DOCUMENT RESUME

ED 389 693 SP 036 337

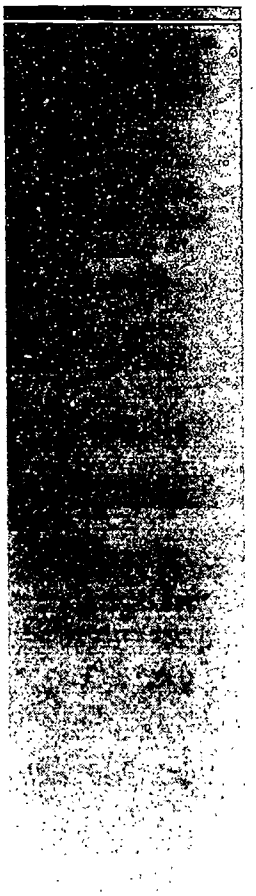
AUTHOR Ehrenberg, Ronald G.; Brewer, Dominic J.
TITLE Did Teachers' Verbal Ability and Race Matter in the 1960s? "Coleman" Revisited. RAND Reprints.
INSTITUTION Rand Corp., Santa Monica, CA. Inst. on Education and Training.
REPORT NO RAND/RP-413
PUB DATE 95
NOTE 27p.; Reprinted from "Economics of Education Review", Vol. 14, No.1, 1995, pp. 1-21.
AVAILABLE FROM Distribution Services, RAND, 1700 Main St., P.O. Box 2138 Santa Monica, CA 90407-2138.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Academic Achievement; Black Students; *Black Teachers; Data Analysis; Elementary School Students; *Elementary School Teachers; *Equal Education; High Schools; High School Students; *Race; Racial Differences; Research Methodology; *Secondary School Teachers; Teacher Characteristics; *Verbal Ability; Whites; White Students
IDENTIFIERS *Coleman Report

ABSTRACT

This paper reanalyzed data from the classic 1966 study "Equality of Educational Opportunity," or "Coleman Report." It addressed the issue of whether teacher characteristics, including verbal ability and race, influenced "synthetic gain scores" of students (mean test scores of upper grade students in a school minus mean test scores of lower grade students in the school), in the context of an econometric model that allow for the possibility that teacher characteristics in a school are endogenously determined. The study found that verbal aptitude scores of teachers influenced synthetic gain scores for both black and white students. Verbal aptitude mattered as much for black teachers as it did for white teachers. Finally, holding teacher characteristics other than race constant, in some specifications black teachers were associated with higher gain scores for black high school students, but lower gain scores for white elementary and secondary students. The study noted that because these findings were for American schools in the mid-1960s they could not be directly applied to contemporary experience. However, the results do raise issues that should be addressed in hiring policies in American education. (Contains 83 references.) (JB)

RAND



Did Teachers' Verbal Ability and Race Matter in the 1960s?

Coleman Revisited

Ronald G. Ehrenberg and Dominic J. Brewer

Reprinted from *Economics of Education Review*

Institute on Education and Training

Published 1995 by RAND
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138

RAND

REPRINTS

RAND/RP-413

Did Teachers' Verbal Ability and Race Matter in the 1960s?

Coleman Revisited

Ronald G. Ehrenberg and Dominic J. Brewer

Reprinted from *Economics of Education Review*

Institute on Education and Training

Pergamon

0272-7757(94)00031-X

Did Teachers' Verbal Ability and Race Matter in the 1960s? *Coleman* Revisited

RONALD G. EHRENBERG* and DOMINIC J. BREWER†

* New York State School of Industrial and Labor Relations, Cornell University, Ithaca, NY 14853-3901, U.S.A. and National Bureau of Economic Research
†The Rand Corporation, 1700 Main Street, Santa Monica, CA 90407, U.S.A.

Abstract — Our paper reanalyzes data from the classic 1966 study *Equality of Educational Opportunity*, or *Coleman Report*. It addresses whether teacher characteristics, including verbal ability and race, influenced "synthetic gain scores" of students (mean test scores of upper grade students in a school minus mean test scores of lower grade students in a school), in the context of an econometric model that allows for the possibility that teacher characteristics in a school are endogenously determined.

We find that verbal aptitude scores of teachers influenced synthetic grain scores for both black and white students. Verbal aptitude mattered as much for black teachers as it did for white teachers. Finally, holding teacher characteristics other than race constant, in some specifications black teachers were associated with higher gain scores for black high school students, but lower gain scores for white elementary and secondary students. Because these findings are for American schools in the mid-1960s, they do not directly apply to our contemporary experience. However, they do raise issues that should be addressed in discussions of hiring policies in American education. [*JEL* 121]

I. INTRODUCTION

MOTIVATED by the poor academic performance and high drop-out rates of many minority elementary and secondary school students vis-a-vis their white counterparts, as well as the fact that the racial/ethnic distribution of public school faculty often does not reflect the racial/ethnic distribution of their students, many school districts have aggressively sought to increase their hiring of minority faculty. This policy has been purchased even in the face of a declining pool of minorities seeking to enter careers in education and evidence that new minority teachers tend to fail the National Teacher Examination at a higher rate than new white teachers.[1] Confronted by fiscal stringency, many school districts have also begun to institute early retirement plans to encourage older, more experienced, and often white, teachers to retire, thereby creating vacancies for lower paid new, or relatively inexperienced, minority teachers.

These policies raise a host of issues. Minority teachers are thought by many to be more effective teachers of minority students because the former may serve as role models for, may interact better with, may have more favorable attitudes towards and higher expectations for, and may provide more positive feedback to, minority students.[2] Ultimately, however, society must be concerned about minority teachers' impacts on the educational (test scores, completed schooling levels) and post-educational (labor market success) *outcomes* of *both* minority and white students. Only if minority teachers improve, or at least leave unchanged, the outcomes of *both* groups (as compared to what white teachers would generate) can minority recruitment policies in public education be judged *pareto optimal* in terms of their impacts on students. If they improve the outcomes for minority students but reduce the outcomes for white students, the debate over these policies will shift to their distributional (across student group) consequences. If minority teachers are shown to have no impact on the outcomes for minority students and to adversely

influence those for white students, the debate will shift to one over the importance society places on providing employment opportunities for minority teachers to help remedy historical inequities and perceptions of current discrimination against potential minority teachers.

Of course, minority and white teachers differ, on average, on a number of dimensions other than race. They come from different socioeconomic backgrounds, have different levels of experience, have different degree levels, and tend to have received their degrees from different institutions. They also tend to score differently on standardized aptitude and achievement tests. While issues relating to the "cultural bias" in test scores have been raised, studies do suggest that student's academic performance is related, on average, to their teachers' performance on standardized tests.[3] Comparisons of the effectiveness of minority and white teachers must control, if possible, for these other characteristics.

Research on the relative effectiveness of minority teachers has been conducted primarily by sociologists, psychologists, and educational researchers. Most studies have focused on teachers' attitudes, teachers' expectations, teachers' placement of students, and the feedback (positive and negative) that teachers provide students.[4] Only a few have addressed educational outcomes and none has addressed subsequent labor market success.[5] Many have been studies of a single school district and these typically failed to control for other teacher characteristics. Only a few studies used representative national data bases, only a few attempted to model the process by which teachers get assigned to schools, and none controlled for this process in the estimation of teacher effects. To our knowledge, none addressed whether the effects of teachers' verbal ability vary either with the race of the teachers or the race of the students they are teaching.

Our paper begins to address some of the issues we have raised, by reanalyzing data from the classic 1966 study *Equality of Educational Opportunity*, or *Coleman Report*. As we describe in the next section, these data permit us to estimate how, during the mid-1960s, the characteristics of teachers of different races (verbal aptitude, degree levels, years of experience) influenced an estimate of the change in test scores over a three grade level period, for students of different races. They also permit us to test whether controlling for the process by which teacher characteristics (including race) get assigned to different schools influences our estimated relationships.

After discussion of the *Coleman Report* in the next section, sections III and IV present our empirical analyses. The two final sections then discuss the significance of our findings and their implications for future research.

II. THE COLEMAN REPORT

The *Coleman Report* represented an important step in educational research. Its statistical analyses, based on data from over 570,000 pupils, 60,000 teachers, and 4,000 principals, represented the beginning of the "educational production function" literature.

The methodological approaches used in the *Coleman Report* were severely criticized and numerous reanalyses of the data took place within a few years of the *Report*'s publication.[6] Most social scientists, and public attention, focused on its conclusions concerning the extent of race segregation in schools and the importance of family background characteristics in explaining variations in student achievement. Less well known (or well remembered) is that the underlying data set contained information on teacher verbal ability (as measured by scores on a verbal aptitude test) and that the average verbal aptitude of teachers in a school was seen to be positively correlated with student test scores. Both the original *Coleman Report* and subsequent reanalyses of its data found this correlation and some researchers concluded that the correlation appeared stronger at higher grade levels.[7]

The *Coleman Report* data appear to be unique among existing micro level data sets in that they contain a measure of individual teachers' verbal ability. One serious weakness of the data, however, is that they represent a "snapshot" at a single point in time and that only a current year test score measure exists for each student. Subsequent educational research by economists has stressed that to more fully control for unobservable student, family, and community characteristics that influence student achievement, one should relate school characteristics, including teacher ability levels, to student gain scores, or changes in test scores over

time — not to student test score levels at a point in time.[8]

While all of the prior research that used the *Coleman Report* data estimated current year test score equations, the data do in fact contain information on third and sixth graders at each elementary school and ninth and twelfth graders at each high school. Moreover, one can identify the subset of third and sixth graders who spent their entire elementary school careers at a given elementary school and similarly identify the subset of twelfth graders who spent their entire high school careers at a given high school. Restricting one's attention to these subsets of student respondents, computing mean scores for each school by grade, and assuming that within a school, the unobservable characteristics do not vary across grades, one can compute "synthetic" gain scores as the difference between the mean test scores in the two grades for each school. For example, the difference between the mean test scores of sixth graders in a school and the mean test scores of third graders in a school at the survey date can be taken as an estimate of how much third graders in the school would learn if they remained in the school for three more years.[9] In cases where the schools have a significant number of both white and black students, these gain scores can also be computed separately for each racial group.[10]

These gain scores are used as dependent variables in the next section in the estimation of educational production functions in which the gain scores by school are related to student family, community, school, and teacher characteristics. Of primary interest to us will be the effect of the racial composition of teachers in a school and their verbal abilities on the gain scores of students of each racial group.[11] Given prior mentioned concerns about the alleged "cultural bias" of tests, we also address whether increasing the verbal ability of teachers of each race has the same impact on the gain scores of students of different races.

