NBER WORKING PAPER SERIES

DO TEACHERS' RACE, GENDER, AND ETHNICITY MATTER?: EVIDENCE FROM NELS88

Ronald G. Ehrenberg
Daniel D. Goldhaber
Dominic J. Brewer

Working Paper No. 4669

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 1994

Our research was supported by grants from the William H. Donner Foundation and the Finance Center of the Consortium for Policy Research in Economics (CPRE). CPRE is a consortium of Cornell University, Harvard University, Michigan State University, Rutgers University, the University of Southern California, and the University of Wisconsin-Madison and is funded by grant No. R1178G10039 from the U.S. Department of Education's Office of Educational Research and Improvement. The views expressed here are strictly our own and do not represent the views of Cornell University, the NBER, the William H. Donner Foundation, CPRE, or the U.S. Department of Education. This paper is part of NBER's research program in Labor Studies. Any opinions expressed are those of the authors and do not necessarily represent those of the National Bureau of Economic Research.

NBER Working Paper #4669
March 1994

DO TEACHERS' RACE, GENDER, AND ETHNICITY MATTER?: EVIDENCE FROM NELS88

ABSTRACT

Our study uses a unique national longitudinal survey, the National Educational Longitudinal Study of 1988 (NELS), which permits researchers to match individual students and teachers, to analyze issues relating to how a teacher's race, gender, and ethnicity, per se, influence students from both the same and different race, gender, and ethnic groups. In contrast to much of the previous literature, we focus both on how teachers subjectively relate to and evaluate their students and on objectively how much their students learn.

On balance, we find that teachers' race, gender, and ethnicity, per se, are much more likely to influence teachers' subjective evaluations of their students than they are to influence how much the students objectively learn. For example, while white female teachers do not appear to be associated with larger increases in test scores for white female students in mathematics and science than white male teachers "produce", white female teachers do have higher subjective evaluations than their white male counterparts of their white female students. We relate our findings to the more general literature on gender, race, and ethnic bias in subjective performance evaluations in the world of work and trace their implications for educational and labor markets.

Ronald G. Ehrenberg
ILR-Cornell University
Ithaca, NY 14853-3901
and NBER

Daniel D. Goldhaber
ILR-Cornell University
Ithaca, NY 14853-3901

Dominic J. Brewer
ILR-Cornell University
Ithaca, NY 14853-3901

I. INTRODUCTION

Why should public school systems aggressively pursue policies to recruit and retain teachers from underrepresented groups? In part, these policies derive from distributional considerations and the desire to provide employment opportunities for members of groups that have been historically discriminated against. More importantly, they are motivated by the poor academic performance and high drop-out rates of many minority students vis-a-vis their white counterparts and the belief by many that teachers from underrepresented minority groups are more effective teachers of minority students. The latter is thought to occur because minority teachers may serve as role models for minority students, interact better with them, have more favorable attitudes and higher expectations for them, and provide more positive feedback to them.[1]

Research on the relative effectiveness of minority teachers in educating minority students has been conducted primarily by sociologists, psychologists, and educational researchers and has focused on teachers' attitudes towards, expectations for, and placement of minority students, as well as the feedback that they provide to the students.[2] Most do not address the students' educational outcomes and also fail to control for other teacher characteristics, such as verbal ability, experience and degree levels.[3] They also fail to ask about the effects that underrepresented minority teachers have on non-minority students.

Females are underrepresented in many scientific and engineering fields at the collegiate level, both as students and faculty.[4] A major reason for this is that by the time females reach the ends of their high school careers, they perform, on average, poorer than males in many science and mathematics classes and on standardized tests.[5] Many decry the absence of female role models in science and mathematics as part of the explanation for this

outcome and call for increased efforts to recruit and retain female high school mathematics and science teachers. However, empirical research on the role that teacher gender actually plays again typically focuses on attitudinal types of measures and not educational outcomes.[6] The few studies that do address outcomes focus on the correlation between teacher gender and students' test scores at a point in time, rather than on the correlation between teacher gender and some value added measure.[7]

Our study uses a unique national longitudinal survey, the National Educational Longitudinal Study of of 1988 (NELS) to analyze issues relating to how a teacher's race, gender, and ethnicity influence students from both the same race, gender, and ethnicity group and students from other groups. In contrast to much of the previous literature, we focus both on how teachers subjectively relate to and evaluate their students and on objectively how much their students learn.

After discussing the unique nature of the NELS data in the next section, section III presents our empirical results. On balance, we find that teachers' race, gender, and ethnicity are much more likely to influence teachers' subjective evaluations of their students than they are to influence how much their students objectively learn. A final section relates our findings to the more general literature on gender, race, and ethnic bias in subjective performance evaluations and traces through the implications of our findings for future research on both education and labor markets.

II. THE NATIONAL EDUCATIONAL LONGITUDINAL STUDY (NELS)

The National Educational Longitudinal Study of 1988 (NELS) is a unique national data base.[8] Its initial survey wave was conducted between February 1 and June 30 of 1988 when students were enrolled in the 8th grade. Data were gathered from students, parents, teachers, and school administrators. For each student, two subject areas were chosen from among English/reading, mathematics, science, and history/social studies. The student was then administered a cognitive test in these two areas and the student's teachers in these two areas surveyed. NELS thus permits one to directly link data on each 8th grade student with data on two of his or her teachers, as well as data from his or her parents and school administrator. The initial wave of NELS included responses from 24,599 students, 22,651 parents, and 5,193 teachers at 1,035 schools.

