Topics in Economic Analysis & Policy

Volume 4, *Issue* 1 2004 *Article* 12

Has Curriculum Closed the Test Score Gap in Math?

Heather Rose*

*Public Policy Institute of California, rose@ppic.org

Has Curriculum Closed the Test Score Gap in Math?*

Heather Rose

Abstract

This study examines the extent to which convergence in mathematics course-taking behavior is responsible for narrowing the Hispanic-white and the black-white test score gaps during the 1980s. Mathematics curriculum is measured in detail using high school transcript data from both High School and Beyond and the National Education Longitudinal Study of 1988. After controlling for demographic, family, and school characteristics, changes in curriculum account for about 60 percent of the narrowing Hispanic-white test score gap between 1982 and 1992. However, the black-white test score gap did not drop significantly.

*I would like to thank Julian Betts for helpful comments on earlier drafts of this paper. Heather Rose, Public Policy Institute of California, 500 Washington Street, Ste. 800, San Francisco, CA, 94111.

1. Introduction

Test scores are becoming the standard yardstick for measuring the success of students, teachers, and schools alike. For students, test scores are important for several reasons. First, high scores are increasingly crucial for college admission. Second, to the extent that test scores represent acquired skills, students with higher scores are rewarded more lucratively in the labor market. And finally, there is a growing trend to reward schools for high student test scores, thereby indirectly benefiting the students themselves.

A disconcerting pattern is that, on average, black and Hispanic students consistently score lower than white students on standardized tests and are thus at a disadvantage in terms of college admission and wages.[1] Data from the National Assessment of Educational Progress shows that these test score gaps have been shrinking over time. Although scores for all ethnicities display an upward trend from the early 1980s onward, the average score for black students has increased relatively more than for white students. (Grissmer et al., 1988).

Another trend emerging during the 1980s was the convergence in course-taking patterns among students of different ethnicities. The completion rate of the New Basics program increased substantially among all ethnic groups (Green et al., 1995).[2] By the early 1990s, no difference remained between completion rates of black and white students, but Hispanic completion rates still lagged somewhat. Not only was the *number* of math courses taken converging, but so was the *type*. Grissmer et al. (1988) report that black enrollment rates were rising more rapidly than white rates in courses at or above the algebra level. They conjecture that this course-taking behavior could partly explain the convergent test scores in math.

Although the collage of evidence indicates that curriculum could play a role in the narrowing test score gap, conclusive evidence still eludes researchers. Rather than simply noting temporal consistencies in trends between average test scores and curriculum, I systematically analyze the extent to which curriculum could be responsible for narrowing the test score gap between 1982 and 1992. Previous research has shown that family, school, and socioeconomic factors account for only one quarter of the reduction in the test score gap.[3] While controlling for these characteristics, I determine that about 60 percent of the remaining test score gap between white and Hispanic students can be attributed to a convergence in course-taking behavior. However, curriculum appears to have little effect on changes in the black-white test score gap.

I focus on the effects of mathematics curriculum on the math test score gap, because research by Murnane, Willett, and Levy (1995) and Grogger and Eide (1995) shows that between the 1970s and the 1980s the relative importance of math test scores in

---

[1] See Green et al. (1995), Hedges and Nowell (1998), Jencks and Phillips (1998), Raisinski et al. (1993) for descriptions of the test score gaps.

[2] The New Basics curriculum was recommended by the famous "A Nation at Risk" report by the National Commission on Excellence in Education (1983). This report advised that high school students take four years of English, three years of math, three years of science, half a year of computer science, and two years of foreign language.

[3] See Grissmer et al. (1998) and (1994), Hedges and Nowell (1998), Ferguson (1998), and Cook and Evans (2000).

determining earnings grew substantially. Those studies also conclude that math achievement is a better predictor of adult outcomes than are other types of test scores commonly available. As Rose and Betts (2004) demonstrates, math curriculum is an important predictor of earnings; more-advanced math courses lead to higher earnings 10 years after students leave high school.

2. Data Sources and the Definition of the Test Score Gap

The data for this study come from two principal sources. The first is the High School and Beyond Sophomore Cohort: 1980-92 (HSB). This longitudinal study surveyed over 30,000 high school sophomores in 1980 and followed up approximately 15,000 of them in 1982, 1984, 1986, and 1992. The second is the National Education Longitudinal Study: 1988-1994 (NELS). This study surveyed nearly 25,000 eighth graders in 1988 and biennially thereafter. High school transcripts were gathered for approximately 17,000 students once they left high school.

The restricted versions of both datasets provide extremely detailed high school transcript data, which include every course taken by the student, the term it was taken, the grade received, and the number of credits earned. The coding of courses is identical in the two data sets, classifying every math course a student took during high school into one of 42 categories using the standard Classification of Secondary School Courses. Such detail is quite appealing but analytically daunting. I aggregate these 42 classes into seven broader categories based on a classification system provided by the National Center for Education Statistics. In increasing level of difficulty, these categories are: vocational math, pre-algebra, algebra/geometry, intermediate algebra, trigonometry, pre-calculus, and calculus. The appendix provides more detail about the type of courses in each category. The number of credits students earned in each course is standardized such that a typical one-year course is assigned one credit and a half-year course is assigned 0.5 credits (these credits are Carnegie units). Combining these two variables yields my primary measures of curriculum for the regression analysis: the number of credits earned by each student in each of the seven math course categories. For expositional purposes, Section 3 of this paper focuses on the highest level of math course students completed in high school as a way of describing curriculum trends.

Both data sets also contain math test scores. The content and scoring method is not the same for the HSB and NELS tests, so actual student point scores (i.e., test score levels) and gaps in actual point scores between two groups of students cannot be compared across the data sets. However, the test score gaps can be compared by first standardizing them within each data set. I measure the test score gap between two different groups of students at time t as:

$$Gap_t = \frac{M_{x,t} - M_{white,t}}{SD_{white,t}}, \tag{1}$$

where $M_{x,t}$ is the mean test score of students in group x (either Hispanic or black) at time t, and $SD_{white,t}$ refers to the standard deviation of white students' test scores at time t. This measure indicates how many white standard deviations the Hispanic or black average score is from the white average score. Because it is scale free, this measure of the gap can be compared across the two data sets. Furthermore, this measure exhibits another important property: it is invariant to changes in the ethnic composition of the student population.[4]

Similarly, the test scores that I use in my analysis, $TS_{i,t}$, for an individual student i at time t refer to standardized scores, i.e., original test scores ($TS_{i,t}^{orig}$) that have been converted by subtracting the mean white test score and then normalizing by the white standard deviation. The following formula clarifies this relationship:

$$TS_{i,t} = \frac{TS_{i,t}^{orig} - M_{white,t}}{SD_{white,t}}. \qquad (2)$$

The test score in the HSB data is based on a math test in which students had 21 minutes to answer 38 questions. The questions required students to compare two quantities and determine which one was greater, whether they were equal, or whether the relationship was indeterminable based on the given data. The tests were administered when the students were sophomores and again when they were seniors. I use the HSB-computed Item Response Theory (IRT) scores from the senior test as my measure of math test score.[5]

The mathematics content and testing procedure were different in NELS. Students had 30 minutes to answer 40 questions. In addition to quantitative questions like those in HSB, the NELS math test included questions involving word problems, graphs, equations, and geometric figures (Quin, 1994). Furthermore, in the NELS, all eighth

[4] Previous research generally normalizes the difference in average test scores by the standard deviation of all test scores, not just the scores of white students. However, shifts in the composition of student groups can affect such a measure even if the true test score gap between the groups is constant. (I thank Mark Appelbaum for bringing this issue to my attention.) For example, if all group *a* students score 70 points and all group *b* students score 100 points on the same tests in both 1982 and 1992, the true test score gap between group *a* and group *b* students does not change. However, if the number of group *a* students in the population doubles during the 1980s but the population of group *b* students remains constant, the pooled standard deviation in 1992 will be greater than it was in 1982 as a result of the influx of lower scoring students. The constant difference in test scores coupled with a larger overall standard deviation would yield a smaller gap in 1992, implying that the average group *b* test score is fewer standard units away from the group *a* score in 1992 than it was in 1982 even though the absolute gap has not changed. Normalizing the difference in average test scores by the white standard deviation rather than the pooled standard deviation eliminates this problem.

