonetheless, while the patterns vary across regions, the amount of racial contact in the schools has increased over the past decades (Welch and Light 1987).

3. The racial composition of the schools has minimal effects on student test performance, other things being equal.

4. Quality of the schools may be correlated with racial composition, although this is not particularly well documented.

5. Limited progress has been made in addressing the important issues of how school policies interact with racial disparities in performance, largely because the data available for analysis have not been at all adequate.

Kain and Singleton have embarked on a data construction effort in the state of Texas that directly addresses point 5 and holds promise for filling in the details on points 3 and 4. Their data set, which is still under construction, could become the richest data set ever compiled to address central issues of educational policy, particularly as related to race and space. Until now, the largest and most comprehensive data base has been the one for the original Coleman Report, even though it has a number of serious flaws for the investigations of interest here. The Kain and Singleton data set will clearly leapfrog that data base. Without repeating their description, the key features include the extraordinarily large samples, the ability to follow individual students over time, and the ability to link school resources rather closely to individual students.

The Kain-Singleton Analysis

The analysis in this paper largely concentrates on a series of very important descriptive issues. While this analysis considers only a small part of what they can eventually exploit, the authors begin to provide important insights that motivate analyses yet to come.

The basic starting point is a finding that clear and systematic differences exist in student test performance by race and ethnic background. While not surprising in light of other data, this finding sets the scene for the central analysis. An important point, however, is that the differences are larger for low-income blacks and Hispanics. This interaction between income and race is less well known or documented in past work.

Most of the new analytical efforts within this paper are devoted to understanding the distribution of school resources across schools in Texas. Before doing this, however, they present what I believe is the key table for interpreting all of the results—their Table 6. Table 6 presents the only estimates in the paper of the determinants of student performance. These 6th-grade results are clearly preliminary and subject to modification with further refinements. Nonetheless, they are rather remarkable. The first column presents estimates of achievement models that employ just income-race interactions (plus student gender and age). The fifth column presents estimates of this same model with individual school fixed effects, that is, a dummy variable for each of the about 2,000 separate campuses. At least at the visual level, the estimated differences in performance by race appear *independent of school level inputs*. In other words, the racial differences are not affected by differences in school level resources.

This finding does not particularly surprise me, because I have long held that school quality is not closely related to expenditure or conventional school inputs (Hanushek 1986). But it does provide a somewhat different interpretation of much of the Kain-Singleton analysis.

The focus of attention of their study is how school resources vary by race of the school. They examine scores on teachers' tests (TECAT), master's degrees, teaching experience (novice or old), and class size. The analysis is very clever and demonstrates the power that comes from their data set. They investigate how resources differ by race, holding constant overall district factors through the use of district fixed effects. The general form of the regressions calls for regressing each of the school resource measures on percent black, percent Hispanic, and interactions with income along with a district fixed-effect term.

Several aspects of these analyses stand out. First, and most important, these resources consistently are distributed such that more resources go to schools with low minority populations. Schools with high proportions of blacks and Hispanics simply get less of each of these resources.

Second, and somewhat unexpected, the pattern and the magnitude of these race effects are very similar across districts. Large MSAs as a group or individual large districts look quite similar to all districts in the state. (Again, these conclusions are not based on formal statistical tests but instead on qualitative summaries of the estimated models.) The apparent uniformity belies conventional views that such race effects are larger and more intense in the big urban centers.

Third, their careful consideration of the measurement of inputs is admirable. They work hard at constructing solid estimates of teacher test scores. They also provide an interesting supplemental analysis of how class size varies widely by type of instruction and grade level, adding a real caution about inherent conceptual difficulties in measuring class sizes for districts. Average class size for a district, for example, will be a very poor measure of potential performance effects if there are nonlinearities in how class sizes affect performance.[1]

As mentioned, the interpretation by many people of these resource variations is that they indicate disparities in the quality of schooling received by students. My interpretation is different, because the evidence on resources indicates that master's degrees and class size are not closely related to student performance. For example, in 277 separate estimates of the effects of teacher-pupil ratios on student outcomes, 15 percent find statistically significant positive effects while 13 percent find statistically significant negative effects (Hanushek, Rivkin, and Taylor 1995). The remaining 72 percent are statistically insignificant; that is, we are not very confident that student outcomes are affected by teacher-pupil ratios. Teacher experience shows somewhat stronger effects but, as Kain and Singleton point out, causality is not well sorted out. The evidence on test scores tends to be stronger: 26 of the 36 studies with estimated effects on student achievement are positive and 15 of those are statistically significant.[2] Thus, past work might suggest taking the TECAT variations more seriously in terms of potential effects on student outcomes.

Kain and Singleton's data set could become the richest ever constructed to address central issues of educational policy, particularly as related to race and space.

But remember Table 6. That table suggests that school-level differences do not affect racial differences in student performance. By implication this supports a finding of "no effect" of these factors, because we know that these factors are themselves distributed in a systematic manner by race and ethnicity.

The overall patterns of resource variations remain inherently interesting. If these hold up to further refinement of the data and analyses, they suggest

[1] Some people, beginning with Glass and Smith (1979), argue that class sizes above some level have little effect on performance but have significant effects below a cut-off—roughly 15 students per teacher in the Glass and Smith analysis. Ferguson (1991) argues from Texas data that class size effects become more important when pupil-teacher ratios rise above a threshold. Specifically, "the number of 'students per teacher' is important only when it exceeds eighteen" (p. 477). Both of these studies imply nonlinear responses to variations in class size, and suggest that aggregation across grades and schools within districts will lead to significant biases.

[2] This summary omits the five studies that report statistically insignificant effects but do not report the sign of the estimated relationship.

systematic discrimination in the operation of schools. Resources that are conventionally thought to affect student achievement are systematically distributed toward the majority whites in Texas and away from the blacks and Hispanics. We can thus be thankful that these resources in reality do not appear to have much to do with student performance.

Finally, I must conclude with a statement of anticipation. Kain and Singleton have constructed a data base that is likely to become the source of much new knowledge about schooling. Issues ranging from the effects of school desegregation to the impacts of student migration to the effects of special education and other distinct programs all can be brought under the spotlight of their data. They should be encouraged to work faster, so we can have the answers sooner.

References

Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfeld, and Robert L. York. 1966. *Equality of Educational Opportunity*. Washington, D.C.: Government Printing Office.

Ferguson, Ronald F. 1991. "Paying for Public Education: New Evidence on How and Why Money Matters." *Harvard Journal on Legislation*, vol. 28, no. 2 (Summer), pp. 465–98.

Glass, Gene V. and Mary Lee Smith. 1979. "Meta-analysis of Research on Class Size and Achievement." *Educational Evaluation and Policy Analysis*, vol. 1, no. 1, pp. 2–16.

Hanushek, Eric A. 1986. "The Economics of Schooling: Production and Efficiency in Public Schools." *Journal of Economic Literature*, vol. 24, no. 3 (September), pp. 1141–77.

Hanushek, Eric A., Steven G. Rivkin, and Lori L. Taylor. 1995. "Aggregation and the Estimated Effects of School Resources." Working Paper No. 397, Rochester Center for Economic Research, University of Rochester, February.

U.S. Commission on Civil Rights. 1967. *Racial Isolation in the Public Schools*. Washington, D.C.: Government Printing Office.

Welch, Finis, and Audrey Light. 1987. *New Evidence on School Desegregation*. Washington, D.C.: U.S. Commission on Civil Rights.