A second wave of NELS was conducted between January 26 and June 30 of 1990, when the vast majority of the initial cohort of students were enrolled in the 10th grade.[9] Once again, two subjects, typically the same as in 1988, were chosen for each student and the student was administered cognitive tests in these areas. The student's current teacher in these two subjects were again surveyed, as were school administrators; however, parents were not surveyed in 1990. The second wave of NELS included responses from 18,221 students, 15,908 teachers, and 1,291 school administrators.

The cognitive tests administered to the students each year were developed by the Educational Testing Service. The number of questions varied across tests, ranging from 21 in reading comprehension and 25 in science to 30 in history/social studies and 40 in mathematics. While all students who took a subject area test in the 8th grade were

administered the same test, six versions of the tests, which differed in their mathematics and reading difficulty levels, were administered in the 10th grade. Each student's 10th grade tests were determined by his or her scores on the base year mathematics and reading tests. The purpose of the multi-level design of the 10th grade test was to guard against "ceiling" and "floor" effects that might otherwise have occurred.

In the next section, we estimate gain score equations to ascertain whether teachers' race, gender and ethnicity per se, influence how much their students learn. Because different students were administered different tests in the 10th grade, which differed in their degree of difficulty, unadjusted gain scores can not be used for this purpose. Fortunately, these different tests were made comparable by the designers of NELS through the use of Item Response Theory (IRT). IRT is a method that uses the pattern of right, wrong, and omitted responses to the questions actually administered on each test, and the difficulty, discriminating ability, and "guessability" of each question, to place each student, regardless of the test he or she was administered on a continuous scale.[10] The gain scores we actually use in our analyses are the difference between a student's 10th grade test IRT estimated number right and the student's rescaled IRT estimated number right on the 8th grade test.[11]

We also ascertain in this section whether a teacher's subjective evaluation of a student depends upon the match between the student's and the teacher's race, gender, and ethnicity. Teachers in the 10th grade survey were asked a set of questions about their perceptions of each surveyed student in their class. These included whether they thought the student would probably go to college, would recommend the student for academic

honors, believed the student relates well to others, spoke to the students outside of class, and believed the student works hard. These responses (1=yes, 0=no for each) were aggregated by us into a single teacher's subjective evaluation variable and this variable used in our analyses.

III. EMPIRICAL ANALYSES

A) Gain Score Analyses

Our analyses of students' gain scores for each subject area are restricted to white, black, and Hispanic students enrolled in public schools in both the 8th and the 10th grades who took the subject area tests in both years, and for whom teacher characteristics (in both years), school variables (in the 10th grade), and parental survey responses (in the 8th grade) were all present. These restrictions reduced the number of observations in our analyses to 1,776 in history, 2,848 in reading, 3,029 in mathematics, and 2,445 in science.[12]

Table 1a presents the mean values of the 8th grade subject area test scores (number of correct answers) for these students, stratified by race, gender, and ethnicity. Since the number of questions on the tests varied across subject areas, comparisons of absolute scores across tests are not very useful. These data do suggest, however, that white students out-performed other students, on average, on all four tests, and that male students in each racial/ethnic group slightly out-performed female students in the group on the mathematics test and under-performed female students on the reading test.

Table 1b presents the mean adjusted (as described above) gain scores on each test between the 8th and 10th grades for students in each demographic group. These means

range across groups from roughly 1 to 2.8 in science, 1.8 to 2.5 in reading, 2.2 to 3.2 in history, and 4.6 to 5.1 in mathematics. These mean gain scores should be kept in mind when one evaluates the estimates of the importance of teacher race, gender, and ethnicity per se that appear below.

Table 2 presents data for each race/gender/ethnic group of 10th grade students and each of the four subject matter areas, on how the students were distributed across teachers of different race, genders, and ethnicities. Across groups of students and subject matter areas, between 74 and 97 percent of the teachers are white. White students almost exclusively have white teachers in the sample. Numerous male and female teachers appear for all groups of students in all four subject areas. Finally, while black and Hispanic students have primarily white teachers, occasionally over 10 percent of the students from these groups have teachers in a subject matter area from a particular underrepresented minority group. For example, 11 percent of black history students had black male teachers in the sample. Similarly, 16 percent of black male English students and 21 percent of black female English students had black female teachers in the sample.

Small sample sizes and cells in which very few teachers from a group are present make it a priori unlikely that we will observe statistically significant effects.[13] Hence, our prior's are that we will be much more likely to estimate the impact of white female vis-s-vis white male teachers on each student group, then we will be able to estimate the impacts of black and Hispanic teachers on these groups. In cases where a relatively large proportion of minority teachers were present, however, such as the three noted above, statistically significant effects might also be expected to be observed.

Our analytical approach is to estimate for each racial/gender/ethnic group of students and each of the four subject areas in which the tests were given, gain score equations of the form

$$(1) \qquad G^{t}_{ijk} - G^{t-1}_{ijk} = a_0 + a_{1jk}X_{ijk} + a_{2jk}S_{ijk} + a_{3jk}T_{ijk} + \sum_{r=1}^{7} b_{rjk}d_{rijk} + \epsilon_{ijk}$$

In equation (1), the subscript i indexes individuals, the subscript j racial/gender/ethnic groups of students (6), and the subscript k subject matter areas (4). G^t is the student's 10th grade adjusted subject test score and G^{t-1} the student's 8th grade adjusted score. The X's, S's, and T's are vectors of variables that control respectively for personal and family characteristics, school level variables for the student's high school, and characteristics of the 10th grade subject teacher and subject class. a_1, a_2, and a_3 are vectors of parameters and ϵ is a random error term.