[5] Item Response Theory is a "method of estimating achievement level by considering the pattern of right, wrong, and omitted responses on all items administered to an individual student. Rather than merely counting right and wrong responses, the IRT procedure also considers characteristics of each of the test items, such as their difficulty, and the likelihood that they could be guessed correctly by low-ability individuals. IRT scores are less likely than simple number-right formula scores to be distorted by correct guesses on difficult items if a student's response vector also contains incorrect answers to easier questions." See Ingels (1995, p. M-4). For a comprehensive explanation of the IRT method, see Rock and Pollack (1995).

graders took the same test in the base year. Yet in the sophomore and senior follow-up surveys, three different math tests were administered depending on whether the student had scored in the lowest quartile, the interquartile range, or the top quartile on the previous test. These test scores were put on the same scale using IRT methods, yielding a score that is comparable across students within NELS.

The multiple, adaptive math tests used in NELS permitted a broader range of math achievement to be measured by substantially raising the ceiling and slightly lowering the floor (Green et al., 1995). This change in the test structure could have been responsible for a widening test score gap between 1982 and 1992, had a widening occurred. However, the change in structure did not pose a problem for this study, because test score gaps actually narrowed during this time period, reflecting real change. If anything, the narrowing gaps may be understated because of the test score structure in NELS.

In addition to the scholastic information, both HSB and NELS provide a wealth of personal and family characteristics. Both data sets also include high school dropouts in the transcript and the follow-up surveys, allowing them to be included in the analysis. Throughout the paper, references to the 1982 and 1992 high school classes mean the 1980 and 1990 sophomore cohorts two years later, including students who graduated and those who dropped out after their sophomore year.

Obviously, those students for whom math curriculum or test score data are missing are excluded from this analysis. I also limit the data in two other ways. I restrict the sample to those students who attended public schools because of the inherent selection problems in analyzing public-private school choice. I also exclude students who transferred between schools during high school. A model of test score gaps should account for school characteristics; in the case of students who attended two schools, it would not be clear which set of school characteristics to use. Further details regarding the data are located in the appendix.

3. Ethnic Trends in Course-Taking Behavior and Test Scores

This section documents the degree to which mathematics course-taking patterns converged among Hispanic, black, and white students between 1982 and 1992. It also reports the contemporaneous trends in test scores. Longer-term trends from the National Assessment of Educational Progress (NAEP) supplement the HSB and NELS trends. The review of NAEP helps ensure that the trends in HSB and NELS are not due to peculiarities in these datasets' test measures or timing, and it helps paint a picture of test scores that begins before 1982 and extends after 1992.

3.1 Course-Taking Behavior

One of the most straightforward ways to measure course-taking behavior is by the highest level of math course that students completed during high school. Table 1 presents the percentage of students whose highest math course during their four years at high school was the course in the left-most column. The table presents overall completion rates for 1982 and 1992 seniors and also presents results by ethnicity.

Table 1
Percentage of Public High School Students Completing
Specific Math Course as their Highest

	1982				1992			
Highest Course	Overall	White	Black	Hispanic	Overall	White	Black	Hispanic
Vocational	27.4	22.1	43.4	41.4	9.8	8.5	17.7	10.9
Pre-Algebra	8.8	7.7	11.8	11.8	7.9	7.0	10.7	12.0
Algebra/Geometry	29.5	30.8	23.8	28.3	24.7	23.4	28.5	34.3
Intermediate Algebra	15.2	16.8	11.5	10.1	24.4	25.2	21.5	21.2
Trigonometry	12.0	14.0	7.1	5.6	14.2	15.5	10.5	9.7
Pre-Calculus	3.3	4.0	1.3	1.1	9.7	10.5	6.4	6.8
Calculus	3.9	4.6	1.2	1.6	9.2	9.9	4.7	5.1
Algebra or Higher	63.8	70.2	44.8	46.8	82.3	84.5	71.6	77.1
Trigonometry or Higher	19.2	22.6	9.6	8.4	33.2	35.9	21.6	21.6
Sample Size	9,146	5,409	1,157	2,018	11,885	8,395	1,102	1,461

Notes: The highest level of math class is the one in which a student completed at least 0.5 credits (Carnegie Units) with a passing grade. The sample size is the number of observations in each category. Observations were weighted to compute the percentages. To be consistent with the regression analysis and across datasets, I used the weights that produce weighted twelfth grade students statistics in cross-sectional analyses (FU1WT from HSB and F2QWT from NELS). In addition to the three ethnic groups listed, the overall category includes Asian, American Indian, and students from other ethnicities.

Overall, students completed more-advanced math courses in 1992 than they did in 1982. The percentage of students completing calculus, the most advanced math course, more than doubled during this decade. Overall completion rates jumped from nearly 4 percent to over 9 percent. The decade also witnessed a sharp rise in the share of students whose final course was at the level of intermediate algebra, trigonometry, and pre-calculus. The percentage of students whose highest math course was intermediate algebra rose dramatically from 15 percent to 24 percent. Coupled with the rise in academic course completion rates was a dramatic drop in the percentage of students who did not progress beyond vocational math. Whereas 27 percent of students took only vocational math courses in 1982, a decade later that number dropped by almost two-thirds to about 10 percent.

In some respects, the trend toward a more rigorous curriculum is more pronounced for Hispanic and black students than for white students. Whereas 70 percent of white students completed an algebra class or higher in 1982, about 45 percent of black and Hispanic students did so, leaving a gap of about 25 percentage points. A decade later, however, Hispanic students lagged by only 7 percentage points, and black students lagged by only 13 percentage points. In general, the 1992 completion rates for black and Hispanic students look remarkably like the white completion rates in 1982, suggesting that black and Hispanic students are catching up with a lag of about a decade. The table also shows that the percentage of students completing trigonometry or higher grew for students of all ethnicities between 1982 and 1992. Although the absolute percentage point gap between white and minority students remained constant at about 14 points in this set of courses, in proportional terms minorities were catching up.

Given these gains in the higher-level courses, it is not surprising that the percentage of minority students who completed only vocational math courses dropped substantially. In 1982, 41 percent of Hispanic students took only vocational math. A decade later, however, only 11 percent did not progress beyond this level. Similarly, the percentage of black students who took only vocational math fell from 43 percent to 18 percent. White students also made improvements, but measured as an absolute decrease of 13 percentage points, the improvement was not as dramatic. Nonetheless, in relative terms, black students in 1992 were still twice as likely as white students to take only vocational math. However, Hispanic students were only slightly more likely than white students to take only vocational math.

In summary, the black-white and Hispanic-white percentage point gaps closed more at the algebra/geometry or higher level than they did at the trigonometry or higher level. This is logical because as students progress beyond vocational math courses, initially they are likely to move just a few steps up the sequence of courses. Nonetheless, minorities kept pace with the white percentage point gains at high-level math courses as well. Although both minority groups made improvements, on average Hispanic students improved the rigor of their math curriculum more than black students did.

3.2 Test Score Gaps in 1982 and 1992

Several reports have documented the test score gap between students of different ethnicities. Table 2 presents some representative results in addition to my estimates from this study. Most notably, all of the studies find that the test score gaps for black and Hispanic students were very large in the 1980s. My results indicate that black high school seniors scored 0.87 standard deviations below their white counterparts in 1982. This gap declined only somewhat to 0.81 by 1992. The Hispanic-white test score gap shrank considerably more, falling 0.15 standard deviations from 0.80 to 0.65 standard units.

Table 2
Test Score Gap Findings

Source	Sample	Black-white Gap	Hispanic-white Gap
This Study	1982 HSB Seniors 1992 NELS Seniors	.87 .81	.80 .65
Green et al. (1995)	1980 HSB Seniors 1992 NELS Seniors	.92 .87	NA
Hedges and Nowell (1998)	1980 HSB Seniors 1982 HSB Seniors 1992 NELS Seniors	.85 .85 .82	NA
Raisinski et al. (1993)	1980 HSB Sophomores 1990 NELS Sophomores	.89 .75	.77 .59

Note: These are unconditional test score gaps. They do not account for any student characteristics.

There are several reasons why my results differ slightly from the other studies. I use a sample of public schools, whereas the other studies generally include both public and private schools. I normalize differences in test score means by the standard deviation of white students' test scores; other studies generally normalize by the pooled standard deviation. The many studies use different weighting schemes depending on the group of students to which they extrapolate their results. Whereas I only include students who were in the senior follow-up studies, some other studies include all students who participated in the base year surveys, augmenting their sample by 15,000 students in the HSB data.