While conceptually such an analysis is straightforward, an important statistical issue exists. Teachers are not randomly assigned to schools and school districts; teachers with higher test scores may be more easily attracted to higher paying districts, districts with smaller class sizes, and districts whose families are highly educated. Similarly, teachers may prefer to work with students who come from the same racial group or from similar socioeconomic backgrounds. To make sure that such nonrandom assignment does not lead to biased estimates of the effects of teacher characteristics on student gain scores, an instrumental variable approach is employed in section IV to control for the process by which teachers and schools are matched.[12]

III. ESTIMATING SYNTHETIC GAIN SCORE EQUATIONS

A. Descriptive Statistics

The *Equality of Educational Opportunity (EEO)* data tapes that we received from the National Archives contained data for third grade students at 2499 schools, sixth grade students at 2389 schools, ninth grade students at 930 schools and twelfth grade students at 787 schools.[13] We restricted our attention to the subset of elementary schools for which data were reported for both elementary grades and the subset of high schools for which data were reported for both secondary grades. We also required that data for each school were reported on all of the explanatory variables used in the analyses that follow, including the characteristics of teachers.[14] All data were aggregated to obtain school level mean values for the entire sample, for white students at the subset of schools that had some white students in attendance, and for black students at schools that had some black students in attendance.[15]

As Table 1 indicates, the restrictions left us with a maximum sample of 969 elementary schools and 256 high schools.[16] Of the former, 799 had at least one white student in both grades and 514 had at least one black student in both grades. Of the latter, 178 had at least one white student in both grades and 183 had at least one black student in each grade. Because the elementary school sample sizes are so much larger, the majority of the analyses that follow use the elementary school data.

Students in each grade were administered verbal aptitude, nonverbal aptitude, reading and mathematics tests. The weighted (by number of students taking the tests) mean percentage of correct answers on the four tests across schools was 58.18 for third graders and 52.28 for ninth graders. The weighted mean synthetic gain scores, the mean for the sixth grade minus the mean for the third grade and the mean for the twelfth grade minus the mean for the ninth grade, were 1.55 and 0.90, respectively.

At the elementary school level 31% of the students and 27% of the teachers were black, while

Table 1. Descriptive statistics: weighted means (by school size)

	Elementary schools			High schools		
	(1E) All schools	(2E) White students	(3E) Black students	(1H) All schools	(2H) White students	(3H) Black students
BYTEST	58.18	62.89[a]	50.90	52.28	59.10[a]	38.93
GAIN	1.55	5.86[a]	−4.42	0.90	0.86	−0.53
BLACKS	31.12	6.63[a]	77.84	29.24	6.32[a]	82.06
BLACKT	27.00	4.00[a]	71.00	25.00	3.00[a]	77.00
EXP	15.81			14.89		
MA	17.69			34.23		
VERB	75.48			76.71		
BEXP		14.37	15.50		13.95	14.94
BMA		22.17	18.35		42.72	25.28
BVERB		73.54[a]	64.03		80.20[a]	66.91
WEXP		15.96	14.89		14.71	14.49
WMA		17.66	17.83		37.05	39.73
WVERB		81.13	78.76		81.17	76.62
P_B		0.38	0.79		0.42	0.77
P_W		0.92	0.62		0.92	0.55
N	969	799	514	256	178	183

where

BYTEST — mean percentage of correct answers of third grade students in the school on the group of tests (ninth grade for high schools)

GAIN — mean percentage of correct answers of sixth grade students in the school on the group of sixth grade tests minus the mean percentage of correct answers of third grade students in the school on the group of third grade tests (twelfth and ninth grades for high schools)

BLACKS — percentage of black students in the third and sixth grades in the school (ninth and twelfth for high school)

BLACKT — percentage of black teachers in the school

EXP — mean years of teaching experience of teachers in the school

MA — mean percentage of teachers in the school with at least a master's degree

VERB — mean percentage of correct answers of teachers in the school on the verbal test

BEXP, BMA, BVERB — mean values for black teachers in schools with positive numbers of black teachers in the sample

WEXP, WMA, WVERB — mean values for white teachers in schools with positive numbers of white teachers

P_B — proportion of schools with at least one black teacher

P_W — proportion of schools with at least one white teacher

N — number of schools

[a] mean for white students is significantly different from the mean for black students at the 0.05 level of significance, two-tail test.

Source: Authors' computations from the *Equality of Educational Opportunity* survey data tapes.

at the high school level the comparable percentages were 29 and 25.[17] Elementary school teachers in the sample averaged close to 16 years of teaching experience, about 17% of them had earned at least a masters degree and, on average, they answered correctly slightly more than 75% of the questions on a verbal aptitude test that was administered to them. High school teachers were quite similar on their experience and their verbal aptitude scores, but over 34% of them had at least a masters degree.

Table 1 also summarizes these data separately for white students and for black students, indicating in each case how their teachers' characteristics varied by race. White students tended to have higher base year test scores and larger gain scores than black students at both the elementary and secondary levels.[18] The typical white elementary school student was enrolled in a school in which about 6% of the students and 4% of the teachers were black, while the typical black elementary school student

found 77% of his or her classmates and 71% of his or her teachers black. The respective values for high school students were quite similar.

Both white and black teachers' verbal aptitude scores tended to be higher if they were employed at schools in which white students were enrolled than if they were employed in schools in which black students were enrolled. In addition, white teachers' verbal aptitude scores were higher than black teachers' verbal aptitude scores in both types of schools, although at the high school level the difference at schools in which white students were enrolled was quite small. Experience and degree differences across the two types of schools and, within a type, across the two types of teachers were less uniform and were not always in favor of white teachers. For example, in both elementary and secondary schools at which white students were present, black teachers were more likely than white teachers to have advanced degrees.[19]

B. Elementary School Analyses

Estimates are presented in column (1) of Table 2 of synthetic gain score equations of the form,

$$GAIN_i = \alpha_0 + \alpha_1 X_i + \alpha_2 S_i + \alpha_3 BYTEST_i + \alpha_4 T_i + \epsilon_i. \quad (1)$$

Here X_i represents a vector of characteristics of school i's students, their families, and the community in which the school is located; S_i represents a vector of characteristics of the school; and T_i represents a vector of characteristics of the school's teachers. $BYTEST_i$ is the average test score of grade 3 students in the school and $GAIN_i$ is the difference between the average test score of grade 6 students in the school and the average test score of grade 3 students in the school.

Included in X_i are the percentages of the school's students that are female (FEM), black (BLACKS), have no father or no mother in the home (FNHH, MNHH), have a telephone in the household (PHONE), and receive free lunches (FLNCH), the mean income of the families of the school's students (INCOME), the mean education levels of the fathers and mothers of the school's students (FED, MED), and whether the school is located in a central city (CITY), rural (RURAL), or suburban (the omitted category) area.[20] The school characteristics are the number of books per pupil in the school's library (BOOKS) and the pupil/teacher ratio in the school (PUPT). Finally, the teacher characteristics are the percentage that are black (BLACKT), the mean years of teaching experience (EXP), the percentage with at least a masters degree (MA), and the mean verbal test score of teachers in the school (VERB). Since the schools in the sample vary considerably in size, the method of weighted least squares is used to obtain the estimates.

Gain scores prove to be higher in schools with a greater percentage of female pupils, a smaller percentage of black students, fewer families with only one parent in the household, more families with telephones, fewer families receiving free lunches, and higher parental education levels. Relative to suburban schools, gain scores are higher in rural schools and lower in central city schools. Higher pupil/teacher ratios are associated with lower gain scores. Finally, "regression to the mean" is present, as higher base year test scores are associated with lower gain scores.

Of primary concern to us is the role that teacher characteristics play. In this model, increasing the percentage of black teachers by 10 percentage points is associated with a 0.4 decrease in the school's gain score. Teachers' experience is positively associated with the gain score, but teachers' degree level does not appear to matter. Crucially, higher verbal aptitude scores for teachers are associated with higher gain scores for students. If teachers' verbal aptitude scores could be increased by 10 percentage points, gain scores are predicted to be 0.9 points higher.[21] The latter should be contrasted to a mean gain in the sample of 1.55.

The remaining columns in Table 2 ascertain the sensitivity of these results to changes in the model's specification. While inclusion of the base year (third grade) test score on the right-hand side of equation (1) is justified because how much students learn over time depends on where they are starting from, there are well-known statistical problems that result. On the one hand, if the base year test score and the gain score are both influenced by a common set of variables and any of these variables are omitted from equation (1), then inclusion of BYTEST may lead to biased estimates of the coefficients of other variables in the model.[22] On the other hand, if BYTEST measures students' true abilities with error, its coefficient will be biased towards zero and thus our estimate of the extent of regression to the mean overstated.

One way to handle this problem is to omit BYTEST from the model and see whether this

Table 2. School level synthetic gain score equations: Grades 3 to 6[a,b] (absolute value *t* statistics)

	All students			Students who never changed schools	
	(1)	(2)	(3)	(4)	(5)
INTER	11.483 (3.8)	−13.057 (3.9)	10.260 (3.2)	12.174 (3.9)	9.771 (3.0)
FEM	0.059 (3.1)	0.067 (2.9)	0.060 (3.1)	0.043 (3.0)	0.042 (2.9)
BLACKS	−0.060 (5.8)	−0.011 (0.9)	−0.057 (5.5)	−0.083 (6.5)	−0.081 (6.3)
FNHH	−0.062 (3.3)	−0.029 (1.3)	−0.061 (3.1)	−0.037 (2.1)	−0.037 (2.1)
MNHH	−0.243 (9.0)	−0.194 (6.5)	−0.242 (8.8)	−0.109 (4.2)	−0.105 (4.0)
PHONE	0.100 (7.5)	0.051 (3.2)	0.098 (7.2)	0.123 (9.2)	0.122 (9.2)
INCOME	0.100 (0.9)	0.229 (1.0)	0.233 (1.1)	0.277 (1.5)	0.344 (1.9)
FLNCH	−0.029 (2.6)	−0.008 (0.6)	−0.029 (2.6)	−0.026 (2.0)	−0.024 (1.9)
FED	1.079 (4.7)	0.501 (1.8)	1.102 (4.8)	0.796 (4.0)	0.791 (4.0)
MED	0.358 (1.3)	−0.046 (0.1)	0.317 (1.1)	0.534 (2.2)	0.537 (2.2)
CITY	−0.894 (1.9)	1.062 (1.9)	−0.876 (1.8)	−0.677 (1.2)	0.395 (0.7)
RURAL	1.080 (2.3)	1.017 (1.8)	1.052 (2.2)	0.915 (1.8)	0.829 (1.6)
BOOKS	0.010 (0.3)	0.015 (0.3)	0.012 (0.3)	0.013 (0.3)	0.012 (0.3)
PUPT	−0.081 (2.3)	0.002 (0.0)	−0.076 (2.2)	−0.076 (2.0)	0.061 (1.6)
STAY			0.009 (1.0)		0.030 (2.7)
BYTEST	−0.638 (20.4)		−0.634 (20.0)	−0.653 (19.9)	0.656 (20.0)
BLACKT	−0.042 (3.8)	−0.060 (4.9)	−0.045 (3.9)	−0.042 (3.0)	0.045 (3.2)
EXP	0.068 (2.0)	−0.023 (0.5)	0.063 (1.8)	0.058 (1.5)	0.046 (1.2)
MA	−0.004 (0.3)	0.006 (0.5)	−0.003 (0.3)	−0.001 (0.1)	0.001 (0.1)
VERB	0.093 (4.1)	0.063 (2.4)	0.096 (4.2)	0.076 (3.1)	0.081 (3.2)
n	969	969	953	928	928
$\bar{R}^2$	0.728	0.616	0.727	0.698	0.700

where:

INTER	intercept term
FEM	percentage of the school's students that are female
BLACKS	percentage of the school's students that are black
FNHH	percentage of the school's students without a father in the household
MNHH	percentage of the school's students without a mother in the household
PHONE	percentage of the school's students with a telephone in the household
INCOME	mean income of the families of the school's students (in thousands)
FLNCH	percentage of the school's students that receive free lunches
FED	mean years of education of fathers of the school's students
MED	mean years of education of mothers of the school's students
CITY	1 = central city school, 0 = other
RURAL	1 = rural school, 0 = other
BOOKS	number of books in the school's library (000s) per pupil
STAY	percentage of sixth grade students who have not changed schools since the first grade
BYTEST	mean grade 3 test score (percentage of correct answers) in the school
BLACKT	percentage of black teachers in the school
PUPT	pupils per teacher in the school
EXP	mean years of experience of teachers in the school
MA	percentage of teachers with at least a master's degree in the school
VERB	mean verbal test score of teachers in the school

[a]Each student's test score is the simple average of the percentage of correct answers the student received on verbal, nonverbal, reading, and mathematics tests.

[b]Weighted least squares regressions. The weight used is $[(N_3 \times N_6)/(N_3 + N_6)]$ where N_3 (N_6) is the number of third (sixth) grade students taking the test in the school.

Source: Author's computations from the *Equality of Opportunity* survey data tapes

substantially influences the other coefficients. This is done in column (2). While mean teacher experience is now no longer significant, teacher verbal ability is still positively, and the percentage of black teachers negatively, associated with a school's gain score.[33]

Does mobility of students into and out of a school influence the amount of learning that goes on in the

Table 3 Selected coefficients from school level synthetic gain score equations for students disaggregated by race: Grades 3 to 6[a] (absolute value *t* statistics)

	White students				Black students			
	(1)	(1w)	(3w)	(5w)	(1)	(1b)	(3b)	(5b)
BLACKS	0.017 (1.2)	0.018 (1.2)	0.018 (1.2)	0.055 (2.5)	0.025 (1.5)	0.013 (0.7)	0.021 (1.1)	0.006 (0.2)
BLACKT	0.157 (8.4)	0.334 (4.2)	−0.309 (3.6)	0.276 (2.5)	−0.018 (1.2)	−0.013 (1.9)	−0.157 (2.2)	−0.174 (1.9)
EXP	0.111 (3.6)				0.024 (0.4)			
MA	−0.039 (3.8)				0.052 (2.7)			
VERB	0.080 (3.3)				0.086 (2.5)			
BTEXP		−0.220 (1.0)	0.229 (1.0)	0.224 (0.8)		0.008 (0.1)	0.035 (0.4)	0.002 (0.1)
BTMA		0.070 (1.0)	0.098 (1.3)	0.138 (1.5)		0.095 (4.3)	0.089 (3.9)	0.094 (3.5)
BTVERB		0.378 (3.6)	0.352 (3.2)	0.301 (2.1)		0.134 (3.2)	0.137 (3.2)	0.112 (2.2)
WTEXP		0.111 (3.6)	0.087 (2.6)	0.070 (1.7)		0.014 (0.1)	0.016 (0.1)	0.090 (0.5)
WTMA		0.046 (1.9)	0.048 (4.6)	0.048 (3.7)		0.087 (2.3)	0.068 (1.5)	−0.045 (0.7)
WTVERB		0.047 (1.9)	0.052 (2.1)	0.025 (0.9)		0.019 (0.2)	−0.013 (0.1)	−0.068 (0.7)
n	799	799	741	705	514	514	438	398
R^2	0.448	0.448	0.364	0.617	0.584	0.617	0.625	0.561

where

BTEXP — proportion of black teachers in the school multiplied by mean years of experience of black teachers in the school

BTMA — proportion of black teachers in the school multiplied by the percentage of black teachers with at least a master degree in the school

BTVERB — proportion of black teachers in the school multiplied by the mean black teacher verbal score in the school

WTEXP, WTMA, WTVERB — defined as above save that each variable now relates to the proportion of white teachers in the school and the characteristics of white teachers in the school

[a]Also included in the equations are all variables that appear above the first dotted line in columns (1), (3) and (5), respectively, in Table 2. Weighted regressions are again employed as described in footnote b of Table 2, with the weights now based on the numbers of students in the racial group in the school taking the tests in grades three and six.

school? Column (3) adds as an explanatory variable the percentage of sixth grade students in a school that spent their entire elementary school careers in the school (STAY). This variable does not prove to be statistically significant and its inclusion does not substantially influence any of the other coefficients in the model (compare columns (1) and (3)).

Of course, as noted in the introduction, it would be desirable to confine the computation of the synthetic gain scores to students who had always remained in the same elementary school. This is done in columns (4) and (5) where the gain score is now computed as the mean test score for sixth grade students who spent their entire school career at the school minus the mean test score for third grade students who spent their entire careers in the school.[24]

The estimated association between the percentage of black teachers and teachers' verbal aptitude scores and the synthetic gain scores in this restricted sample are quite similar to those found in the unrestricted sample. One new finding, however, is that the synthetic gain scores of these stayers are larger, the larger is the proportion of sixth grade students in the school who spent their entire careers at the school. Put another way, the more turnover there is in a school's student population, the lower the gain scores are for the students who remained at the school.

The analyses reported in Table 2 group all students together. They do not permit us to address an issue that is of key importance to us — whether a teacher's race and verbal ability differentially influence the academic achievement of students of different races. To address this issue, Table 3 reports estimates of selected coefficients from equations that were estimated separately for black and white students. In each case the synthetic gain score is now computed for each school using only data for students of the given race.

Columns (1) for the white and black student samples report coefficients from equations specified identically to column (1) in Table 2. In this model, a higher percentage of black teachers in a school is associated with a lower synthetic gain score for white students but is not associated with a higher gain score for black students. Teacher verbal ability is positively related to gain scores for both groups of students and the magnitude of the relationship is about the same. Teacher experience has a payoff only for white students and having more teachers with advanced degrees enhances learning for black students, but perversely lowers it for white students.

To estimate whether the effects of teacher verbal ability, experience and degree level vary for each group of students with the race of the teachers, expanded versions of the equation underlying columns (1), (3), and (5) of Table 2 were estimated that allowed for interactions. For example, in the case of column (1), the estimating equation became:

$$GAIN_{ij} = \alpha_{0j} + \alpha_{1j}X_{ij} + \alpha_{2j}S_i + \alpha_{3j}BYTEST_{ij} + \alpha_{4j}((BLACKT_i/100)*T_{iB}) + \alpha_{5j}((1-BLACKT_i/100))*T_{iW} + \epsilon_{ij} \quad (2)$$

Here j, equal to 1 or 2, indexes the black or white student equation and T_{iB} (T_{iW}) is the vector of characteristics of black teachers (white teachers) in school i.

Selected coefficients from these models appear in columns (1w), (3w) and (5w) for white students and columns (1b), (3b), and (5b) for black students. The estimates in column (1) for each group come from equations that are restricted versions of the equations that underlie columns (1w) and (1b) and hence one can test for each group whether the restrictions are valid. Formal F tests suggests they are not.[25] That is, we can reject the hypothesis that, for each group of students, the effects of all the teacher characteristics variables are the same for black and white teachers.

The results in this table are striking. In most specifications the percentage of students that are black does not affect either black or white students' gain scores. Higher verbal scores for black teachers are associated with higher gain scores for *both* black and white students. In contrast, white teachers' verbal scores matter only for white students. Higher white teacher experience levels are associated with higher gain scores for only white students and black teachers' experience levels do not appear to have any impact on either group of students' gain scores. Finally, while an increase in the percentage of black teachers in a school with at least a masters degree increases the gain scores of black students, an increase in the comparable percentage for white teachers again perversely is associated with lower gain scores.[26]

Several extensions warrant brief reporting here.[27] First, when the analyses underlying tables 2 and 3 were repeated separately for each of the four individual tests (verbal aptitude, reading, nonverbal

aptitude, and mathematics), teacher verbal ability was found to be positively associated with the synthetic gain scores on all four tests, with the magnitude of the relationship being smallest for the mathematics gain score.

Second, as discussed in footnote 9, students were administered different tests in the third and the sixth grades. It is, therefore, not obvious that the metric we have chosen, the synthetic gain score, is necessarily the correct outcome variable to use.[28] To test for the sensitivity of our findings to the metric chosen, we reestimated the equations found in column 1 of Table 3 using first the difference in the logarithms of the mean sixth and third grade test scores in a school as the outcome variable and then the difference in the school's percentile rank on the sixth grade tests (among all schools in the sample) and its percentile rank on the third grade tests.[29] We found that teacher verbal ability continued to be positively associated with the outcome variable in each of these specifications.

C. Secondary School Analyses

Table 4 contains weighted least square estimates of synthetic gain score equations for the high schools in our *EEO* sample. The outcome variable is now the mean percentage of correct answers of twelfth grade students in the school on their group of tests minus the mean percentage of correct answers of ninth grade students in the school on their group of tests. Because the sample sizes are smaller, fewer coefficients than in the elementary school sample prove to be statistically significant and fewer analyses are reported.