The personal and family variables included in our actual empirical analysis are parents' education levels, family size, family income, the student's base year 8th grade adjusted test score, and whether the student was learning disabled or had limited English proficiency. The school level variables included are total enrollment, the percentage of the school's graduates who enroll in college, the racial distribution of the student body, the percentage of teachers with at least a master's degree, and the highest salary paid to full-time teachers in the school. Finally, the class and teacher variables included are the number of students and the proportion of these that were minority in the student's subject area class, as well as the subject teacher's years of experience, degree level, certification in the subject

and subject matter background. Control variables of these types are often found in prior "educational production function" studies.[14]

Of key concern to us is whether teacher race, gender, and ethnicity per se influence how much students learn. So, included in (1) also is a vector of dichotomous variables, d, that indicate whether the student's 10th grade subject matter teacher was a black male, black female, Hispanic male, Hispanic female, other (primarily Asian American) male, other female or white female. The omitted category of teachers is white male, so the coefficients (b) of these variables reflect the impact of each group of teachers on the students' adjusted gain scores vis-a-vis the impact of white male teachers.

The gain scores refer to students' improvement in their academic performance that occurred depending upon when the two tests were administered, sometime between February and June of their 8th grade year and February and June of their 10th grade year. To the extent that teacher characteristics influence student gain scores, the characteristics of the students' 9th grade subject teachers should also be included in the analyses. Similarly, the characteristics of the 8th grade subject area teachers should also be included, both because the 8th grade test was administered to many of the students before the end of the year (which provided time for many 8th grade teachers to influence how much the students learned after the test that year) and because they may also have influenced their students' interest in, and motivation for, future study in the subject area.

No data on the characteristics of 9th grade teachers were collected in NELS. Hence, 9th grade teachers' characteristics could not be included in equation (1) and this omission

may bias our estimates. We report below the results of our attempts to include 8th grade teacher characteristics in the model.

Estimates of the coefficients for the teacher racial/gender/ethnicity dichotomous variables appear in Table 3 for each race, gender, and ethnic group of students by subject area. Coefficients of the control variables, when statistically significant, were typically similar in sign to those found in other studies.[15]

Turning to table 3, for only 11 (out of 130) of these coefficients can we reject the hypothesis that the coefficient is equal to zero at the .10 level of significance or above, and thus conclude that teacher race, gender, and ethnicity per se may influence gain scores. Indeed, for only one of the 24 race/gender/ethnicity/subject matter groups, Hispanic female science students, can one reject at the .05 level of significance the null hypotheses that all of the teacher race/gender/ethnicity variables have no effect on students' gain scores. Together this is strong evidence that, on balance, teacher's race, gender, and ethnicity per se do not play an important role in how much students learn in this sample.

The pattern of the small number of statistically significant coefficients does warrant mention. In comparison to white male teachers, black male teachers are associated with higher history gain scores for black male, white male and white female students, but lower reading scores for Hispanic male students. Black female science teachers are associated with higher science scores for Hispanic female and white female science students.[16] White female teachers are associated with lower reading and history scores for Hispanic male students, but higher science scores for Hispanic female students. Finally, quite strikingly, there is no evidence that, as compared to white male teachers, white female teachers

increase, or decrease, the scores of either white male or white female students in any subject. Given the large sample sizes for white students and white teachers in our analyses, our failure to find significant effects of teacher gender here can not be attributed to small samples.

Several extensions warrant brief mention. Use of a smaller number of dichotomous variables in which gender was not interacted with race or ethnicity (black, Hispanic, other race, female) did not lead to a larger number of statistically significant effects. When the race, gender, and ethnicity of the students' 8th grade subject matter teacher was added to the model, the coefficients of these variables never proved to be jointly statistically significant. Finally, adding dichotomous variables that represented the race, gender, and ethnicity of the student's second observed 10th grade subject area teacher did not improve the fit of the model.

Together these results provide, at best, little support for the notion that teachers' race, gender, and ethnicity per se influence how much students objectively learn. Indeed, in only one case, black male history teachers and students, do we find any evidence that the match of teacher and student gender and race or ethnicity enhances the students' gain scores.

B) Subjective Teacher Evaluations

How can our findings be reconciled with those studies cited earlier that purported to show that teachers' attitudes towards, expectations for, placement of, and feedback to students depends upon the match of teacher and student race, gender, and ethnicity? One

strategy is to ask a related question, do such relationships exist in the NELS data, where our findings suggest that such matches do not influence how much students objectively learn?

To answer the latter question, we reestimated variants of equation (1) in which the student's gain score was replaced by a variable that summarizes the student's 10th grade subject teacher's evaluation of the student.[17] This variable was constructed as the sum of a set of five yes (=1), no (=0) answers to whether the teacher thought the student would go to college, would recommend the student for academic honors, believed the student related well to others, spoke to the student out of class, and believed the student worked hard.

Table 4 shows the mean teacher evaluations of the students by subject area, on a scale of 0 to 5. Female students of each race and ethnicity are rated more highly than male students of the same race and ethnicity in each subject matter area. Hispanic and black students are rated about the same in each subject matter, however, white students tend to be rated higher than the other two groups. Whether this difference reflects differences in background characteristics (the control variables) or different subjective evaluations of a teaching staff that is predominantly white (see Table 2) will be learned from the variants of equation (1) that we reestimated.

Table 5 presents the estimated coefficients of the dichotomous variables for the subject matter teacher's race, gender, and ethnicity from these equations. Quite strikingly 23 of these coefficients are now statistically significantly different from zero, over twice the number observed to be so in the gain score equations. Moreover, the match of teacher and student race or ethnicity often is associated with teachers having higher subjective

evaluations of the students. For example, as compared to white male teachers, black male teachers are associated with significantly higher subjective evaluations for black male students in reading and science and for black female students in mathematics and science. Hispanic teachers of either gender are associated with significantly higher subjective evaluations for Hispanic students of either gender in mathematics. Finally, white female teachers are associated with significantly higher subjective evaluations for white female students in reading, mathematics, and science and for white male students in reading and science, but lower evaluations for white male students in history.