The patterns found in the HSB and NELS data resemble those in the longer-term NAEP data. Table 3 presents test score gaps for 17-year olds in the NAEP data for several years between 1978 and 1996. As with HSB and NELS, the NAEP data show that the Hispanic-white gap closed by much more than the black-white gap between 1982 and 1992. However, the NAEP test score gaps closed by roughly 0.05 standard units more than the other data sets indicate. The gaps in NAEP also tend to be slightly larger than those in HSB and NELS. These differences may arise due to sampling issues in NAEP, differences in test content, and the different ages of the sample members (NAEP surveys 17 year olds, which means that its sample contains a large proportion of high school juniors, whereas HSB and NELS surveys sophomores and seniors). Nonetheless, NAEP results are sufficiently similar to mine to give assurances that the trends measured using HSB and NELS are not due to peculiarities in these datasets' test measures or timing.

Table 3
NAEP Mathematics Test Score Gaps for 17 Year Olds
(Measured in White Standard Deviations)

	1978	1982	1986	1990	1992	1994	1996
Black-White Gap	1.16	1.05	0.99	0.71	0.92	0.94	0.96
Hispanic-White Gap	0.92	0.89	0.84	0.88	0.69	0.75	0.76

Note: The data are from the table entitled *NAEP 1996 Long-Term Mathematics Results - Data Almanacs for Age 17 Student Data, Weighted Average Scale Scores (Mean), Standard Deviations, and Percentile Distributions* from the National Center for Education Statistics' website. I compute test score gaps using equation (1).

4. Curriculum's Effect on Test Scores

Although my ultimate model encompasses a wide range of explanatory factors, the following model provides a foundation by combining the 1982 HSB and 1992 NELS data to obtain differences in test scores by ethnicity. This baseline model is:

$$TS_i = \alpha + \beta_1 B_i + \beta_2 H_i + \beta_3 D92_i + \beta_4 (B_i * D92_i) + \beta_5 (H_i * D92_i) + \beta_6 R_i + \varepsilon_i, \qquad (3)$$

where TS_i is student *i*'s test score measured as the number of white test score standard deviations student *i* scored away from the white test score mean. The variable B_i equals one if student *i* is black; it equals zero otherwise. The variable H_i equals one if student *i* is Hispanic; it equals zero otherwise. The variable $D92_i$ equals one if student *i* is from the 1992 senior cohort; it equals zero if the student is from the 1982 cohort. The interaction terms, $(B_i * D92_i)$ and $(H_i * D92_i)$, equal one if both components equal one; otherwise, they equal zero. Dichotomous variables for the Asian, American Indian, and other ethnic categories, as well as the interaction of those variables with $D92_i$, are represented by the vector R_i.

This simplified model restates the variation in mean test scores between different ethnic groups of students.[6] Because test scores are measured as the number of white standard deviation units away from the white mean test score, by construction the mean white test score equals zero in each year (i.e., α and β_3 are zero). The test score gap for 1982 is measured as the coefficient on a particular ethnicity variable. For example, β_2 indicates the average number of standard deviation units Hispanic students scored away from white students in 1982. A particular ethnic test score gap in 1992 is equal to the

[6] Estimating model (3) is equivalent to estimating two separate models for 1982 and 1992 data in which test scores are a function of a vector of ethnic dummy variables. Model (3) can be interpreted as a 1982 model to which I add a dummy variable indicating the year the data are from and a series of interaction terms interacting that dummy variable with every explanatory variable in the original model. Model (3) can be thought of as a differences in differences model.

coefficient on the ethnicity dummy variable plus the coefficient on that ethnicity variable's interaction term. It follows that the coefficient on the interaction term (i.e., β_4 or β_5) measures how the test score gap changed between 1982 and 1992. I focus on how these coefficients change as more controls are added to the model.

I add demographic, family, school, and curriculum characteristics. Sequentially adding theses factors allows me to assess what portion of the gaps within each year and what portion of the change in the gaps are accounted for by particular factors. To the extent that changes in an explanatory factor from 1982 to 1992 led to a narrowing test score gap, β_4 and β_5 should approach zero when that explanatory factor is added. I focus mainly on what happens to the gaps once the curriculum variables are added to the model. The order in which explanatory factors are added to the model has implications for how much of the test score gap they can explain. Adding curriculum last arguably yields a lower bound of its effect, because the demographic, family, and school variables may have already reduced the gap. Adding curriculum before adding any other factors should yield an upper bound of its effect on test score. In this case, some of the effects attributed to curriculum are most likely due to its correlation with other explanatory factors. I present results for both of these extreme cases.

The fully specified model is:

$$\begin{aligned} TS_i = \alpha + \beta_1 B_i + \beta_2 H_i + \beta_3 D92_i + \beta_4 (B_i * D92_i) + \beta_5 (H_i * D92_i) \\ + \beta_6 R_i + \beta_7 Demo_i + \beta_8 Fam_i + \beta_9 Sch_i + \beta_{10} Curric_i + \varepsilon_i \,, \end{aligned} \quad (4)$$

where $Demo_i$ is a vector of student *i*'s demographic characteristics as well as each of those variables interacted with the 1992 dummy variable. Similarly, Fam_i is a vector of student *i*'s family characteristics and each of those variables interacted with the 1992 dummy variable. The vector Sch_i contains school characteristics describing student *i*'s high school as well as each of those variables interacted with the dummy variable for 1992. $Curric_i$ is a vector of the math credits student *i* earned in each of seven math courses, as well as those variables interacted with the 1992 dummy variable. Unlike the previous section, which focused on the highest math course a student took, the seven math curriculum variables in the regression analysis are not mutually exclusive. Interaction terms are included so that the effect of any control variable can change between 1982 and 1992.[7]

To explain why test score differences between minority and white students narrowed, it is important to consider how test score gaps may be the result of underlying family,

[7] Demographic variables include age as a sophomore and gender. Family characteristics include six categories of parental income, six categories of parental education, and the number of siblings. School characteristics include the student-teacher ratio, the days in the school year, the percentage of teachers with a master's degree, the percentage of disadvantaged students in the school, average spending per pupil in the school district, the school's geographic region, and the urbanicity of the school. The appendix provides a crosswalk between the HSB and NELS variables.

school, and student characteristics that are correlated with both ethnicity and test scores.[8] For example, suppose that students who come from families with high incomes or high parental education levels receive instruction and support at home that enables them to perform well on tests. If black and Hispanic students are overly represented at low levels of family income and parental education relative to white students, then the perceived relationship between ethnicity and test scores may really stem from this difference in family background rather than from ethnic differences in ability or educational experience. Similarly, suppose that better school resources such as smaller class sizes and more educated teachers explain test scores. If, relative to white students, black and Hispanic students tend to go to schools that are more crowded or have a less educated teaching force, ethnicity and test scores will appear to be related if school characteristics are not taken into account.

Any study of test scores should address the effect that omitting the student's ability from the model could have on the interpretation of the results. If the goal is simply to understand the effect that curriculum has on test scores in any given year, then it is important to keep in mind that some of the effect that model (4) estimates could be due to variations in underlying student abilities. I test the robustness of my results by also estimating models that control for ability using the student's math grade point average (GPA). However, even without controlling for math GPA, omitted ability bias may not affect the results of this study. Assuming that ability does not vary across ethnicities within a given year, the difference in test scores between two ethnic groups is unaffected by omitted ability bias. Further, even if there is variation in omitted ability across ethnicities, changes in the test score gap over time will be unaffected as long as the distribution of abilities across and within ethnicities is constant throughout the 1980s.

4.1 The Relationship Between Math Curriculum and Test Scores

The relationship between math curriculum and test score gaps is best understood in the context of how math courses are related to math test scores. Table 4 shows the estimated coefficients and standard errors of the curriculum variables from the fully specified model in (4).[9] The first column shows the 1982 curriculum coefficients, the second column shows the change in the coefficients between 1982 and 1992, and the third column shows the 1992 effects (computed as the sum of the first two columns). The effects in 1982 and 1992 are the estimated gains in test scores, as measured by white standard deviation units, from completing an additional credit in the particular math course.

[8] One extreme view of test score gaps is that they are caused by ethnic variation in cognitive abilities. Racial bias in testing represents the other extreme. With either interpretation, it is necessary to control for the characteristics listed above.

[9] Complete regression results are located in appendix Table A.1.