In these high school data, higher verbal aptitude scores of teachers are associated, on average, with higher gain scores for white students, but not for black students. When teacher characteristics are broken down by race, white teachers' verbal aptitude scores appear to matter for both groups of students, but black teachers' verbal aptitude scores do not. Increasing the percentage of black teachers in a school with at least a masters degree is

Table 4. School level synthetic gain score equations: Grades 9 to 12 (absolute value *t* statistics)

	All students	White students		Black students	
	(1)	(1)	(1w)	(1)	(1b)
INTER	2.405 (0.7)	4.161 (0.8)	6.705 (1.2)	4.013 (0.8)	−13.326 (1.4)
FEM	0.021 (1.4)	−0.036 (2.0)	−0.031 (1.8)	−0.021 (0.5)	0.021 (0.8)
BLACKS	−0.025 (1.2)	0.057 (2.0)	0.063 (2.2)	−0.002 (0.1)	−0.017 (0.5)
FNHH	0.099 (3.6)	−0.129 (3.1)	−0.139 (3.3)	−0.085 (2.5)	−0.076 (2.3)
MNHH	0.059 (1.2)	−0.044 (0.6)	−0.056 (0.8)	−0.056 (1.0)	−0.047 (0.8)
PHONE	0.054 (3.3)	0.004 (0.1)	−0.005 (0.2)	0.064 (3.2)	0.064 (3.2)
INCOME	0.695 (2.6)	0.719 (2.2)	0.802 (2.4)	0.888 (1.8)	0.978 (2.0)
FLNCH	0.007 (0.4)	−0.043 (1.5)	−0.049 (1.7)	0.023 (1.6)	0.036 (1.7)
FED	0.773 (2.1)	0.694 (1.5)	0.741 (1.6)	1.603 (2.5)	1.488 (2.2)
MED	0.541 (1.3)	1.268 (2.3)	1.125 (2.0)	−0.849 (1.4)	−0.754 (1.2)
CITY	0.138 (0.3)	0.255 (0.4)	0.496 (0.8)	0.204 (0.3)	−0.060 (0.1)
RURAL	0.328 (0.7)	−0.063 (0.1)	−0.031 (0.1)	0.556 (0.8)	0.645 (1.0)
BOOKS	0.028 (0.6)	0.004 (0.1)	0.022 (0.4)	−0.031 (0.4)	−0.097 (1.2)
PUPT	0.088 (2.1)	0.054 (1.1)	0.051 (1.0)	0.006 (0.9)	0.017 (0.2)
BYTEST	−0.556 (14.7)	−0.621 (12.2)	−0.621 (12.2)	−0.544 (9.4)	−0.553 (9.7)
BLACKT	−0.026 (1.4)	0.102 (2.4)	0.106 (0.3)	0.001 (0.0)	0.229 (2.2)
EXP	0.023 (0.4)	−0.000 (0.0)		0.020 (0.2)	
MA	−0.011 (1.1)	−0.017 (1.4)		0.035 (1.9)	
VERB	0.081 (2.5)	0.104 (2.4)		−0.008 (0.2)	
BTEXP			−0.671 (1.1)		−0.061 (0.6)
BTMA			0.335 (2.0)		0.057 (2.5)
BTVERB			−0.233 (0.6)		−0.046 (0.8)
WTEXP			−0.008 (0.1)		0.321 (1.7)
WTMA			−0.025 (1.9)		0.004 (0.1)
WTVERB			0.092 (2.2)		0.176 (1.5)
n	256	178	178	183	183
R^2	0.488	0.470	0.480	0.360	0.378

[a]See Tables 2 and 3 for variable definitions. BYTEST is now the mean grade 9 test score (percentage of correct answers) in the school.

associated with higher gain scores for both black and white students, although again the percentage of white teachers with at least an MA degree is negatively associated with white students' gain scores.

There is also evidence that white students' gain scores are positively associated with the percentage of students in the school that are black and negatively associated with the percentage of teachers that are black. In contrast, while black students' gain scores are not related to the percentage of students in the school that are black, they do appear to be positively associated with the percentage of black teachers. That is, in the *EEO* data, other things held constant, black teachers do appear to improve the gain scores of black students at the high school level.

IV. CAN SCHOOL AND TEACHER CHARACTERISTICS BE TREATED AS EXOGENOUS?

A. Conceptual Issues

Differences in school or teacher characteristics are not randomly determined across schools. Families choose where to live, and hence their children's schools, based on their own preferences and resource constraints.[30] Teacher characteristics depend upon factors such as the salaries teachers are offered, and the pecuniary and nonpecuniary characteristics of the community in which the school is located.[31] These considerations suggest that failure to treat teacher and school characteristics as endogenous may lead to biased estimates of their affects.[32] Yet to date, virtually all studies of teacher and school affects have treated these characteristics as exogenous.[33]

These biases might arise if the teacher and school characteristics used in the synthetic gain score equations (equations (1) and (2)) are correlated with the error terms in the equations due to an omitted variable problem. For example, suppose that we are interested in the effects of teachers' verbal aptitude, that we assume (for now) that teacher verbal aptitude actually does *not* influence gain scores, and that the omitted variable is a measure of the value that the parents of students place on education. Presumably parents who value education highly will invest more in their children at home (thus leading to higher gain scores) and will also reside in school districts that pay high salaries to attract and retain teachers with high verbal aptitude scores (if they believe, erroneously in our example, that high teacher verbal aptitude enhances learning). Other things held constant, estimation of equations (1) or (2) by least squares would yield a positive relationship between gain scores and teacher verbal aptitude even though we hve assumed (for now) that the true relationship is zero.

The bias arises in our example because of the endogeneity of families' locational decisions coupled with our inability to fully control for unobserved variables that simultaneously influence students' gain scores and their families' locations (which in turn determined teachers' verbal aptitude). We address this problem below by using an instrumental variable estimation method to obtain instruments for the school and teacher characteristics variables, conducting formal statistical tests to ascertain which of these characteristics can be legitimately treated as being exogenous and which must be treated as endogenous, and then reestimating the synthetic gain score equation using the original variables for the exogenous characteristics and the instrumental variable estimates for the endogenous characteristics.

B. Institutional Variable Approach

Appendix Table A2 summarizes the equations we estimated to generate instruments for the school and teacher characteristics variables. These equations were estimated primarily to obtain instruments for the school and teacher variables and they should *not* be thought of as structural equations. In each case, the actual value of these variables were regressed on a set of characteristics of the families of these students at the school, a broader set of characteristics of residents of the county or SMSA in which the school was located, and an estimate of the starting teacher salary in each school district.[34] The county level variables are suitable for use as instruments because they are less likely to be correlated with the error term in the gain score equations than are school or district level variables. The county and SMSA variables were obtained from the *1965 City and County Databook* and starting salaries were obtained from within-school district teacher salary equations that were estimated by us, using the individual teacher data from the *EEO*.[35]

The instruments obtained from the coefficients in Appendix Table A2 were used to test whether the assumptions that all the teacher and school charac-

teristics can be treated as exogenous are valid, using Durbin-Wu-Hausman specification tests in the following two-stage fashion.[36] First, an expanded version of the model that appears in column (1) of Table 2 was estimated that included both the original values and the instruments for all six teacher and school characteristics variables. The coefficients of the instruments for MA and PUPT each had t statistics that exceeded 1.9 in absolute value and a formal F test suggested that one can reject the hypothesis that the vector of coefficients of the six instruments as a set are all zero, and hence that all of the school and teacher characteristics variables should be treated as exogenous.[37] These tests imply that *at least* MA and PUPT should be treated as endogenous in the estimation of this gain score equation.

Second, a version of the model was estimated in which the instruments for MA and PUPT replaced the original values of these variables, but both the original values and the instruments for the other teacher and school characteristics (BOOKS, BLACKT, EXP, VERB) were included. None of the coefficients of the instruments for the latter four variables in this model had a t statistic that exceeded 1.5 in absolute value and a formal F test suggested that one cannot reject the hypothesis that the entire vector of these latter four instruments' coefficients are all equal to zero.[38] This test implies that in a formal statistical sense it is legitimate to treat BOOKS, BLACKT, EXP, and VERB as exogenous in the estimation of the gain score equation.

As a result, the synthetic gain score equation found in column (1) of Table 2 was first reestimated with instruments used only for PUPT and MA. A comparison of the coefficients of the school and teacher characteristics variables that were obtained when weighted least squares was used on the original data (column 1, Table 5) and when a weighted instrumental variable procedure was used with these instruments (column 2, Table 5) suggests that the percentage of black teachers continues to be negatively, and teachers' verbal aptitude positively, associated with the synthetic gain scores. Moreover, the magnitudes of the coefficients of the two variables are roughly the same in the two specifications. Hence, if we take our endogeneity tests at face value, our findings about the roles that teacher ability and race play do not appear to be biased by endogeneity issues.

We note, however, that treating BOOKS, BLACKT, EXP, and VERB as exogenous is *not* consistent with either our underlying economic model or the empirical evidence provided by others.[39] If one also treats BOOKS, BLACKT, EXP, and VERB as endogenous and reestimated the model, the results in column 3 are obtained. While the estimated effect of teacher verbal ability increases substantially, the coefficient of the proportion of black teachers now switches sign and is statistically insignificant. Thus, our conclusion about the effects that teachers' skin color *per se* had on students in the 1960s hinges on the accuracy of our specification tests.[40]

IV. SIMULATIONS

Did teachers' verbal ability and race matter in the 1960s? Our reanalysis of the *EEO* data suggest that the answer to the first question is definitely yes,

Table 5. Comparison of WLS and WIV estimates of school level synthetic gain score equations: Grades 3 to 6[a] (absolute value t statistics)

	(WLS)[b]	(WIV)[c]	
	(1)	(2)	(3)
BOOKS	0.010 (9.3)	0.064 (1.4)	0.479 (1.3)
PUPT	0.081 (2.3)	0.392 (3.0)	0.415 (2.3)
BLACKT	0.042 (3.8)	0.039 (2.9)	0.076 (1.1)
EXP	0.068 (2.0)	0.039 (0.9)	0.105 (0.5)
MA	0.004 (0.3)	0.219 (4.6)	−0.284 (3.9)
VERB	0.093 (4.1)	0.098 (3.6)	0.448 (2.1)

[a] All equations also contain all of the variables used in column (1) of Table 2.