We experimented with several alternative specifications. Since the subjective evaluation variable can only take integer values between 0 and 5 for each individual, the linear model we estimated is not strictly appropriate and a multinomial probit model was also estimated. Since the first two subjective evaluation questions included in our in our index (did the teacher believe the student would attend college and did the teacher recommend the student for academic honors?) are probably conceptually, and canonical correlation analysis suggest also empirically, more important than the others, we also reestimated separate least squares and probit equations for the (1,0) answers to these questions, as well as ordered probit models for the sum of the answers to these two more important questions.[18] In each case the results were qualitatively similar to those reported in Table 5, although somewhat fewer significant coefficients emerged in some of these specifications.

IV. CONCLUDING REMARKS

Do teachers' race, gender, and ethnicity matter? Our analyses of the NELS data suggest that for the most part these teacher characteristics do not affect how much students learned between the 8th and 10th grade in four subject matter areas. They do, however, sometimes seem to influence 10th grade teachers' subjective evaluations of their students, even after one controls for the student's subject matter test scores in the 8th grade. So, for example, while white female teachers do not appear to be associated with higher gain scores for white female students in mathematics and science than white male teachers "produce," white female teachers do have higher subjective evaluations than their white male counterparts of their white female students.

These findings are subject to a number of qualifications that were imposed upon us by the NELS data. No information was available in the data about the characteristics of 9th grade teacher, no measures of teacher ability were present, and teacher and school characteristics had to be treated as pre-determined.[19] However, if one takes our findings at face value, there are two conflicting ways in which one might interpret them.

On the one hand, one might argue that what is crucial is how much students learn in classrooms, in which case one might conclude that teacher race, gender, and ethnicity per se do not matter. On the other hand, one might argue that teachers' subjective evaluations of students may be reflected in the encouragement they provide these students and the "track" on which they place the students or to which they encourage them to aspire. If the latter view is correct, our results suggest that in some cases teachers' race, gender, and ethnicity do matter.

Resolving which interpretation is correct must await the release of subsequent waves of NELS88. In particular, students were resurveyed during the first half of 1992 when they either were seniors in high school or had dropped out of school. When released, these data will allow us to test whether the 10th grade teachers' subjective evaluations of their students influenced how much these students learned between the 10th and 12th grades and the types of classes in which these students were placed, all conditional on the students' 10th grade test scores. They also will allow us to estimate the role that teachers' race, gender, and ethnicity per se play in students' drop-out decisions. Later years NELS data will permit us to analyze teacher affects on actual college going behavior.

It is interesting to speculate about the implications of our findings for whether it is important to match employees and supervisors by race, gender, or ethnicity in the employment relationship? Suppose that one thinks of employees as being analogous to students and supervisors as being analogous to teachers. Similar to our results, an extensive literature in the field of human resources shows that gender, racial, or ethnic differences often exist in supervisor subjective performance evaluations.[20]

However, as in the education case, there is virtually no evidence on whether a match of supervisors and employees by race, gender, and ethnicity objectively influences how well employees perform. Similarly there is no evidence on whether the match, or lack of such, between supervisor and employee characteristics influences an employee's long-run earnings and productivity at a firm. Research addressing these issues should be high on the priority list of those concerned with the progress of women and minorities in the labor market.[21]

Footnotes

1. See, for example, Jaqueline Jordan Irving (1985, 1986).

2. Many of these studies are cited in Ronald Ehrenberg and Dominic Brewer (1993).

3. Studies that do address educational outcomes are cited in Ehrenberg and Brewer (1993), which itself reanalyzed the Coleman report data and found little evidence to support the view that, on average, black students benefitted in the 1960s from having black teachers rather than white teachers.

4. See Ronald G. Ehrenberg (1992).

5. An example illustrates this point. To achieve gender balance in the allocation of National Merit Scholarships to high school seniors, a student's performance on the PSAT verbal aptitude test is weighted twice as heavily as his or her performance on the PSAT mathematics aptitude test in the competitions for these awards.

6. See, for example, Jere Brophy (1985).

7. See, for example, Mark Evans (1992) and Eve Humrich (1988).

8. For a description of NELS, see Steven J. Ingels (1992).

9. Some respondents could not be resurveyed because they had moved and could not be located and some had dropped out of school by 1990 and were administered a separate drop-out survey.

10. See, D.A. Rock and J.M. Pollock (1991) for a discussion of IRT and the NELS data.

11. All students took the same test in the base year. However, the base year test scores also had to be rescaled because the questions on the base year test differed in their degree

of difficulty, discriminating ability and guessability. Thus, two test takers with the same number of correct answers, but different questions incorrect, may have different adjusted base year scores.

12. Since each student was tested in two areas, over 5,000 students were included in the analyses. Approximately 75 percent of the NELS students were in public schools in both years which reduced the original NELS sample to about 13,000 students. The remaining reductions came from missing responses on individual questions from the school, teacher, or student surveys in the 10th grade and from the parent, teacher, or student surveys in the 8th grade, as well as from students being tested in a subject in 10th grade only if they were enrolled in a course in that subject.

13. The estimated variance of a dichotomous variable in a regression equation is given by $S_e^2/(NS_p^2(1-R_{p.x}^2))$ where N is the sample size, S_e^2 is the variance of the error term in the equation, S_p^2 is the variance of the dichotomous variable (which will be small if the variable rarely takes on the value of one), and $R_{p.x}^2$ is the proportion of the variation in the dichotomous variable that can be "explained" by the other variables in the model. See Hugh Pitcher (1979) for an elaboration of this point.