Table 4
The Effect of an Additional Math Credit on Test Scores
(Measured in Standard Deviation Units)

	Without GPA			With GPA		
	1982	Change	1992	1982	Change	1992
Vocational Math	-0.042 **	-0.131 **	-0.173	-0.090 **	-0.154 **	-0.244
	(0.009)	(0.017)		(0.009)	(0.017)	
Pre-Algebra	0.171 **	-0.208 **	-0.037	0.133 **	-0.217 **	-0.084
	(0.012)	(0.020)		(0.012)	(0.020)	
Algebra/Geometry	0.352 **	-0.130 **	0.222	0.306 **	-0.148 **	0.158
	(0.009)	(0.016)		(0.009)	(0.016)	
Intermediate Algebra	0.496 **	-0.090 **	0.407	0.415 **	-0.117 **	0.299
	(0.015)	(0.024)		(0.015)	(0.024)	
Trigonometry	0.479 **	0.013	0.492	0.390 **	-0.040 *	0.350
	(0.014)	(0.023)		(0.014)	(0.024)	
Pre-Calculus	0.653 **	0.012	0.665	0.526 **	-0.048	0.478
	(0.033)	(0.042)		(0.032)	(0.041)	
Calculus	0.662 **	0.071 *	0.733	0.539 **	0.018	0.557
	(0.030)	(0.042)		(0.029)	(0.042)	
Math GPA	---	---	---	0.198 **	0.082 **	0.280
				(0.008)	(0.014)	

Notes: ** indicates significance at the 5 percent level; * indicates significance at the 10 percent level. Standard errors are omitted from column three because the significance of the 1992 value is determined by the significance of the change from 1982 to 1992. The R-Squared is 0.625 without GPA in the model and 0.650 with GPA in the model. There are 8,140 observations from HSB and 9,610 from NELS. When GPA is added to the model, 46 observations are lost due to missing data. All models control for demographic, family, and school characteristics.

In 1982, a student who took an additional algebra/geometry course is expected to score 0.35 standard deviation units higher than the average student. All else equal, a student who completed one additional credit in each of algebra/geometry, intermediate algebra, pre-calculus, and calculus is predicted to score about 2.20 standard units higher than an average student who did not complete these additional math courses. This effect is quite large compared with the existing ethnic test score gaps, which Table 2 shows range from 0.65 to 0.89 of a standard deviation.

Although the estimated effects of math courses increase uniformly with the rigor of the given math course in both 1982 and 1992, the magnitude of the effects varies across these two years. Vocational math through intermediate algebra are predicted to have significantly lower effects on test scores in 1992 as compared to 1982. Two explanations for this drop seem plausible. First, such a change could result from the technical change that the math test underwent. Because the NELS test eliminated some of the floor and ceiling effects that were present in 1982, its larger scope could have automatically reduced the value of taking lower level math courses for performance on the test.

Second, suppose that math courses act as a signal of ability to college admission offices. As high-ability students saw an influx of low ability students into their academic courses, they might have perceived the need to take even more advanced math courses to set themselves apart from their newfound peers.[10] If, on average, the high-ability students took more advanced math courses in 1992, the math courses would be expected to have larger coefficient to the extent that taking the courses is correlated with ability. Whereas in 1982 very high ability might have manifested itself in the completion of the mid-level courses such as algebra/geometry or intermediate algebra, by 1992 high ability might have signaled itself through the completion of courses at the trigonometry level or higher. If this were the case, higher-level math courses would have a stronger relationship with test scores in 1992.

To mitigate such ability effects, I add the student's math GPA to the model. The second trio of columns in Table 4 shows the results. Controlling for GPA, the predicted math effects drop but are still large and significant. The coefficients on vocational math, pre-algebra, and algebra/geometry drop by less than 0.05 in 1982. The intermediate algebra and trigonometry coefficients drop by slightly more than 0.08, whereas the coefficients on pre-calculus and calculus drop by just over 0.12. To the extent that ability is correlated with taking higher-level math courses, it is not surprising that the coefficients on the higher-level math courses drop more than those on the lower level math courses once the model includes GPA.[11]

The predicted curriculum effects dwarf the predicted effects of the family and school characteristics.[12] The full set of regression results is located in Table A.1. Recall that a student who completes one additional class in each of algebra/geometry, intermediate algebra, pre-calculus, and calculus is predicted to score about 2.2 standard units higher than a comparable student who does not complete any of these math courses. In contrast, coming from a family not in the lowest income bracket is expected to increase a student's test scores by about 0.12 to 0.15 standard test score units in 1982. These family effects drop by about 0.08 to 0.10 units in 1992. Having a mother who has at least some education beyond a high school diploma is expected to increase test scores by 0.09 to 0.12 standard units, regardless of the student's cohort. In 1982, having a father who completed a higher degree is expected to increase test scores by about 0.2 standard units -

[10] They may also have sensed this need if they saw their lower-academic courses being watered down.

[11] To test whether the predicted curriculum effects are merely the result of a high correlation between the type of curriculum students take and their drop-out behavior, I estimated a version of model (4) that includes a dummy variable (as well as its interaction with the 1992 dummy variable) indicating whether the student dropped out of high school before grade 12. Including this new variable barely changes the curriculum coefficients. Furthermore, excluding dropouts entirely only minimally changes the results. This gives some reassurances that any bias resulting from the higher rate of missing data for high school dropouts in the 1992 NELS data is minimal. In the 1980 data, test scores are missing for approximately 12 percent of both high school graduates and high school dropouts. For the 1990 data, 16 percent of the test score data are missing for high school graduates, whereas about 62 percent of the test score data are missing for high school dropouts.

[12] This finding is consistent with that of Lee et al. (1998).

about the same effect as a pre-algebra course, but not as much as any of the advanced math courses. As substantial as the paternal education effects appear, they drop by about 0.15 units in 1992.

The only non-geographic school characteristic that is statistically significant at the 5 percent level is the percentage of disadvantaged students in a student's school. However, its predicted effect on test scores is small. In both 1982 and 1992, a two standard deviation increase in the percentage of disadvantaged students is associated with a decline in test scores of 0.1 standard units. Clearly, curriculum is much more strongly associated with test scores than are the other measurable factors.

4.2 The Relationship Between Curriculum and the Hispanic-White Test Score Gap

Table 5 demonstrates how the Hispanic-white test score gap changed between 1982 and 1992. The first column, which is based on equation (3), shows that without conditioning on any other variables, Hispanic students scored an average of 0.80 standard units less than white students in 1982, but they only scored 0.65 standard deviation units less in 1992. In other words, the test score gap narrowed by 0.15 units. This change is statistically significant at the 5 percent level.

Table 5
The Hispanic-White Test Score Gap

	(1)	(2)	(3)	(4)	(5)	(3a)	(4a)
1982 Gap	-0.804 **	-0.553 **	-0.502 **	-0.316 **	-0.425 **	-0.404 **	-0.283 **
1992 Gap	-0.652 **	-0.285 **	-0.236 **	-0.224 **	-0.364 **	-0.120 **	-0.169 **
Change in Gap	0.151 **	0.268 **	0.265 **	0.092 **	0.061 *	0.284 **	0.114 **
Standard Error	(0.050)	(0.048)	(0.050)	(0.037)	(0.034)	(0.045)	(0.036)
Controls:							
Ethnicity	X	X	X	X	X	X	X
Demographic		X	X	X		X	X
Family		X	X	X		X	X
School			X	X		X	X
Curriculum				X	X		X
Math GPA						X	X
R-Squared	0.116	0.288	0.302	0.625	0.591	0.449	0.650

Notes: ** indicates significance at the 5 percent level and * indicates significance at the 10 percent level. An "X" indicates that the control is in the model. The sample size is 17,750 (8,140 from HSB plus 9,610 from NELS) in columns 1 through 4. Because of missing GPA data, columns 3a and 4a have 46 fewer observations. The regression is weighted by the appropriate senior follow-up weight from each dataset. All of the standard errors are less than 0.05 for the 1982 gaps. The change in gaps is the 1992 gap minus the 1982 gap and measures how much the 1982 test score gap has narrowed. The standard error refers to the standard error of this change.

Column 2 adds demographic and family characteristics to the previous model. Accounting for these factors in 1982, Hispanic students are predicted to score 0.55 standard units lower than white students. This gap is smaller than the baseline of 0.80 units in column 1, indicating that these factors explain 31 percent of the baseline gap. In 1992, the baseline gap of 0.65 drops to 0.29 once the demographic and family controls are added, indicating these factors explain 56 percent of the test score gap in that year. Accounting for family characteristics, then, it appears as though the test score gap narrowed from 0.55 standard units in 1982 to 0.29 in 1992, constituting a reduction of 0.27 standard units in the test score gap.