[b] Coefficients from Table 2, column (1).

[c] Instrumental variable estimates – estimates in column (2) use instruments for MA and PUPT only, while the estimates in column (3) use instruments for all five variables.

while the answer to the second question is much more sensitive to the statistical model used. On balance, increases in the verbal aptitude scores of both black and white teachers were associated, other things held constant, with higher synthetic gain scores. On balance, increases in the percentage of black teachers in a school were associated, other things held constant, with lower gain scores for white students at both the elementary and secondary level and higher gain scores for black students at the secondary level. The latter results do not always hold up, however, once one attempts to control for the endogeneity of teacher and school characteristics.

A number of simulations that are summarized in Table 6 provide the reader with estimates of the quantitative importance of these characteristics in the weighted least squares analyses. These simulations make use of the estimates that were obtained separately for black and white students and, thus, that allowed the influence of teachers' characteristics to vary with the race of the teachers (Tables 3 and 4). To given the reader a sense of the magnitudes that follow, we note that the results in column 1 of Table 2 suggest that reducing class size by 10 students per teacher for elementary school students, would be associated with roughly a 0.8 point increase in the students' gain scores.

The mean percentages of black teachers in the sample were 4 (3) for white elementary (secondary) school students and 71 (77) for black elementary (secondary) school students, respectively. The first three simulations ask what the impact would have been on students if the percentage of black teachers had been 10 percentage points higher? Those reported in row 1 hold constant the mean values of black and white teachers' other characteristics (MA, EXP, VERB) at their sample values. However, since the mean values of black and white teachers' characteristics differed, especially for VERB, this first simulation provides no information on whether teacher skin color *per se* would matter if other teacher characteristics were the same. The simulations reported in rows 2 and 3 address this issue. The former assumes that all teachers have the mean sample value of black teachers' characteristics, while the latter assumes that all teachers have the mean sample value of white teachers' characteristics

In fact, the three sets of simulations yield quite similar findings. Depending upon the particular

Table 6. Estimated changes in gain scores from changes in teacher race and verbal ability

	Grades 3 to 6						Grades 9 to 12	
	White students			Black students			White	Black
Simulation	(1W)	(3W)	(5W)	(1B)	(3B)	(5B)	(1W)	(1B)
(1) Increase the percentage of black teachers by 10 (holding constant all other teacher characteristics)	-1.20[a]	-1.09[a]	0.79[a]	0.27[b]	-0.23	0.36[b]	-0.95[b]	0.21
(2) Increase the percentage of black teachers by 10 (assuming black and white teachers both had the mean characteristics of black teachers)	-1.12[a]	-1.01[a]	-0.74[a]	-0.24	-0.25	-0.47[b]	0.93[b]	0.37
(3) Increase the percentage of black teachers by 10 (assuming black and white teachers both had the mean characteristics of white teachers)	-0.98[a]	-0.90[a]	-0.66[b]	-0.08	-0.04	-0.20	-1.22[a]	0.25
(4) Increase the verbal aptitude scores of black teachers by 10 points	0.15[a]	0.14[a]	0.12[a]	0.95	0.97[a]	0.80[a]	-0.08[a]	0.04
(5) Increase the verbal aptitude scores of white teachers by 10 points	0.45[b]	0.50[a]	0.24	0.06	-0.04	-0.20	0.09[a]	0.04

[a]Estimated change is statistically significant different from zero at the 0.05 level of significance, two-tail test
[b]Estimated change is statistically significantly different from zero at the 0.10 level of significance, two-tail test
Source: Authors' calculations from data in Table 1 and coefficient estimates in the indicated columns from Table 3 (for grades 3 to 6) and Table 4 (for grades 9 to 12).

elementary school equations used in the simulation, increasing the percentage of black teachers by 10 percentage points is estimated to *reduce* the synthetic gain scores of white elementary school students by roughly 0.7 to 1.2 points and to *reduce* the gain scores of black elementary school students by roughly 0 to 0.5 points. This should be contrasted to mean gain scores of 5.86 and −4.42 for the two groups, respectively. At the high school level, the increase would *reduce* the synthetic gain scores of white students by roughly 0.9 to 1.2 points, but *increase* the gain scores of black students by roughly 0.2 to 0.4 points. This should be contrasted to mean gain scores of 0.86 and −0.53, respectively.

The mean verbal aptitude test scores of white elementary (high school) teachers and black elementary (high school) teachers were 81.13 (81.17) and 73.54 (80.20) respectively for white students and 78.76 (76.62) and 64.03 (66.91) respectively for black students. The simulations reported in the fourth and fifth rows of the table ask what the impact on the synthetic gain scores would have been if all black teachers' verbal aptitude scores were increased by 10 points (row 4) and if all white teachers' scores were increased by 10 points (row 5). The standard deviation of teachers' verbal aptitude scores is approximately 20 points for each of the white and black teacher samples, so a 10-point increase represents a change of about one-half of a standard deviation. Given that white teachers taught primarily white students and black teachers taught primarily black students in the 1960s, one should expect that improving only the verbal test scores of teachers of one race would influence primarily the gain scores of students of that race. This in fact occurs.

Improving the verbal aptitude scores of black teachers by 10 points is estimated to increase the synthetic gain scores of white elementary school students by roughly 0.14 points and of black elementary school students by roughly 0.9 points. Similarly increasing the verbal aptitude scores of white elementary school teachers by 10 points is estimated to increase the synthetic gain scores of white elementary students by between 0.24 and 0.50 points, but to have very little, or even a small negative effect, on black elementary school students' scores. At the high school level, neither white nor black students gains scores are predicted to change very much in response to either a change.

VI. CONCLUDING REMARKS

What conclusions should one come away with from our findings? Teacher race and verbal aptitude did matter in the 1960s in the sense that they were associated with synthetic gain scores! Verbal aptitude scores of teachers nationwide have declined substantially during the last two decades.[41] If one adds a quadratic term in ability to our gain score equation, one finds that the marginal affect of ability increases as ability declines, so tht the payoff to improving teachers' verbal scores probably is even higher today.[42]

We must caution, however, that our results are for synthetic gain scores in the mid-1960s. The mid-1960s was a period of change for schools in the United States and this may reduce the appropriateness of our synthetic gain score approach, which assumes that the same "fixed" unobservable variables identically affected the test scores of students in different grades at a school at a point in time. Furthermore, synthetic gain scores, especially for the high school data, are dependent on drop-our rates. Other things equal, assuming the drop-outs come from the lower tail of the test score distribution, the higher the drop-out rate in a school between the ninth and twelfth grades, the higher the twelfth grade score will be and thus the larger the gain score will be. Put another way, our results may be subject to a form of selection bias. While no data exist on drop-out rates in the *EEO*, results we report elsewhere based on analyses of data from *High School and Beyond* suggest that teacher race and a proxy for teacher ability do not influence a student's drop-out probability, so perhaps this is not a problem.[43]

Changes in student test scores over their school careers are not the sole outcome of interest. Do teacher verbal aptitude and race influence college-going behavior, college completion rates, or post-educational labor market outcomes? Do they influence noneconomic outcomes such as attitudes students hold towards individuals from other racial groups? These questions cannot be answered with the *EEO* data, although one recent study on a related topic did suggest that school quality measures do affect labor market outcomes.[44]

The racial and ethnic distributions of students and teachers are different today than they were 30 years ago when the *EEO* survey was undertaken. So, perhaps, are the attitudes about and expectations of

black and white students and teachers towards members of the other race. Although our reanalyses of the *EEO* data found in some specifications that increasing the percentage of black teachers in a school, other things held constant, was associated with lower white student gain scores at both the elementary and secondary levels and higher black student gain scores only at the high school level, these findings were sensitive to assumptions about the endogeneity of teacher and school characteristics. Furthermore, estimated relationships obtained from "educational production function" analyses are often not the same across studies and there is nothing that guarantees that these relationships would hold today.[45] For example, one recent study of the actual gain scores between the second and sixth grades on reading and vocabulary tests for low-income black students in Gary, Indiana found that, holding other characteristics of teachers constant, black elementary school teachers did enhance the performance of these black students.[46] Clearly, before drawing any policy conclusions about the importance of teacher skin color, *per se*, it is necessary to replicate our analyses using more recent data.[47]

Conceptually, however, the issues we have raised should not be ignored. Minority teachers may, on average, improve the academic performance of black students but adversely influence the academic performance of white students today. Teacher verbal aptitude may matter today, on average, both for white and minority teachers. To the extent that the latter tend to have lower test scores, hiring minority teachers with lower verbal aptitude scores than white teachers may adversely affect the gain scores of both minority and white students. The case for expanding the number of minority teachers in public education rests on distributative as well as efficiency considerations. However, contemporary empirical evidence of the type we have presented surely should be part of the policy debate.

Acknowledgements — Our research was supported by the William H. Donner Foundation and the Finance Center of the Consortium for Policy Research in Education (CPRE). CPRE is a consortium of the University of Southern California, Rutgers University, Cornell University, Harvard University, Michigan State University, and the University of Wisconsin-Madison, and is funded by grant no. R1178G10039 from the U.S. Department of Education, Office of Educational Research and Improvement. The views expressed here are solely the authors' and do not reflect the views of Cornell, the NBER, the Rand Corporation, the William H. Donner Foundation, CPRE, or the U.S. Department of Education. Without implicating them for what remains, we are grateful to numerous colleagues at Cornell, the NBER and around the nation for their comments on earlier drafts. The original *Coleman* data set has been deposited at the Interuniversity Consortium for Political and Social Research (ICPSR).