14. See, for example, Eric Hanushek (1986).

15. A table of representative results is available from the authors on request.

16. As Table 2 indicates, only one black female teacher taught a Hispanic female science student in the sample. This fact, plus the large magnitude of the estimated impact of this pairing on the student's gain score (over 5), suggests that the estimated coefficient may also be capturing the impact of other omitted factors associated with the pairing. Other

impacts of this magnitude are observed for black teachers, but in each case that they occurred they were based on at least seven teacher/student observations.

17. Eighth grade teachers in the NELS survey were not asked to provide these subjective evaluations.

18. Should our variable be calculated as an unweighted sum of the scores on the individual teacher subjective evaluation questions? One way to address this question is to ask what linear combination of the scores on the five evaluation questions is most highly correlated with a linear combination of the variables on the right-hand side of equation (1). This is what the method of canonical correlation does, although a weakness of the method is that one can not perform tests of statistical significance for coefficients of individual variables using it. When we employed this method, using data from the various race/gender/ethnicity/subject area groups, the weights placed on the first two subjective evaluation variables were typically close to one and the weights placed on the remaining three subjective evaluation variables were typically much smaller. For a discussion of canonical correlation analysis, see Bruce Thompson (1984).

19. Ehrenberg and Brewer's (1993) reanalysis of the Coleman Report data make use of teacher ability measures and tests whether treating school and teacher characteristics (including teacher race) influences their findings. The lack of geographic identifiers in the NELS data set preclude similar analyses here.

20. Robert Bretz, Jr., George Milkovich, and Walter Read (1992), and George Milkovich and John Boudreau (1994) present summaries of the literature on gender, race, and ethnicity bias in performance appraisals.

21. Donna Rothstein is addressing some of these issues in her Cornell dissertation which is in progress.

References

Bretz, Robert D., Jr., George Milkovich, and Walter Read, "Performance Appraisal Research and Practice" (Center for Advanced Human Resources Working Paper 92-15, Cornell University, 1992).

Brophy, Jere, "Interactions of Male and Female Students with Male and Female Teachers" in Louise Cherry Wilkinson and Cora B. Marrett, eds., Gender Influences in Classroom Interaction (New York: Academic Press, 1985): pp. 115-142.

Ehrenberg, Ronald G., "The Flow of New Doctorates," Journal of Economic Literature 30 (June 1992): pp. 830-875.

Ehrenberg, Ronald G. and Dominic J. Brewer, "Did Teachers' Verbal Ability and Race Matter in the 1960s?: Coleman Revisited" (mimeo, Cornell University, December 1993).

Evans, Mark O. "An Estimate of Race and Gender Role-Model Effects in Teaching High School," Journal of Economic Education 23, No. 3 (Summer 1992): pp. 209-217.

Hanushek, Eric, "The Economics of Schooling," Journal of Economic Literature 24 (September 1986): pp. 1-10.

Humrich, Eve, "Sex Differences in the Second IEA Science Study – U.S. Results in an International Content," Bethesda, MD: ERIC Document Reproduction Service, ED 292 649, April 1988.

Ingels, Steven J. et al., NELS: First Follow-Up (Washington, DC: U.S. Department of Education, 1992).

Irving, Jaqueline Jordon, "Teacher Communication Patterns as Related to the Race and Sex of the Student," Journal of Educational Research 78 (1985): pp. 338-345.

Irving, Jaqueline Jordon,"Teacher-Student Interactions: Effects of Student, Race, Sex, and Grade Level," Journal of Educational Psychology 78 (1) (1986): pp. 14-21.

Milkovich, George T. and John W. Boudreau, Human Resource Management (Burr Ridge, IL: Irwin, 1994).

Pitcher, Hugh M., "A Sensitivity Analysis to Determine Sample Sizes for Performing Impact Evaluation of the CETA Program" in Ronald Ehrenberg, ed., Research in Labor Economics: Supplement 1 (Greenwich, CT: JAI Press, 1979): pp. 37-78.

Rock, D.A. and J.M. Pollock, Psychometric Report for the NELS: 88 Base Year Test Battery (Washington, DC: National Center for Education Statistics, 1991).

Thompson, Bruce, Canonical Correlation Analysis: Uses and Interpretation (Beverly Hills, CA: Sage Publications, 1984).

Table 1

A) Mean 8th Grade Test Scores by Gender, Race, and Ethnicity
(standard deviation)

	History [30]*	Reading [21]	Math [40]	Science [25]
Black Males	14.706	15.052	30.482	9.737
	(4.772)	(5.914)	(9.563)	(3.750)
n=	66	140	114	118
Black Females	15.153	16.466	30.079	9.222
	(4.452)	(6.063)	(11.212)	(3.261)
n=	55	156	132	137
Hispanic Males	17.733	15.927	33.443	10.631
	(5.512)	(5.916)	(10.714)	(4.146)
n=	82	164	163	127
Hispanic Females	14.708	16.858	32.189	9.693
	(4.688)	(6.165)	(10.746)	(3.122)
n=	80	177	147	146
White Males	18.062	19.295	39.717	13.188
	(5.658)	(6.948)	(11.648)	(4.619)
n=	724	1106	1201	926
White Females	17.634	20.818	39.088	12.022
	(4.856)	(7.029)	(11.133)	(4.105)
n=	771	1192	1288	988
Total n=	1776	2848	3029	2445

B) Mean Adjusted Gain Scores by Gender, Race and Ethnicity of Students Between the Eighth and Tenth Grades
(standard deviation)