This change in the gap is larger than the 0.15 unit reduction in the baseline model, because family characteristics account for more of the 1992 gap than the 1982 gap. In other words, the test score gap converged more than it would have had the effect of family characteristics remained the same. This change in the effect of family characteristics stems from two main sources. First, the coefficients on the family characteristics variables change between 1982 and 1992; second, the family characteristics themselves changed over the decade. For example, the gap in real family income between Hispanic and white students grew between 1982 and 1992. More Hispanic students were from lower- income families and more white students were from higher-income families in 1992 relative to 1982. Similarly, the gap in parental education for Hispanic and white students was more pronounced in 1992. Whereas Hispanic students were one and a half times more likely than white students to have a father without a high school degree in 1982, a decade later they were three times more likely than white students to be in that situation. Because the family characteristics account for some of the variation in test scores, and because the ethnic gap in family characteristics widened throughout the 1980s, these variables account for more of the test score gap in 1992. This serves to magnify the change in the unexplained portion of the Hispanic-white test score gap from 1982 to 1992.

The third column in Table 5 adds school characteristics to the previous model. Adding these factors causes the unexplained portion of both the 1982 gap and the 1992 gap to fall by the same amount. Therefore, the narrowing of 0.27 in the unexplained Hispanic-white test score gap between 1982 and 1992 is nearly identical to that in the previous column.

The model in the fourth column indicates that curriculum can explain much of the 0.27 narrowing in the test score gap from the previous column. Adding curriculum to the model reduces the 1982 test score from 0.50 to 0.32, a reduction of 37 percent. However, curriculum causes the 1992 gap to shrink by only 5 percent, from 0.24 to 0.22. Curriculum accounts for more of the 1982 gap, because curriculum varied more across Hispanic and white students in 1982 than it did in 1992 (and the more an explanatory factor varies, the more variation in the dependent variable it can explain). Recall from Table 1 that the percentage of Hispanic students completing an algebra course or higher was much closer to the completion rate for white students in 1992 compared to 1982.

Had course-taking patterns not changed between 1982 and 1992, and had curriculum affected test scores in the same way in both years, adding curriculum to the model would have caused the 1992 gap to drop as much as the 1982 gap dropped. In this case, the change in the gap between 1982 and 1992 would have been the same as in column 3, and curriculum would have explained none of the convergence in the gap. However, because curriculum changed, the unexplained test score gap narrows from 0.32 to 0.22. Thus, the convergence in the unexplained test score gap with curriculum in the model is 0.09 rather than 0.27 without curriculum, indicating that that curriculum accounts for two-thirds of the convergence in the test score gap in column 3.

The decline in the test score gap in either 1982 or 1992 that occurs between column 3 and column 4 represents a lower bound on the predicted effect of curriculum in either year, because both models already control for family and school characteristics that might be positively correlated with curriculum. At the other extreme, adding curriculum to the baseline model presented in column 1, before adding the family and school factors, yields an upper bound on the predicted effect of curriculum in either year.[13] Column 5 shows the results from adding only curriculum to the baseline model in column 1. Adding curriculum without any other controls in the model causes the 1982 gap to shrink from 0.80 to 0.43, a reduction of 47 percent. The 1992 gap shrinks from 0.65 to 0.36, a 44 percent reduction. In the column 5 model, the change in the unexplained gap from 1982 to 1992 is 0.06, down from 0.15, suggesting that the convergence in curriculum is associated with a 60 percent convergence in the test score gap. Interestingly, the earlier analysis that accounted for ethnic, demographic, and school characteristics leads to a similar conclusion: in proportional terms, a converging curriculum is estimated to account for nearly two-thirds of the narrowing in the gap between white and Hispanic students. However for the model in column 5, the change between the 1982 and 1992 gap is no longer significant at the 5 percent level. Thus, for this model, we cannot reject the null hypothesis that curriculum explains the entire convergence in test scores.

Once again, it is important to consider the role that ability could be playing in these results. To isolate curriculum's effects, columns 3a and 4a duplicate columns 3 and 4, respectively, but they also include the student's math GPA as a control for ability. Although adding GPA to column 3 causes both the 1982 gap and the 1992 gap to fall, the difference between the 1982 and the 1992 gap is 0.28, nearly identical to its value in column 3. The same story holds when adding GPA to column 4. The Hispanic-white test score gap in each year falls from its column 4 value, but the 0.11 difference between the 1982 and 1992 gaps in column 4a is similar to the difference of 0.09 in column 4. Thus, even accounting for math GPA, adding curriculum to a model that already controls for demographic, family and school characteristics accounts for about 60 percent of the narrowing test score gap.

4.3 The Relationship Between Curriculum and the Black-White Test Score Gap

As Table 6 demonstrates, the black-white test score gap mimics the Hispanic-white gap in some ways. Family, school, and curriculum factors explain a large share of the black-white test score gap in both 1982 and 1992. However, regardless of the model specification, the change in the test score gap between 1982 and 1992 is quite small. For

[13] These extremes do not represent the lower and upper bounds on the *change* in the test score gap between 1982 and 1992. That difference would depend on how much curriculum lowers the gap in one year versus the other.

Table 6
The Black-White Test Score Gap

	(1)	(2)	(3)	(4)	(5)	(3a)	(4a)
1982 Gap	-0.874 **	-0.615 **	-0.511 **	-0.344 **	-0.511 **	-0.376 **	-0.289 **
1992 Gap	-0.806 **	-0.530 **	-0.487 **	-0.323 **	-0.434 **	-0.302 **	-0.244 **
Change in Gap	0.068	0.085 *	0.024	0.021	0.076 **	0.074 *	0.045
Standard Error	(0.048)	(0.045)	(0.048)	(0.036)	(0.033)	(0.043)	(0.035)
Controls:							
Ethnicity	X	X	X	X	X	X	X
Demographic		X	X	X		X	X
Family		X	X	X		X	X
School			X	X		X	X
Curriculum				X	X		X
Math GPA						X	X
R-Squared	0.116	0.288	0.302	0.625	0.591	0.449	0.650

Notes: ** indicates significance at the 5 percent level and * indicates significance at the 10 percent level. An "X" indicates that the control is in the model. The sample size is 17,750 (8,140 from HSB plus 9,610 from NELS) in columns 1 through 4. Because of missing GPA data, columns 3a and 4a have 46 fewer observations. The regression is weighted by the appropriate senior follow-up weight from each dataset. All of the standard errors are less than 0.08 for the 1982 gaps. The change in gaps is the 1992 gap minus the 1982 gap and measures how much the 1982 test score gap has narrowed. The standard error refers to the standard error of this change.

instance, in the baseline model that conditions solely on ethnicity, the black-white test score gap drops from 0.87 in 1982 to 0.81 in 1992, but the change of 0.07 is not statistically significant. Thus, there is very little convergence in test scores for family, school, and curriculum to explain.

Including demographic and family characteristics in the baseline model leads to about a 30 percent decrease in the black-white test score gap in both 1982 and 1992. Because the gap falls by the same amount in each year, the change in the gap is almost the same as in column 1. The convergence in the test score gap is still small and only borderline significant. Adding school characteristics leads to an additional 17 percent drop in the gap in 1982 but only an 8 percent drop in 1992 from the column 2 levels. Because of the unequal effect of school characteristics in these two different years, the change in the unexplained black-white gap from 1982 to 1992 narrows from 0.09 to 0.02 once these controls are added. The standard errors on these estimates are so large that it is unclear whether school characteristics really contribute to a narrowing gap.

Adding curriculum to the model explains about one third of the gap remaining in column 3 – in both 1982 and 1992. Consequently, the difference between the 1982 and 1992 gaps is quite similar in models both with and without curriculum, indicating that curriculum plays little role in explaining changes in the test score gap over time. Repeating the models in columns 3 and 4, but including GPA to control for the student's ability, indicates that curriculum may play a small role in narrowing the test score gap. In this pair of models, adding curriculum causes the change in the gap to drop from 0.07 to 0.05.

These results are quite different from those for Hispanic students. This difference may be surprising given that both black and Hispanic students improved their curriculum. In 1982, the test score gap was just over 0.8 for both black and Hispanic students. By 1992 this gap had dropped to 0.65 for Hispanic students, but the gap remained relatively unchanged for black students. Even after controlling for demographic, family, and school factors, the Hispanic gap decreased more between 1982 and 1992 than did the black gap (see column 3 of Table 5 and Table 6). Given that these factors do not explain why the overall results are different for black and Hispanic students, curriculum becomes the next likely candidate.