NOTES

1. On the declining pool of potential minority teachers see Cole (1986), Irvine (1988), Berger (1990), and Zinn (1990). For evidence that minority teachers are more likely to fail the National Teachers Examination (NTE), and, in Texas, the Texas Educational Assessment of Minimum Skills Examination, see Cole (1986), Ferguson (1991a), and Thernstrom (1991).
2. See, for example, Irvine (1988).
3. On the potential cultural bias of tests and the validity of the NTE see National Research Council (1989), Ayers and Qualls (1979), Ayers (1988), Haney *et al.* (1987), Darling-Hammond and Wise (1983), and Sheehan and Marcus (1978). Studies that show that teachers' "ability", as measured by test scores, does affect student academic achievement include Armor (1972), Boardman, *et al.* (1978), Coleman *et al.* (1966), Ferguson (1991a; 1991b), Sheehan and Marcus (1977), and Strauss and Sawyer (1986). In contrast Summers and Wolfe (1977) find no evidence that teachers' "ability" matters.
4. These studies include Aaron and Powell (1982), Aloia, Maxwell and Aloia (1981), Banks (1988), Barnes (1979), Baron (1985), Beady and Hansell (1981), Braun (1976), Brophy (1981), Brown *et al.* (1970), Byalick and Bersoff (1979), Carew and Lightfoot (1979), Coates (1972), Cooper *et al.* (1975), Cooper and Tom (1984), Cornbleth and Korth (1990), Dusek and Joseph (1983), Eaves (1975), Feldman (1986), Gottlieb (1964), Haller (1985), Heath (1971), Holiday (1985), Irvine (1985; 1986; 1990), Jackson and Cosca (1974), Leiter (1976), Mathis (1976), Meier *et al.* (1989), Natriello and Dornbusch (1983), Simpson and Erickson (1983), Sizemore (1981), Tobias *et al.* (1983), and West and Anderson (1976).
5. See, for example, Alexander *et al.* (1987), Bridge *et al.* (1979), Crain and Mahard (1978), Crain *et al.* (1982), Cunningham (1975), Darkenwald (1975), Farkas *et al.* (1990), Ferguson (1991a; 1991b),

Glick (1971), Maynor (1970), Murnane (1975), Ohberg (1972), Pascarella *et al.* (1979) Rossel and Hawley (1983), Sanders (1982), Sheehan and Marcus (1977), St John (1971), Touliatos *et al.* (1977), and Yando *et al.* (1971).

6. See, for example, Bowles and Levin (1968), Cain and Watts (1970) and the set of papers published in Mosteller and Moynihan (1972).
7. See, for example, Coleman *et al.* (1966), Armour (1972), Hanushek (1972), Jencks (1972). Thernstrom (1991) has recently reminded people of this finding.
8. See Hanushek (1986).
9. Students in each grade were administered a battery of subject and aptitude tests. For each subject, different tests were administered to students in each grade so that one cannot infer anything about the absolute amount students in a school learn by comparing say, the mean third grade and mean sixth grade test scores. However, one *can* infer something about how much students in a school were learning in relative terms by comparing gain scores across schools. So, for example, if the mean third grade score on a test was 80% in each of two schools and the mean sixth grade scores were 80% and 90% in the two schools, respectively, the implication is that the students learn more between the third and sixth grade in the second school.
10. A major finding of the *Coleman Report* was how segregated by race schools were in 1966. For example, almost 80% of all white pupils in the first and twelfth grades attended schools that were 90 to 100% white, while more than 65% (85%) of black students in the first (twelfth) grade attended schools that were between 90 and 100% black (Coleman *et al.*, 1966, p. 3). In the research that follows, gain scores are computed by race for a school if at least one student of that race are present in the base year grade. Since weighted least square analyses are used (with the weights based on the number of students in the racial group in the grades in the school), schools with only a few students from a racial group are given very little weight in the race-specific analyses.
11. While only a small fraction of white students were taught by black teachers in 1966, a greater fraction of black students were educated by white teachers. For the nation as a whole, the average black elementary (secondary) student attended a school in which 35% (41%) of the teachers were white (Coleman *et al.*, 1966, p. 3).
12. Prior analyses of how teachers are sorted across school districts and how they decide whether to remain in the profession have been undertaken. See, for example, Antos and Rosen (1975), Ferguson (1991b), and Murnane and Olsen (1990). However, Ehrenberg and Brewer (1994) is the first study to treat teacher characteristics as endogenous in the estimation of educational production functions.
13. The *EEO* data tapes are very poorly documented and considerable effort had to be expended by us to "clean" the data. A data appendix, available from us on request, discusses a number of the problems we faced and the actions we took.
14. Although over 60,000 teachers were surveyed in the original *EEO* survey, the data set we received from the National Archive contained information on only 44,193 teachers and came with a notice that two teacher files were missing. When teacher data was missing, it appeared to be missing for all the teachers in a school and all of these schools are necessarily excluded from our analyses. This restriction alone reduced the total number of schools in the sample from about 4000 to 2075. This implies that most of the schools that failed to report teacher data were relatively small.
15. Any school that enrolled both white and black students, will appear in both the white student and the black student samples. Test scores for students of other ethnicities/races, primarily Native Americans, Hispanics, and Asian Americans, are used to compute the mean test scores in the overall sample, but not in the black or white samples.
16. Part of the reason for the small high school sample is that many ninth graders attended junior high schools and thus were not enrolled in schools with twelfth grades. A comparison of the descriptive statistics for the samples actually used in our analyses (Table 1) and the largest samples available if we ignore our need for teacher characteristics data, suggests that the two samples were quite similar, for both elementary and secondary schools, in terms of their base year test scores, synthetic gain scores, and proportions of black teachers and students.
17. The high percentages of black students and teachers in our sample occur because black schools were over-sampled in the original *EEO* survey. No sample weights were found by us in the data so we could not control for problems relating to choice-based sampling in our statistical analyses.
18. Gain scores can be negative because, as discussed in footnote 9, the tests administered in each grade were different.
19. Requiring black teachers to have more education than white teachers at schools with white students may reflect either discrimination in hiring or a compensating differential for their lower test scores.
20. The mean values of the students' characteristics in a school were very similar for third and sixth grade students in each school and the latter were used in these analyses. Family income data were actually not available in the *EEO* data and had to be estimated by us. We computed the occupational distribution (at the one-digit level) of fathers in each school and then used this distribution and data

on the 1970 median male earnings in each occupation in the census region (4) in which the school was located to obtain an estimate of family incomes in each school.

21. A 10 percentage point increase in teachers' verbal aptitude scores is equivalent to a change of a little under one standard deviation in the overall distribution of teachers' scores.
22. Appendix Table A1 contains mean test score level equations for the four elementary and secondary grades that are specified similarly to column 1 of Table 2. While these test level equations suffer from omitted variable bias (see the introduction), it is clear that many of the *measured* variables that influence the synthetic gain scores also influence the test score levels.
23. An alternative way to handle the problem is to treat BYTEST as endogenous, obtain an instrumental variable estimate for it, and then reestimate the gain score equation in column (1) using the instrumental variable. We obtained an instrument by regressing BYTEST on the characteristics of the school's students, their families and the community in which the school was located (the X_i) and similar variables for the larger county or SMSA in which the school was located. When the gain score equation was reestimated using this instrument, the coefficient of the instrument proved to be insignificant but the coefficients of BLACKT (−0.063) and VERB (0.066) were very similar to the corresponding coefficients found in column 1 of Table 2 and both remained statistically significant. Consequently, in what follows, we treat BYTEST as exogenous.
24. In some schools, either all students failed to report whether they spent their entire careers in the school or all students reported that they had not. Hence, the smaller sample sizes in columns (3) through (5).
25. For the white students sample, the computed statistic is $F(3,776) = 5.87$ and for the black student sample it is $F(3,491) = 7.93$. Both of these values exceed the 0.99 critical values of $F(3,120)$ or $F(3,\infty)$ of 3.95 and of 3.78, respectively.
26. One cannot infer the effect of black teachers *per se* in these models from the coefficients of BLACKT alone since this variable also interacts with the other teacher characteristics (see equation (2)). We conduct simulations, however, using all of these characteristics and their coefficients in Section V.
27. Tables of the results of these analyses are available from the authors.
28. We are grateful to Fran Blau for suggesting this point to us.
29. In each case, the mean third grade value of the variable (logarithm of mean test score or percentile rank on the third grade test) appears on the right hand side of the equation.
30. See Tiebout (1956).
31. Again, see Antos and Rosen (1975) and Ferguson (1991b).
32. Evans, Oates, and Schwab (1992) make a related point in the context of estimating "peer group effects" on drop-out rates. While they treat peer group measures, such as the percentage of disadvantaged students in a school as endogenous, they do not explore the influence of teacher or school characteristics.
33. Ehrenberg and Brewer (1994) is the exception. Ferguson (1991b) treats teacher and school characteristics as exogenous when he estimates district-level educational production functions, but then goes on to show how these characteristics vary with underlying socioeconomic and demographic variables.
34. We are grateful to Marshall Smith, currently Undersecretary at the U.S. Department of Education, for helping us to develop an algorithm to identify the county or SMSA in which each school in the *EEO* survey was located. The data appendix provides details.
35. For each district, we regressed the logarithm of teacher's salary on the teacher's experience, and degree level, using all teachers we observed in the district as observations. The estimated intercepts from these equations are estimates of each district's logarithm of starting salary.
36. See Russell Davidson and James M. MacKinnon (1993), pp. 237–242 for a more formal treatment.
37. The computed F statistic was $F(6,681) = 6.56$ which exceeds the critical value of roughly 2.80 for rejecting the hypothesis at the 0.99 level.
38. The computed F statistic was $F(4,683) = 1.44$ which is less than the critical value of roughly 3.78 for rejecting the hypothesis at the 0.99 level.
39. See Antos and Rosen (1975) and Ferguson (1991b).
40. Two other extensions warrant brief reporting here. First, when one eliminates the estimated starting salary in a district from the instrumental variable equations (on the grounds that it is endogenous), one obtains virtually identical results to those reported in Table 6 and the text. Second, when one repeats the analyses eliminating the base year test score from the gain score equations (on the grounds that it too may be correlated with unobserved variables), one again obtains a similar pattern of results. That is, Durbin-Hausman-Wu tests suggest that only BOOKS, PUPT, and MA should be treated as endogenous and the estimated VERB and BLACKT coefficients obtained from such a specification are very similar to those obtained in column 2 of Table 2.
41. See Murnane *et al.* (1991), Chapter 2.

42. When VERB squared is added to the model specified in column 1 of Table 2 and then the expanded equation reestimated, the coefficient of VERB and VERB squared, respectively become, with the absolute values of their *t* statistics in parentheses, 0.366 (2.5) and −0.002 (1.9). Although the marginal effect of VERB declines as aptitude increases in this model, it remains positive until VERB reaches 91.5.
43. Ehrenberg and Brewer (1994).
44. See Card and Krueger (1992).
45. See Hanushek (1986).
46. Hanushek (1992).
47. The data used by Ferguson (1991) on Texas school districts would be extraordinarily valuable if characteristics of teachers and students by race/ethnic group for each district could be made available.

REFERENCES

AARON, R. and POWELL, G. (1982) Feedback practices as a function of teacher and pupil race during reading group instruction. *The Journal of Negro Education* **51**(1), 50–59.