	History	Reading	Math	Science
Black Males	2.950	2.198	4.586	1.277
	(4.694)	(5.316)	(6.716)	(3.303)
n=	66	140	114	118
Black Females	2.491	1.882	4.640	.925
	(3.333)	(4.679)	(6.429)	(3.170)
n=	55	156	132	132
Hispanic Males	3.170	2.611	4.517	2.030
	(4.664)	(4.178)	(6.358)	(3.761)
n=	80	177	147	146
Hispanic Females	2.238	2.826	4.753	1.438
	(3.690)	(4.789)	(5.838)	(2.952)
n=	80	177	147	146
White Males	2.569	2.501	4.970	2.780
	(4.214)	(5.601)	(7.679)	(3.673)
n=	724	1006	1201	923
White Females	2.609	2.396	5.144	2.240
	(3.594)	(5.051)	(6.435)	(3.429)
n=	771	1192	1288	980

* [ ] number of questions on the test.
Source: Authors' calculations from the NELS data.

Table 2

Racial, Ethnic, and Gender Distribution of 10th Grade Students' Teachers, by Student Group and Subject Matter
[share of the student group's teachers]

		Teacher Group							
Students		BM	BF	HM	HF	OM	OF	WM	WF
Black Males	H	7 [.11]	1 [.02]	1 [.02]	0	0	0	35 [.53]	22 [.33]
	E	3 [.02]	22 [.16]	5 [.04]	0	0	0	20 [.14]	89 [.64]
	M	5 [.04]	9 [.08]	1 [.01]	1 [.01]	0	2 [.02]	35 [.31]	61 [.53]
	S	7 [.06]	8 [.07]	0	0	0	0	62 [.53]	39 [.33]
Black Females	H	6 [.11]	4 [.07]	0	0	0	0	29 [.53]	16 [.29]
	E	7 [.04]	33 [.21]	0	0	0	0	17 [.11]	99 [.63]
	M	7 [.05]	9 [.07]	1 [.01]	0	1 [.01]	0	46 [.35]	68 [.52]
	S	7 [.05]	7 [.05]	0	0	0	0	69 [.50]	53 [.39]
Hispanic Males	H	0	2 [.02]	1 [.01]	0	0	0	54 [.66]	22 [.27]
	E	2 [.01]	7 [.04]	1 [.01]	13 [.08]	0	3 [.02]	41 [.25]	93 [.57]
	M	4 [.02]	4 [.02]	8 [.05]	11 [.07]	2 [.01]	3 [.02]	73 [.45]	57 [.35]
	S	2 [.02]	0	10 [.08]	3 [.02]	0	1 [.01]	63 [.50]	48 [.38]

Table 2 (continued)

Students		Teacher Group BM	BF	HM	HF	OM	OF	WM	WF
Hispanic Females	H	0	0	6 [.08]	1 [.01]	0	0	52 [.65]	16 [.20]
	E	2 [.01]	5 [.03]	1 [.01]	21 [.12]	0	2 [.01]	38 [.21]	106 [.60]
	M	0	4 [.03]	13 [.09]	4 [.03]	2 [.01]	4 [.03]	71 [.48]	49 [.33]
	S	1 [.01]	1 [.01]	9 [.06]	2 [.01]	1 [.01]	3 [.02]	71 [.49]	58 [.40]
White Males	H	14 [.02]	8 [.01]	3 [0.0]	1 [0.0]	1 [0.0]	4 [.01]	503 [.69]	189 [.26]
	E	6 [.01]	34 [.03]	1 [0.0]	7 [.01]	1 [0.0]	7 [.01]	346 [.31]	688 [.62]
	M	2 [0.0]	12 [.01]	6 [.01]	4 [0.0]	3 [0.0]	5 [0.0]	676 [.56]	491 [.41]
	S	15 [.02]	7 [.01]	1 [0.0]	2 [0.0]	1 [0.0]	7 [.01]	536 [.58]	343 [.37]
White Females	H	16 [.02]	15 [.02]	6 [.01]	2 [0.0]	1 [0.0]	7 [.01]	506 [.66]	217 [.28]
	E	6 [.01]	41 [.03]	3 [0.0]	2 [0.0]	2 [0.0]	4 [0.0]	332 [.28]	783 [.66]
	M	7 [.01]	8 [.01]	6 [0.0]	1 [0.0]	8 [.01]	7 [.01]	707 [.55]	537 [.42]
	S	8 [.01]	9 [.01]	1 [0.0]	0 [0.0]	0 [0.0]	9 [.01]	567 [.58]	372 [.38]

where BM - black males
BF - black females
HM - Hispanic males
HF - Hispanic females
OM - other males
OF - other females
WM - white males
WF - white females

Source: Author's computations from the NELS data.

Table 3

Impact of Tenth Grade Teachers' Gender, Race, and Ethnicity on Students' Gain Scores[a]
(absolute value of t statistics)