Starting with the column 3 models, curriculum can explain a large portion of the test score gap in 1982 for both black and Hispanic students. For both groups of students, the gap drops from about 0.5 to about 0.3 when curriculum is added. However, curriculum explains very little of the 1992 gap for Hispanic students, because curriculum patterns in 1992 were quite similar for Hispanic and white students. In contrast, curriculum explains a large share of the black-white test score gap in 1992. This explanatory power arises because the curriculum patterns of black and white students were not as similar as the patterns of Hispanic and white students. The biggest distinction is the percentage of students who took only vocational math as their highest course. In 1992, 9 percent of white students took only vocational math and 11 percent of Hispanic students took only that level of math. However, black students were twice as likely as white students to take only vocational math. In addition, the effect of vocational math courses on test scores was more negative in 1992 than in 1982. The combination of these two factors can partially explain why there has not been much change in the black-white test score gap.

Despite this uncertainty about curriculum's role in explaining the slight narrowing in the test score gap, its impact on test scores in a given year is important. To estimate an upper bound of its affect, column 5 presents results from adding curriculum to the baseline model from column 1. Adding curriculum before adding family and school characteristics causes the 1982 black-white test score gap to shrink 42 percent (from 0.87 to 0.51) and the 1992 test score gap to shrink 46 percent (from 0.81 to 0.43). By comparing the test score gap within a given year in columns 3 and 5, it is clear that the predicted effects of curriculum alone are equal to, or larger than, the predicted effects of family and school characteristics combined.

In summary, curriculum has a significant effect on test scores and can, depending upon the order in which explanatory factors are added to the model, account for 30 to 46 percent of the black-white test score gap in either year. However, given that the black-white test score gap did not significantly narrow between 1982 and 1992, little can be said about curriculum's effect on the change in the test score gap over time. Given the attention placed on the perceived narrowing of the black-white test score gap, these results indicate that it is probably important to look at long-term trends (for example using NAEP) rather than just limiting the data to two points in time. The HSB and NELS tests alone provide insufficient evidence that the black-white test score gap has really closed.

5. Conclusion

This paper reviews two very striking trends that emerged during the 1980s. First, students of all ethnicities took more advanced high-school mathematics courses in 1988-1992 when compared with 1978-1982, the periods when the HSB and NELS sophomore cohorts attended high school. However, black and Hispanic completion rates in academic math courses increased substantially more than those of white students. Although both of these underrepresented ethnic groups improved their curriculum, Hispanic students made larger gains than black students did in some areas. For example, even though the percentage of black students who took only vocational math dropped from 44 percent to 18 percent, in 1992 they were still twice as likely as white students to stop taking math at that level. In contrast, only 11 percent of Hispanic students stopped at vocational math. Second, during the same time span, the gap in test scores between white and Hispanic and between white and black students also narrowed. Once again, the gap narrowed more for Hispanic students than for black students.

Rather than simply noting these temporal consistencies, this paper picks up where others have left off and statistically tests whether the convergence in the type of math courses students completed has narrowed the ethnic test score gaps. Whereas the causal relationship between math courses and test scores can never be fully unraveled, my research suggests that changes in curriculum can explain about 60 percent of the narrowing in the test score gap between white and Hispanic students that occurred from 1982 and 1992. In a statistical sense, however, the black-white test score gap did not narrow significantly between 1982 and 1992, leaving a negligible change in the gap for curriculum to explain. These different effects for black and Hispanic students likely stem from the fact that curriculum patterns changed much more for Hispanic than for black students. In particular, the higher proportion of black students taking only vocational math, coupled with the more pronounced negative effect of that course, helps explain why the black-white test score gap did not change much between 1982 and 1992.

Even though curriculum does not appear to be related to a narrowing in the black-white test score gap, curriculum is strongly associated with test scores for black students in both 1982 and 1992. After controlling for many demographic, family, and school factors that could influence test scores, adding curriculum to a model of test scores leads to a large reduction in the unexplained test score gap in both 1982 and 1992. For instance, in 1982 adding curriculum leads to a 33 percent reduction in the black-white test score gap that remains after having previously accounted for demographic, family, and school characteristics; in 1992, it leads to a 34 percent reduction. The corresponding values for the Hispanic-white test score gap are 37 percent in 1982 and 5 percent in 1992. These values are much larger in models that do not first control for demographic, family, and school characteristics. Strikingly, adding math curriculum doubles the explanatory power of the models, even after controlling for a rich set of family, personal, and school characteristics.

The vast literature on the impact of school inputs and school demographics on student achievement provides a rather mixed picture of whether school inputs affect achievement (see Hanushek, 1996). My study suggests that at the high school level, measures of school quality such as the student-teacher ratio are not significantly linked to test scores, but the percentage of disadvantaged students in the school is linked to lower achievement. However, differences in math curriculum have a far greater impact on ethnic achievement gaps than do variations in school resources or school demographics.

The dramatic result that curriculum is related to the narrowing of the Hispanic-white test score gap does not necessarily suggest that policymakers can radically reduce ethnic gaps in student achievement simply by prescribing identical courses for all students. Unobserved factors undoubtedly explain some of the covariation in test scores and curriculum. At the same time, the robustness of my central conclusions to a rich set of demographic variables and school characteristics suggests that curriculum enrichment for all students could yield substantial improvements in student outcomes. Moreover, variations in curriculum appear to be far more important determinants of achievement than are variations in school resources that have dominated some of the earlier debate. This finding perhaps signals some possible changes in direction for future research on the determinants of school quality. Policies focusing on increasing the number of minorities enrolled in a rigorous math curriculum may be more successful and cost effective than school reforms aimed at reducing class sizes and welfare programs aimed at increasing socioeconomic status.

6. Appendix

This appendix documents the math course classification system, and it describes the variables used in this paper. The actual variable names from the HSB data set are in capital letters.

6.1 Academic Variables

Math Course Classification System. The NCES provided me with the math course classification system that it used to construct "pipeline composite variables" for the NELS:88 data.[14] These provided the backbone of my classification system. Following is a brief description of each category. The descriptive labels I use throughout the text are followed by the official NCES label in parentheses. Vocational courses (non-academic) include math courses labeled as general, basic, consumer, technical, vocational, or review. Pre-algebra (low-academic) includes pre-algebra, algebra 1 (part 1), algebra 1 (part 2), and informal geometry. Algebra/geometry (middle-academic I) includes algebra 1, geometry (plane and solid), unified 1, and unified 2. Intermediate algebra (middle-academic II) includes algebra 2 and unified 3. Trigonometry (advanced I) includes algebra 3, algebra-trigonometry, analytic geometry, linear algebra, probability, and statistics. Pre-calculus (advanced II) includes only pre-calculus course. Calculus (advanced III) includes calculus (regular and AP) and calculus-analytic geometry.

Math Test Scores. From HSB I used IRT82M (the 1992 IRT calculated score) which was a score for the sophomore cohort two years after their sophomore year. For the NELS:88 data, I used F22XMIRR (Mathematics IRT-Estimated Number Right) which was also a score from the senior year. Personal communication with Judith Pollack of the NCES validated the comparability of these test scores (see Rock and Pollack, 1995, for affiliation).

Grade Point Average and Number of Credits Earned. In both NELS and HSB, I computed the student's math GPA on a scale of 0 to 4.3. I took a weighted average of the student's grade points for each course, where the weights were the number of credits that the student earned for the class. These credits were either 0.25, 0.33, 0.5, or 1, depending on whether the course length was a quarter, trimester, semester, or, more commonly, a year-long course. I converted letter grades to grade points. An "A" received 4 points, a "B" received 3 points, a "C" received 2 points, a "D" received 1 point, and an "F" received 0 points. I added 0.3 points for a "plus" and deducted 0.3 points for a "minus." For example, a "B+" received 3.3 points.

If a student failed a course, but subsequently retook it and passed it, I included the letter grade and credit information from the successful completion of the course in the student's GPA but ignored the information from the failed attempt (it is likely that the *final* level of success in a given course is likely to have the greater effect on a student's academic outcomes). If, however, the student failed a course and did not repeat it successfully, I did include the information in the GPA calculation.

[14] I would like to thank Jeff Owings and Robert Atanda at NCES for providing me with the course classifications.