ALEXANDER, K.L., ENTWISLE, D.R. and THOMPSON, M. (1987) School performance, status relations, and structure of sentiment: bringing teacher back in. *American Sociological Review* **52**, 665–682.

Aloia, G., Maxwell, J.A. and ALOIA, S.D. (1981) Influence of a child's race and the EMR label on initial impressions of regular classroom teachers. *American Journal of Mental Deficiency* **85**(6), 619–623.

ANTOS, J.R. and ROSEN, S. (1975) Discrimination in the market for public school teachers. *Journal of Econometrics* **3**, 123–150.

ARMOR, D.J. (1972) School and family effects on black and white achievement: a reexamination of the USOE data. In *On Equality of Educational Opportunity* (Edited by MOSTELLER and MOYNIHAN). New York: Random House.

AYERS, J.B. and GELNDA S. QUALLS (1979) Concurrent and predictive validity of the national teacher examination. *Journal of Educational Research* **73**(2), 86–92.

AYERS, J.B. (1988) Another look at the concurrent and predictive validity of the national teacher examinations. *Journal of Educational Research* **81**(3), 133–137.

BANKS, J.A. (1988) Ethnicity, class, cognitive, and motivational styles: research and teaching implications. *Journal of Negro Education* **57**(4), 452–466.

BARNES, W.J. (1979) Teachers in desegregated high schools in Texas. *Integrated Education* **97**, 25–26.

BARON, R.M., TOM, D.Y.H. and COOPER, H.M. (1985) Social class, race and teacher expectations. In *Teacher Expectancies* (Edited by DUSEK, J.B.). Hillsdale, N.J.: L. Erlbaum.

BEADY, C.J. JR and HANSELL, S. (1981) Teacher race and expectations for student achievement. *American Educational Research Journal* **1**(2), 191–206.

BERGER, J. (1990) Pessimism in air as schools try affirmative action. *New York Times* **27**, B1.

BOARDMAN, A.E., HOROWITZ, D.A. and LLOYD, A. (1978) Variables affecting the learning of inner city children. In *Urban Schools in Urban Systems, Selected Papers* (Edited by GAPPERT, G.). Philadelphia: Research far better schools, Inc.

BOWLES, S. and LEVIN, H. (1968) The determinants of scholastic achievement: an appraisal of some recent evidence. *Journal of Human Resources* **III**, 3–24.

BRAUN, C. (1976) Teachers expectations: sociopsychological dynamics. *Review of Educational Research* **46**(2), 185–213.

Bridge, R.G., Judd, C.M. and MOOCK, P.R. (1979) *The Determinants of Educational Outcomes: The Impact of Families, Peers, Teachers, and Schools*. Cambridge, MA: Ballinger Publishing Company.

BROPHY, J. (1985) Interactions of male and female students with male and female teachers. In *Gender Influences in Classroom Interaction* (Edited by WILKINSON, L.C. and MARRETT, C.B.). New York: Academic Press, Inc., 115–142.

BROWN, W.E., PAYNE, T.L., LANKEWICH, C. and CORNETT, L.L. (1970) Praise, criticism and race. *The Elementary School Journal* **70**, 373–377.

BYALICK, R. and BERSOFF, D.N. (1979) Reinforcement practices of black and white teachers in integrated classrooms. *Journal of Educational Psychology* **66**(4), 473–480.

CAIN, G. and WATTS, H. (1970) Problems in making policy inferences from the Coleman Report. *American Sociological Review* **35**, 228–242.

CARD, D. and KRUEGER, A.B. (1992) Does school quality matter? Returns to education and characteristics of public schools in the United States. *Journal of Political Economy* **100**, 1–40.

CAREW, J.V. and LIGHTFOOT, S.L. (1979) *Beyond Bias*. Cambridge, MA: Harvard University Press, 69–101

COATES, B. (1972) White adult behavior toward black and white children. *Child Development* **43**, 143–154.

COLE, B. (1986) The black educator: an endangered species. *Journal of Negro Education* **55**(3), 326–335.

COLEMAN, J.S., *et al.* (1966) *Equality of Educational Opportunities*. Washington, DC: U.S. Office of Education.

COOPER, H.M. and TOM, Y.H. (1984) Teacher expectation research: a review with implications for classroom instruction. *The Elementary School Journal* **85**(1), 77–89.

COOPER, C.C. (1988) Implications of the absence of black teachers/ administrators on black youth. *Journal of Negro Education* **57**(2), 123–124.

COOPER, H., BARON, R. and LOWE, C.A. (1975) The importance of race and social class information of expectancies about academic performance. *Journal of Educational Psychology* **67**(2), 312–319.

CORNBLETH, C. and KORTH, W. (1990) Teacher perceptions and teacher–student interaction in integrated classrooms. *Journal of Experimental Education* **58**, 259–263.

CRAIN, R. and MAHARD, R. (1978, 1978b) Desegregation and black achievement: a review of the research. *Law and Contemporary Problems* **42**, 17–56.

CUNNINGHAM, W.G. (1975) The impact of student–teacher pairings on teacher effectiveness. *American Educational Research Association* **12**(2), 169–189.

DARKENWALD, G.G. (1975) Some effects of the 'Obvious Variable': teacher's race and holding power with black adult students. *Sociology of Education* **48**, 420–431.

DARLING-HAMMOND, L. and WISE, A.E. (1983) Teaching standards, or standardized teaching? *Educational Leadership* **41**, 66–69.

DAVIDSON, R. and MACKINNON, J.G. (1993) *Estimation and Inference in Econometrics*. New York: Oxford University Press.

DUSEK, J.B. and JOSEPH, G. (1983) The bases of teacher expectancies: a meta-analysis. *Journal of Educational Psychology* **75**(3), 327–347.

EAVES, R.C. (1975) Teacher race, student race, and the behaviour problem checklist. *Journal of Abnormal Child Psychology* **3**(1), 1–9.

EHRENBERG, R.G. and BREWER, D.J. (1994) Do school and teacher characteristics matter?: Evidence from high school and beyond. *Economics of Education Review* **13**(1), 1–17.

EVANS, W.N., OATES, W.E. and SCHWAB, R.M. (1992) Measuring peer group effects: a study of teenage behavior. *Journal of Political Economy* **100**, 966–991.

FARKAS, G., GROBE, R., SHEEHAN, D. and Shuan, Y. (1990) Cultural resources and school success. gender, ethnicity, and poverty groups within an urban school district. *American Sociological Review* **55**, 127–142.

FELDM[illegible], R.S. (Ed.) (1986) *The Social Psychology of Education: Current Research and Theory*. New York. Cambridge University Press.

FERGUSON, R.F. (1991a) Racial patterns in how school and teacher quality affect achievement and earnings. *Challenge: A Journal of Research on Black Men* **12**.

FERGUSON, R.F. (1991b) Paying for public education: new evidence on how and why money matters *Harvard Journal on Legislation* **28**, 465–488.

GLICK, I. (1971) Does teacher's skin color matter? *Integrated Education* **9**, 26–30.

Gottlieb, D. (1964) Teaching and students: the views of negro and white teachers. *Sociology of Education* **37**, 345–353.

HALLER, E. (1985) Pupil race and elementary school ability grouping: are teachers biased against black children? *American Educational Research Journal* **22**(4), 465–483.

HANEY, W., MADAUS, G. and KRETZER, A. (1987) Charms talismanic: testing teachers for the improvement of American education. *Review of Research in Education* **14**, 169–238.

HANUSHEK, E. (1986) The economics of schooling. *Journal of Economic Literature* **24**, 1141–1176.

Hanushek, E. (1972) *Education and Race: An Analysis of the Educational Production Process* Lexington, MA: Lexington Books.

HANUSHEK, E. (1992) The trade-off between child quality and quantity. *Journal of Political Economy* **100**, 84–118.

HEATH, R.W. (1971) The ability of white teachers to relate to black students and to white students. *American Educational Research Journal* **8**(1), 1–10.

HOLIDAY, B.G. (1985) Differential effects of children's self-perceptions and teachers' perceptions on black children's academic achievement. *Journal of Negro Education* **54**(1), 71–81.

IRVINE, J.J. (1985) Teacher communication patterns as related to the race and sex of the student. *Journal of Educational Research* **78**, 338–345.

IRVINE, J.J. (1988) An analysis of the problem of disappearing black educators. *The Elementary School Journal* **88**(5), 503–513.

IRVINE, J.J. (1986) Teacher–student interactions: effects of student race, sex, and grade level. *Journal of Educational Psychology* **78**(1), 14–21.

JACKSON, G. and COSCA, C. (1974) The inequality of educational opportunity in the southwest: an observational study of ethnically mixed classrooms. *American Educational Research Journal* **11**(3), 219–229.

JENCKS, C.P. (1972) The Coleman Report and the conventional wisdom. In *On Equality of Educational Opportunity* (Edited by MOSTELLER and MOYNIHAN). New York: Random House.

LEITER, K.C.W. (1976) Teachers' use of background knowledge to interpret test scores. *Sociology of Education* **49**, 59–65.

MATHIS, D.W. (1976) Differences in teacher interaction with Afro-American and Anglo-American students in the same classroom. *Dissertation Abstracts International — Humanities and Social Sciences* **36**(9), 5950A.

MAYNOR, W. (1970) Academic performance and school integration: a multi-ethnic analysis. Durham, NC: Duke University, Doctoral Dissertation.

MEIER, K.J., STEWART, J., JR and ENGLAND, R.E. (1989) *Race, Class, and Education: The Politics of Second-Generation Discrimination*. Madison, WI: University of Wisconsin Press.

MOSTELLER, F. and MOYNIHAM, D.P. (Eds) (1972) *On Equality of Educational Opportunity*. New York: Random House.

MURNANE, R.J. (1975) *The Impact of School Resources on the Learning of Inner City Children*. Cambridge, MA: Ballinger Publishing Co.

MURNANE, R.J. and OLSEN, R. (1990) The effects of salaries and opportunity costs on length of stay in teaching. *Journal of Human Resources* **25**, 106–124.

Murnane, R.J. *et al.* (1991) *Who Will Teach? Policies That Matter*. Cambridge, MA: Harvard University Press.

NATIONAL RESEARCH COUNCIL (1989) *Fairness in Employment Testing*. Washington, DC: National Academy Press.

NATRIELLO, G. and DORNBUSCH, S.M. (1983) Bringing behavior back in: the effects of student characteristics and behavior on the classroom behavior for teachers. *American Educational Research Journal* **20**(1), 29–43.