Teachers	Subject			
	History	Reading	Math	Science
A. Black Males				
Black Male	5.133 (2.0)*	-5.152 (1.5)	1.404 (0.4)	1.087 (0.7)
Black Female	-2.037 (0.3)	1.333 (0.7)	.178 (0.1)	.324 (0.2)
Hispanic Male	1.533 (0.3)	1.811 (0.6)	-18.391 (2.5)*	N
Hispanic Female	N	N	-2.333 (0.3)	N
Other Race Male	N	N	N	N
Other Race Female	N	N	3.626 (0.7)	N
White Female	1.901 (1.1)	.084 (0.1)	1.376 (0.9)	.554 (0.8)
F[dof,dof]	1.22 [4,38]	1.04 [4,11]	1.55 [6,84]	.34 [3,90]
B. Black Females				
Black Male	2.360 (1.1)	.474 (0.2)	-4.022 (1.4)	-.016 (0.0)
Black Female	1.007 (0.4)	.688 (0.4)	-1.812 (0.7)	-.776 (0.6)
Hispanic Male	N	N	3.222 (0.5)	n
Hispanic Female	N	N	N	N
Other Race Male	N	N	-6.039 (0.9)	N
Other Race Female	N	N	N	N
White Female	-.421 (0.3)	1.214 (0.9)	.220 (0.2)	.819 (1.4)
F[dof,dof]	.51 (3,28]	.36 [3,132]	.64 [5,103]	.53 [3,109]
C. Hispanic Males				
Black Male	N	-5.732 (1.8)**	-1.274 (0.3)	.341 (0.1)
Black Female	3.095 (1.0)	-1.772 (1.0)	2.244 (0.6)	N
Hispanic Male	4.116 (0.9)	1.362 (0.3)	-1.168 (0.4)	.532 (0.3)
Hispanic Female	N	-1.042 (0.6)	.617 (0.2)	-.628 (0.2)
Other Race Male	N	N	-3.433 (0.7)	N
Other Race Female	N	.186 (0.6)	-.937 (0.2)	-.520 (0.1)
White Female	2.050 (1.8)**	-1.386 (1.7)**	.967 (0.8)	-.832 (0.9)
F[dof,dof]	1.79 (3,54]	.95 [6,133]	.29 [7,131]	.26 [97]

Table 3 (continued)

Teachers	Subject			
	History	Reading	Math	Science
D. Hispanic Females				
Black Male	N	-2.145 (0.6)	N	2.406 (0.8)
Black Female	N	1.881 (0.8)	2.864 (0.8)	5.087 (1.8)**
Hispanic Male	-1.550 (0.8)	.512 (0.1)	1.250 (0.6)	.297 (0.3)
Hispanic Female	-.399 (0.1)	-.819 (0.5)	5.700 (1.6)	2.635 (1.2)
Other Race Male	N	N	2.848 (0.6)	.720 (0.2)
Other Race Female	N	-2.196 (0.6)	1.587 (0.5)	-.876 (0.5)
White Female	-1.405 (1.1)	.764 (0.8)	1.094 (0.9)	1.752 (3.2)*
F[dof,dof]	.50 [3,52]	.54 [6,146]	.62 [6,116]	2.2[b] [7,114]
E. White Males				
Black Male	2.240 (2.0)*	1.797 (0.8)	7.800 (1.5)	.650 (0.7)
Black Female	-.149 (0.1)	-.148 (0.2)	.842 (0.4)	2.438 (1.8)**
Hispanic Male	.705 (0.3)	1.811 (0.3)	.574 (0.2)	-1.808 (0.5)
Hispanic Female	.604 (0.1)	2.817 (1.3)	-.996 (0.3)	-1.567 (0.6
Other Race Male	5.649 (1.4)	3.796 (0.7)	2.468 (0.6)	2.968 (0.9)
Other Race Female	-.853 (0.4)	-.987 (0.5)	6.269 (1.9)**	1.175 (0.9)
White Female	.381 (1.1)	.487 (1.3)	.311 (0.7)	-.322 (1.3)
F[dof,dof]	.69 [7,463]	.69 [7,1075]	.88 [7,1169]	1.14 [7,892]
F. White Females				
Black Male	2.411 (2.8)*	3.010 (1.5)	-1.344 (0.6)	1.018 (0.4)
Black Female	-.058 (0.1)	.674 (0.8)	-1.290 (0.6)	-1.010 (0.9)
Hispanic Male	-.865 (0.6)	3.927 (1.4)	.145 (0.1)	-.138 (0.0)
Hispanic Female	.284 (0.1)	3.186 (0.9)	-7.181 (0.8)	N
Other Race Male	.346 (0.1)	3.324 (1.0)	-4.895 (2.2)*	N
Other Race Female	-.063 (0.0)	-3.879 (1.6)	-1.006 (0.4)	-.979 (0.9)
White Female	.385 (1.4)	.402 (1.3)	.537 (1.4)	.090 (0.1)
F [dof,dof]	1.39 [7,739]	1.40 [7,1161]	1.28 [7,1256]	.54 [5,957]

[a]See the text for a description of the other variables in the model.
F represents the F statistic to test the null hypothesis that the vector of teacher gender, race, and ethnicity coefficients are jointly equal to zero.
[b]Reject the null hypothesis at the .05 level.
N = no teachers in this category.
*(**) Coefficient is statistically significantly different from zero at the .05 (.10) level, two-tail test.

Table 4

Mean Teacher Subjective Evaluation of Students' Scores, by Subject Matter and Gender, Race, and Ethnicity of the Students[a]
(standard deviation)

	History	Reading	Math	Science
Black Males	2.102	2.139	2.299	2.056
	(1.141	(1.417)	(1.348)	(1.282)
n=	49	108	77	89
Black Females	2.500	2.906	2.771	2.538
	(1.348)	(1.196)	(1.317)	(1.400)
n=	41	117	105	104
Hispanic Males	2.169	2.158	2.157	2.012
	(1.522)	(1.328)	(1.322)	(1.340)
n=	59	114	115	84
Hispanic	2.237	2.899	2.623	2.648
Females	(1.343)	(1.203)	(1.279)	(1.409)
n=	59	119	114	95
White Males	2.703	2.575	2.685	2.646
	(1.355)	(1.377)	(1.290)	(1.344)
n=	495	790	819	656
White Females	3.025	3.087	2.946	2.996
	(1.251)	(1.280)	(1.233)	(1.297)
n=	511	846	911	722

Source: Authors' calculations from the NELS data.