The credit values used to compute the GPA also served as my primary measures of math curriculum, with one exception. I did not include credits earned in failed courses in the student's tally of math credits. So, for example, a student who passed a one-year calculus course had a tally of 1 for that course category. A student who passed three semesters worth of a geometry course and failed one semester had a tally of 1.5 for that category. In a few cases, students received a grade of "pass" in a course. I did not include these grades in the GPA calculations but did add the number of credits earned to the tally of credits earned by the student.

6.2 Demographic, Family, and School Variables

Race and Gender. From NELS, I used F2RACE1 and from HSB, I used the composite variable RACE. I used the variable F2SEX from NELS and the variable SEX from HSB.

Age. To compute the age of the respondent in NELS, I used the student's birth month and year as recorded by F2BIRTHM and F2BIRTHY. I calculated the age of the student in years (and fractions thereof) as of June 15, 1990 - the date when they should have been high school sophomores. Because no birthday is given, I assumed that they were born on the 15th of whatever month is recorded for them. From HSB, I calculated age as of June 15, 1980 using the series of birthday-related questions that were asked in the base year as well as in the follow-up surveys. When the birthdays were not the same for the different waves of the survey, I used the birthday from the earliest wave.

Family Income. For the NELS sample, I computed the family income of the student using BYFAMINC. I divided these into income categories that corresponded to the HSB family income categories inflation adjusted to 1987 dollars. If BYFAMINC was missing (as it always is for those in the freshened sample), I used F2P74. The latter variable is not adjusted to 1987 dollars, because doing so makes it impossible to properly align income categories. From HSB, I used the 1980 student-reported family income category BB101 (and inflation adjusted it to 1987 dollars). Where this value was missing, I instead used FAMINC, another student-reported family income category obtained after 1980. Because the income categories from these two separate years did not coincide exactly, I condensed them into six better-aligned income categories. A difference between the HSB and NELS family income measure worth noting is that NELS data are provided by the parent, whereas the HSB data are provided by the student. By using a sub-sample of the HSB data for which parental responses are available, I found that the student-reported data are highly correlated with the parent-reported data.

Parental Education. I included a series of dummy variables indicating the highest level of education attained by both the father and mother of the student respondent. I used the student reported data (F1N20A and F1N20B) rather than the parent reported data from NELS so that the data are more comparable to the HSB student reported data (parent reported data were only available for small portion of HSB students). From HSB, I used BB039 and BB042.

Siblings. From NELS, I computed the number of siblings the respondent has by summing up the non-missing values of the second follow-up variables F2N9A, F2N9B, F2N10A, and F2N10B (these variables represent the number of older brothers and sisters and the number of younger brothers and sisters that the respondent has). When all of those variables were missing, I used the sum of the non-missing values of the first follow-up variables F1S90A, F1S90B, F1S91A, and F1S91B. If F1S89 indicated the respondent has a twin, I added one to my previously computed sibling value. I capped the number of siblings at 6 to be comparable to the HSB data. From HSB, I used BB096A-E, FY106-09, and TY5A-M. I used the number of siblings calculated from the 1982 questions, but where that was missing, I used the number calculated from the 1986 variables.

School Characteristics. To compute the length of the school year, I used F1C7 from NELS and SB005 from HSB. To compute the percentage of teachers with a master's degree in NELS, I divided F1C44C (the number of teachers with a master's) by F1C35 (the number of teachers). I used SB042 from HSB. To compute the student-teacher ratio in NELS, I divided F1C2 by F1C35; in HSB, I divided SB002A by SB039C. Finally, F1C30A from NELS and SB037 from HSB represented the percentage of disadvantaged students.

Table A.1
Test Score Regressions: Dependent Variable is Test Score Measured in Standard Deviation Units

	(1)	(2)	(3)	(4)	(5)
Hispanic	-0.803 **	-0.553 **	-0.502 **	-0.316 **	-0.425 **
	(0.028)	(0.026)	(0.027)	(0.020)	(0.019)
Black	-0.874 **	-0.615 **	-0.511 **	-0.344 **	-0.511 **
	(0.028)	(0.027)	(0.029)	(0.021)	(0.020)
Asian	0.291 **	0.343 **	0.334 **	0.032	0.001
	(0.082)	(0.074)	(0.075)	(0.055)	(0.056)
American Indian	-0.656 **	-0.307 **	-0.264 **	-0.145 **	-0.269 **
	(0.083)	(0.075)	(0.075)	(0.055)	(0.057)
Other Race	-0.527 **	-0.071	-0.059	0.028	-0.126
	(0.149)	(0.135)	(0.134)	(0.099)	(0.101)
1992 Dummy Variable	0.000	1.449 **	2.271 **	0.474	-0.071 **
	(0.017)	(0.432)	(0.912)	(0.673)	(0.035)
1992 * Black	0.068	0.085 *	0.024	0.021	0.076 **
	(0.048)	(0.045)	(0.048)	(0.036)	(0.033)
1992 * Hispanic	0.151 **	0.268 **	0.265 **	0.092 **	0.061 **
	(0.050)	(0.048)	(0.050)	(0.037)	(0.034)
1992 * American Indian	-0.252 *	-0.298 **	-0.267 **	-0.098	-0.097
	(0.142)	(0.129)	(0.130)	(0.096)	(0.097)
1992 * Asian	-0.186 *	-0.201 **	-0.165 *	-0.116 *	-0.108
	(0.104)	(0.094)	(0.095)	(0.070)	(0.071)
1992 * Race Other	-0.368	-0.532	-0.474	-0.141	-0.105
	(0.702)	(0.632)	(0.627)	(0.460)	(0.478)
Vocational Credits				-0.042 **	-0.037 **
				(0.009)	(0.009)
Pre-Algebra Credits				0.171 **	0.210 **
				(0.012)	(0.013)
Algebra/Geometry Credits				0.352 **	0.396 **
				(0.009)	(0.009)
Intermediate Algebra Credits				0.496 **	0.538 **
				(0.015)	(0.015)
Trigonometry Credits				0.479 **	0.537 **
				(0.014)	(0.015)
Pre-Calculus Credits				0.653 **	0.719 **
				(0.033)	(0.034)
Calculus Credits				0.662 **	0.788 **
				(0.030)	(0.031)
1992 * Vocational Credits				-0.131 **	-0.148 **
				(0.017)	(0.018)
1992 * Pre-Algebra Credits				-0.208 **	-0.242 **
				(0.020)	(0.021)
1992 * Algebra/Geometry Credits				-0.130 **	-0.150 **
				(0.016)	(0.016)
1992 * Intermediate Algebra Credits				-0.090 **	-0.113 **
				(0.024)	(0.024)
1992 * Trigonometry Credits				0.013	0.007
				(0.023)	(0.024)
1992 * Pre-Calculus Credits				0.012	0.013
				(0.042)	(0.043)
1992 * Calculus Credits				0.071 *	-0.008
				(0.042)	(0.043)

Continued …

Table A.1 (Continued)

	(1)	(2)	(3)	(4)	(5)
Male Dummy Variable		0.131 **	0.126 **	0.130 **	
		(0.017)	(0.017)	(0.012)	
Sophomore Year age		-0.330 **	-0.311 **	-0.173 **	
		(0.015)	(0.015)	(0.011)	
Number of Siblings		-0.022 **	-0.022 **	-0.007 *	
		(0.005)	(0.005)	(0.004)	
Number of Sibs Missing		-0.129 *	-0.080 **	0.005 **	
		(0.076)	(0.076)	(0.056)	
Family Income $10K-$20K		0.286 **	0.271 **	0.128 **	
		(0.035)	(0.035)	(0.026)	
Family Income $20K-$27K		0.387 **	0.369 **	0.128 **	
		(0.037)	(0.037)	(0.027)	
Family Income $27K-$34K		0.415 **	0.388 **	0.145 **	
		(0.038)	(0.038)	(0.028)	
Family Income $34K-$51K		0.411 **	0.383 **	0.131 **	
		(0.039)	(0.039)	(0.029)	
Family Income $51K or Higher		0.405 **	0.373 **	0.119 **	
		(0.043)	(0.043)	(0.032)	
Family Income Missing		-0.037	-0.083 *	0.003	
		(0.043)	(0.043)	(0.031)	
Mother's Ed: Less that High School		-0.067 **	-0.034	0.030	
		(0.028)	(0.028)	(0.020)	
Mother's Ed: Some College		0.133 **	0.142 **	0.093 **	
		(0.026)	(0.026)	(0.019)	
Mother's Ed: College Graduate		0.291 **	0.293 **	0.123 **	
		(0.040)	(0.040)	(0.029)	
Mother's Ed: M.A., Ph.D.		0.352 **	0.339 **	0.111 **	
		(0.051)	(0.051)	(0.037)	
Mother's Ed: Missing, Unknown		-0.143 **	-0.132 **	-0.045 **	
		(0.027)	(0.026)	(0.019)	
Father's Ed: Less that High School		-0.162 **	-0.142 **	-0.032	
		(0.029)	(0.029)	(0.021)	
Father's Ed: Some College		0.223 **	0.227 **	0.084 **	
		(0.029)	(0.028)	(0.021)	
Father's Ed: College Graduate		0.471 **	0.469 **	0.199 **	
		(0.038)	(0.038)	(0.028)	
Father's Ed: M.A., Ph.D.		0.449 **	0.454 **	0.217 **	
		(0.043)	(0.043)	(0.032)	
Father's Ed: Missing, Unknown		-0.018	-0.016	-0.004	
		(0.027)	(0.027)	(0.020)	
Suburban			0.033	0.016	
			(0.024)	(0.017)	
Rural			0.018	0.058 *	
			(0.025)	(0.019)	
New England			0.124 **	-0.055 *	
			(0.042)	(0.031)	
Middle Atlantic			0.132 **	0.040	
			(0.034)	(0.026)	
South Atlantic			-0.153 **	-0.165 **	
			(0.034)	(0.025)	
East South Central			-0.164 **	-0.190 **	
			(0.042)	(0.031)	