OHLBERG, H. (1972) Does the black child need a black teacher? *Integrated Education* **10**, 27–28.

PASCARELLA, E.T., TALMAGE, H. and PINZUR, S. (1979) Summative analysis of longitudinal student achievement data in a desegregated school district, presented at the American Educational Research Association Conference, San Francisco.

ROSSELL, C.H. and HAWLEY, W.D. (1983) *The Consequences of School Desegregation*. Philadelphia, PA: Temple University Press.

SANDERS, J.M. (1982) Faculty desegregation and student achievement. Santa Barbara, CA: University of California, Department of Sociology.

SHEEHAN, D.S. and MARCUS, M. (1978) Teacher performance on the national teacher examinations and student mathematics and vocabulary achievement. *Journal of Educational Research* **71**.

SHEEHAN, D.S. and MARCUS, M. (1977) The effects of teacher race and student race on vocabulary and mathematics achievement. *The Journal of Educational Research* **70**(3), 123–126.

SIMPSON, A.W. and ERICKSON, M.T. (1983) Teachers' verbal and nonverbal communication patterns as a function of teacher race, student gender, and student race. *American Educational Research Journal* **20**(2), 183–198.

SIZEMORE, B. (1981) Do black and white students look for the same characteristics in teachers? *Journal of Negro Education* **50**(1), 48–53.

STRAUSS, R.P. and SAWYER, E.A. (1986) New evidence on teacher and student competencies. *Economics of Education Review* **41**, 41–48.

SUMMERS, A.A. and WOLFE, B.L. (1977) Do schools make a difference? *American Economic Review* **67**(4), 639–652.

THERNSTROM, A. (1991) Beyond the pale. *The New Republic* 22–24.

TIEBOUT, C.M. (1956) A pure theory of local expenditure. *Journal of Political Economy* **64**, 416–424.

TOBIAS, S., ZIBRIN, M. and MUSFELT, C. (1983) Special education referrals: failure to replicate student–teacher ethnicity interaction. *Journal of Educational Psychology* **75**(5), 705–707.

TOLLEFSON, J., LINDHOLM, B. and RICE, A. (1977) Interaction of race with other variables on achievement in school. *Psychology in the Schools* **14**(3), 360–363.

WEST, C.K. and ANDERSON, T.H. (1976) The question of preponderant causation in teacher expectancy research. *Review of Educational Research* **46**(4), 613–630.

YANDO, R., ZIGLER, E. and GATES, M. (1971) The influence of negro and white teachers rated as effective or noneffective on the performance of negro and white lower-class children. *Developmental Psychology* **5**(2), 290–299.

ZINN, L. (1990) In search of tomorrow's minority teachers. *Business Week* 120.

Appendix Table A1. School level mean test score equations: grades 3, 6, 9 and 12 (absolute value *t* statistics)[a]

	GRADE 3	GRADE 6	GRADE 9	GRADE 12
INTER	38.998 (2.8)	25.031 (8.3)	17.646 (3.1)	4.119 (1.0)
FEM	−0.018 (0.9)	0.053 (2.6)	−0.094 (3.6)	−0.031 (1.9)
BLACKS	−0.075 (7.3)	−0.095 (8.9)	−0.059 (1.7)	−0.047 (1.9)
FNHH	−0.056 (2.9)	−0.085 (4.2)	−0.221 (5.3)	−0.219 (6.7)
MNHH	−0.071 (2.6)	−0.257 (9.0)	−0.139 (1.8)	−0.138 (2.2)
PHONE	0.077 (5.8)	0.129 (9.2)	0.051 (2.0)	0.077 (3.7)
INCOME	0.059 (0.3)	0.200 (1.0)	1.634 (3.9)	1.511 (4.6)
FLNCH	−0.028 (2.5)	−0.039 (3.2)	0.009 (0.3)	0.002 (0.1)
FED	0.854 (3.7)	1.386 (5.7)	1.440 (2.4)	1.515 (3.2)
MED	0.639 (2.2)	0.606 (2.0)	0.329 (0.5)	0.799 (1.4)
CITY	−3.161 (6.6)	−1.764 (3.6)	−2.074 (2.5)	0.470 (0.8)
RURAL	−0.131 (0.3)	1.039 (2.1)	0.091 (0.1)	0.502 (0.8)
BOOKS	0.003 (0.1)	0.007 (0.2)	0.069 (1.0)	0.064 (1.2)
PUPT	−0.134 (3.9)	−0.121 (3.3)	−0.053 (0.8)	0.074 (1.4)
BLACKT	0.029 (2.6)	−0.028 (2.4)	0.015 (0.5)	−0.011 (0.5)
EXP	0.133 (3.8)	0.111 (3.1)	0.016 (0.2)	0.071 (1.0)
MA	−0.017 (1.5)	−0.013 (1.1)	−0.010 (0.6)	−0.030 (2.4)
VERB	0.048 (2.1)	0.108 (4.5)	0.152 (2.9)	0.165 (4.1)
n	969	969	256	256
$\bar{R}^2$	0.701	0.886	0.881	0.932

[a]All variables are defined as in Table 2.

Appendix Table A2. Estimation of instrumental variable equations[a] (absolute value of *t* statistics)

	BOOKS	PUPT	BLACKT	EXP	MA	VERB
INTER	−8.139 (0.5)	56.248 (3.2)	1.392 (2.5)	−16.275 (1.1)	−57.792 (1.2)	3.355 (0.1)
FEM	0.029 (1.8)	−0.021 (1.2)	0.000 (0.07)	−0.017 (1.1)	0.054 (1.1)	0.004 (0.1)
BLACKS	−0.022 (4.0)	0.027 (4.4)	0.008 (38.6)	−0.010 (2.1)	0.027 (1.6)	−0.118 (13.2)
FNHH	−0.005 (0.3)	−0.018 (1.0)	0.002 (2.7)	0.033 (2.2)	0.104 (2.1)	0.047 (1.8)
MNHH	−0.029 (1.4)	0.010 (0.5)	0.003 (3.6)	−0.355 (2.0)	−0.198 (3.3)	−0.158 (5.9)
PHONE	−0.007 (0.6)	−0.006 (0.4)	−0.091 (3.3)	0.013 (1.2)	0.045 (1.3)	0.084 (4.4)
INCOME[a]	−0.369 (1.8)	0.276 (1.3)	0.030 (4.5)	−0.175 (1.0)	1.109 (2.0)	−0.402 (1.3)
FLNCH	0.006 (0.5)	−0.023 (2.2)	−0.001 (1.3)	−0.017 (1.9)	0.021 (0.7)	0.020 (1.3)
FED	0.421 (2.2)	−0.260 (1.0)	0.002 (0.3)	−0.198 (1.1)	−0.561 (1.0)	0.193 (0.6)
MED	0.194 (0.7)	−0.261 (1.0)	−0.015 (1.7)	0.381 (1.7)	1.284 (1.7)	0.527 (1.3)
CITY	−0.922 (2.1)	0.991 (2.1)	−0.120 (8.1)	−0.707 (1.9)	0.403 (0.3)	3.325 (5.0)
RURAL	0.016 (0.3)	−0.921 (2.1)	0.007 (0.5)	−0.538 (1.5)	0.382 (0.3)	−0.146 (0.2)
ESAL	2.066 (1.8)	−3.880 (1.8)	−0.146 (2.2)	3.824 (2.3)	7.335 (1.3)	6.708 (2.2)
POP[b]	0.003 (2.3)	−0.006 (3.2)	−0.000 (1.8)	−0.001 (0.6)	−0.006 (1.4)	−0.008 (3.1)
POPDEN	0.000 (1.3)	−0.001 (2.1)	−0.000 (6.1)	0.000 (0.5)	0.005 (4.9)	0.001 (2.2)
PURB	−0.004 (0.3)	0.013 (1.0)	0.001 (2.9)	−0.004 (0.4)	0.104 (2.7)	−0.062 (3.0)
PBLK	0.003 (0.2)	0.051 (3.4)	0.002 (5.0)	−0.004 (0.3)	−0.074 (1.8)	−0.095 (4.3)
P650	−0.357 (3.7)	−0.088 (0.9)	−0.010 (3.0)	0.191 (2.3)	−0.031 (0.1)	0.346 (2.3)
MEDU	−0.541 (1.5)	0.969 (2.5)	−0.013 (1.0)	0.514 (1.6)	−0.415 (0.3)	0.341 (0.5)
PGHS	0.058 (0.9)	−0.103 (1.6)	0.002 (0.8)	−0.124 (2.2)	0.010 (0.1)	−0.070 (0.7)
SEN[a]	−0.011 (2.0)	0.002 (3.6)	0.000 (2.0)	0.009 (1.9)	0.043 (2.8)	0.027 (3.2)
PWHC	0.072 (1.6)	−0.088 (1.8)	0.003 (1.8)	0.056 (1.3)	0.164 (1.2)	0.008 (0.1)
MINC[a]	−0.686 (2.2)	0.858 (2.5)	−0.056 (5.2)	−0.737 (2.7)	5.227 (5.7)	1.629 (3.3)
PUBAS[a]	−0.033 (2.3)	0.027 (1.8)	0.001 (1.9)	−0.034 (2.7)	−0.093 (2.2)	0.042 (1.9)
PSEE	0.724 (0.2)	−4.468 (1.5)	0.251 (2.7)	−3.883 (1.6)	6.749 (0.9)	−3.356 (0.8)
n	889	888	1129	1129	1129	1129
$\bar{R}^2$	0.148	0.207	0.870	0.134	0.156	0.644

ESAL estimated starting salary of teachers in the district
POP area population in 1960

POPDEN	area population per square mile in 1960
PURB	per cent area population living in urban areas in 1960
PBLK	per cent area population that is black in 1960
P650	per cent area population that is age 65 or older
MEDU	median school years completed for area population that is age 25 or older in 1960
PGHS	per cent area population 25 or older in 1960 that completed high school
SEN	total area school enrollment in 1960
PWHC	per cent area adults in white collar jobs in 1960
MINC	median family income in the area in 1960
PUBAS	per cent area families on public assistance in 1964
PSEE	per student expenditures in the area in 1960

All other variables are defined in Table 2.

'Weighted least squares regressions. The weights used in each case are the number of teachers taking the test.

Sources: (a) Authors' computations from the *Equality of Educational Opportunity* survey data tapes. (b) Authors' computations from the *1965 City and County Databook* data tape. The "area" refers to the SMSA in which the school is located, if the school is in SMSA and the county in which the school is located for schools outside SMSAs.