[a]The subjective evaluation is the sum of yes (=1), no (=0) responses by the teachers to the following five questions: Did the teacher
1. think the student would probably go to college?
2. recommend the student for academic honors (i.e.,either honors classes or recognition)?
3. think the student relates well to others?
4. speak to the students outside of class?
5. think the student works hard?

Table 5

Impact of Teachers' Gender, Race, and Ethnicity on the Teachers' Subjective Evaluations of Their Students[a]
(absolute value of t statistics)

Teachers	Subject			
	History	Reading	Math	Science
A. Black Males				
Black Male	-.080 (0.1)	1.890 (1.8)**	-.025 (0.0)	1.323 (1.9)**
Black Female	N	.480 (0.9)	.689 (0.9)	.732 (1.1)
Hispanic Male	-1.112 (0.7)	1.233 (1.5)	N	N
Hispanic Female	N	N	.368 (0.3)	N
Other Race Male	N	N	N	N
Other Race Female	N	N	-.261 (0.2)	N
White Female	-.691 (1.5)	.591 (1.5)	.035 (0.1)	.444 (1.3)
F[dof,dof]	.90 [3,22]	1.43 [4,79]	.20 [5,48]	1.39 [3.61]
B. Black Females				
Black Male	-.388 (0.4)	-.422 (0.7)	1.192 (1.8)**	2.387 (3.3)*
Black Female	-.665 (0.4)	.459 (1.0)	.836 (1.5)	.130 (0.2)
Hispanic Male	N	N	.443 (0.4)	N
Hispanic Female	N	N	N	N
Other Race Male	N	N	N	N
Other Race Female	N	N	N	N
White Female	-.670 (0.8)	.318 (0.9)	.066 (0.2)	.452 (1.5)
F[dof,dof]	.23 [3,15]	1.03 [3,90]	1.30 [4,77]	4.09[b] [3,76]
C. Hispanic Males				
Black Male	N	1.394 (1.0)	.012 (0.0)	.541 (0.5)
Black Female	.336 (0.2)	.224 (0.3)	1.262 (1.5)	N
Hispanic Male	1.897 (1.0)	N	1.253 (2.2)*	.076 (0.1)
Hispanic Female	N	.061 (0.1)	1.465 (2.6)*	-.108 (0.1)
Other Race Male	N	N	1.280 (1.3)	N
Other Race Female	N	2.693 (1.8)**	2.497 (2.9)*	1.425 (0.7)
White Female	-.692 (1.1)	.022 (0.1)	.242 (0.9)	-.540 (1.3)
F [dof,dof]	.78 [3,31]	.85 [5,84]	2.54[b] [7,83]	.85 [5,54]

Table 5 (continued)

Teachers	Subject			
	History	Reading	Math	Science
D. Hispanic Females				
Black Male	N	-.897 (0.9)	N	2.317 (1.4)
Black Female	N	.578 (0.8)	.648 (1.1)	-1.028 (0.6)
Hispanic Male	.706 (0.9)	-.696 (0.5)	1.322 (3.3)*	.284 (0.4)
Hispanic Female	N	.400 (0.9)	1.508 (2.2)*	1.401 (1.1)
Other Race Male	N	N	-.895 (1.0)	1.446 (0.9)
Other Race Female	N	N	0.000 (0.0)	.960 (1.0)
White Female	-.545 (0.9)	-.081 (0.3)	.489 (2.0)*	.360 (0.9)
F[dof,dof]	.90 [2,33]	.82 [5,89]	2.73[b] [6,83]	.74 [7,63]
B. White Males				
Black Male	.809 (1.7)**	1.617 (3.2)*	N	.261 (0.8)
Black Female	-.126 (0.3)	.207 (0.8)	-.129 (0.3)	-.332 (0.5)
Hispanic Male	1.413 (1.7)**	-2.438 (2.0)*	-.094 (0.2)	1.490 (1.2)
Hispanic Female	-.578 (0.5)	-.101 (0.3)	1.258 (1.9)*	-1.432 (1.2)
Other Race Male	-.118 (0.1)	-.308 (0.3)	1.194 (1.0)	.030 (0.0)
Other Race Female	-.307 (0.4)	-.153 (0.3)	.455 (0.9)	.273 (0.5)
White Female	-.333 (2.6)*	.405 (4.1)*	.041 (0.5)	.191 (1.3)**
F[dof,dof]	2.05 [7,463]	4.47[b] [7,759]	.91 [6,788]	1.04 [7,625]
C. White Females				
Black Male	-.875 (2.2)*	.502 (1.1)	.653 (1.2)	-.050 (0.1)
Black Female	.118 (0.2)	.218 (1.0)	.510 (0.9)	.045 (0.1)
Hispanic Male	1.115 (1.6)	.183 (0.3)	-.524 (1.2)	N
Hispanic Female	-.347 (0.3)	-.142 (0.2)	1.324 (0.9)	N
Other Race Male	1.346 (1.2)	-.477 (0.6)	.274 (0.7)	N
Other Race Female	.499 (0.6)	.186 (0.2)	-.563 (1.0)	-.437 (0.9)
White Female	.050 (0.4)	.173 (2.0)*	.152 (2.0)*	.286 (3.0)*
F [dof,dof]	1.45 [7,480]	.82 [7,815]	1.47 [7,879]	2.54[b] [4,694]

F represents the F statistic to test the null hypothesis that the vector of teacher gender, race, and ethnicity coefficients are jointly equal to zero.

[a] Also included in the model are all of the explanatory variables from the gain score equations.

[b] Reject the null hypothesis at the .05 level.

N = no teachers in this category.

*(**) Coefficient is statistically significantly different from zero at the .05 (.10) level, two-tail test.