Continued …

Table A.1 (Continued)

	(1)	(2)	(3)	(4)	(5)
West South Central			-0.182 **	-0.251 **	
			(0.037)	(0.027)	
East North Central			-0.038	-0.067 **	
			(0.033)	(0.024)	
West North Central			0.111 **	-0.017 *	
			(0.039)	(0.029)	
Mountain			-0.085 *	-0.094 **	
			(0.047)	(0.034)	
Days In School Year			0.003	0.002 *	
			(0.002)	(0.001)	
Days In School Year Missing			0.555 *	0.388 *	
			(0.314)	(0.230)	
% Teachers with Master's Degree			-0.018	0.049 *	
			(0.039)	(0.029)	
% Teachers with Master's Missing			0.027	0.033	
			(0.053)	(0.039)	
% Disadvantaged Students			-0.244 **	-0.241 **	
			(0.049)	(0.037)	
% Disadvantaged Students Missing			-0.087 **	-0.019	
			(0.032)	(0.024)	
Student Teacher Ratio			-0.0001	-0.0001	
			(0.0001)	(0.0001)	
Student Teacher Ratio Missing			-0.027	-0.046 *	
			(0.032)	(0.024)	
1992* Male		-0.024	-0.011	0.024	
		(0.028)	(0.027)	(0.020)	
1992* Age		-0.092 **	-0.114 **	0.010	
		(0.026)	(0.026)	(0.020)	
1992* Siblings		0.009	0.009	0.010	
		(0.009)	(0.009)	(0.006)	
1992* Siblings Missing		-0.186 **	-0.195 **	-0.079	
		(0.091)	(0.091)	(0.067)	
1992* Family Income $10K-$20K		-0.147 **	-0.154 **	-0.113 **	
		(0.057)	(0.057)	(0.042)	
1992* Family Income $20K-$27K		-0.179 **	-0.198 **	-0.111 **	
		(0.063)	(0.062)	(0.046)	
1992* Family Income $27K-$34K		-0.176 **	-0.182 **	-0.106 **	
		(0.059)	(0.059)	(0.043)	
1992* Family Income $34K-$51K		-0.113 *	-0.123 **	-0.078 *	
		(0.059)	(0.059)	(0.044)	
1992* Family Income $51K or Higher		-0.004	-0.013	-0.049	
		(0.064)	(0.064)	(0.047)	
1992* Family Income Missing		-0.036	-0.024	-0.100 *	
		(0.076)	(0.075)	(0.055)	
1992* Mother's Ed: Less than High School		-0.020	-0.036	-0.021	
		(0.048)	(0.048)	(0.035)	
1992* Mother's Ed: Some College		0.123 **	0.109 **	0.003	
		(0.041)	(0.041)	(0.030)	
1992* Mother's Ed: College Graduate		-0.009	-0.028	-0.012	
		(0.056)	(0.056)	(0.041)	
1992* Mother's Ed: M.A., Ph.D.		-0.098	-0.093	0.016	
		(0.070)	(0.069)	(0.051)	

Continued …

Table A.1 (Continued)

	(1)	(2)	(3)	(4)	(5)
1992* Mother's Ed: Missing, Unknown		0.135 **	0.126 **	0.051	
		(0.051)	(0.051)	(0.038)	
1992* Father's Ed: Less than High School		0.012	0.008	-0.000	
		(0.049)	(0.049)	(0.036)	
1992* Father's Ed: Some College		-0.133 **	-0.133 **	-0.091 **	
		(0.044)	(0.044)	(0.032)	
1992* Father's Ed: College Graduate		-0.186 **	-0.177 **	-0.148 **	
		(0.055)	(0.054)	(0.040)	
1992* Father's Ed: M.A., Ph.D.		-0.000	-0.020	-0.137 **	
		(0.062)	(0.062)	(0.046)	
1992* Father's Ed: Missing, Unknown		0.081	0.081 *	0.003	
		(0.050)	(0.049)	(0.036)	
1992* Suburban			-0.088 **	-0.025	
			(0.038)	(0.028)	
1992* Rural			-0.078 *	-0.047	
			(0.040)	(0.030)	
1992* New England			-0.005	0.018	
			(0.075)	(0.056)	
1992* Middle Atlantic			0.009	-0.074 *	
			(0.060)	(0.044)	
1992* South Atlantic			0.222 **	0.116 **	
			(0.056)	(0.041)	
1992* East South Central			0.014	-0.008	
			(0.069)	(0.051)	
1992* West South Central			0.272 **	0.120 **	
			(0.061)	(0.045)	
1992* East North Central			0.090 *	0.028	
			(0.054)	(0.040)	
1992* West North Central			0.141 **	0.073	
			(0.065)	(0.048)	
1992* Mountain			0.139 *	0.003	
			(0.070)	(0.052)	
1992* Days In School Year			-0.003	-0.004	
			(0.004)	(0.003)	
1992* Days In School Year Missing			-0.627	-0.654	
			(0.803)	(0.589)	
1992* % Teachers with Master's Degree			0.116	-0.010	
			(0.075)	(0.055)	
1992* % Teachers with Master's Missing			0.085	-0.035	
			(0.072)	(0.053)	
1992* % Disadvantaged Students			0.026	0.027	
			(0.082)	(0.060)	
1992* % Disadvantaged Students Missing			0.083	0.032	
			(0.054)	(0.040)	
1992* Student Teacher Ratio			-0.004	-0.001	
			(0.003)	(0.002)	
1992* Student Teacher Ratio Missing			-0.170 **	-0.021	
			(0.079)	(0.058)	
Intercept	~0	4.939 **	4.180 **	1.579 **	-0.768 **
	(0.011)	(0.253)	(0.398)	(0.294)	(0.017)
R-Squared	0.116	0.288	0.302	0.625	0.591
Number of Observations	17,750	17,750	17,750	17,750	17,750

Note: ** indicates significance at the 5 percent level. * indicates significance at the 10 percent level.

7. References

Cook, Michael D. and William Evans (2000). "Families or Schools? Explaining the Convergence in White and Black Academic Performance," *Journal of Labor Economics*, 18(4):729-54.

Ferguson, Ronald F. (1998). "Can Schools Narrow the Black-White Test Score Gap?" in Christopher Jencks and Meredith Phillips, eds., *The Black-White Test Score Gap*, Washington, D.C.: Brookings Institution Press.

Green, Patricia, Bernard L. Dugoni, and Steven J. Ingels (1995). *Trends Among High School Seniors, 1972-1992*. NCES Statistical Analysis Report 95-380, Washington, D.C.: National Center for Education Statistics.

Grissmer, David, Ann Flanagan, and Stephanie Williamson (1998). "Why Did the Black-White Test Score Gap Narrow in the 1970s and 1980s?" in Christopher Jencks and Meredith Phillips, eds., *The Black-White Test Score Gap*, Washington, D.C.: Brookings Institution Press.

Grissmer, David W., Sheila Nataraj Kirby, Mark Berends, and Stephanie Williamson (1994). *Student Achievement and the Changing American Family*, Santa Monica, California: RAND.

Grogger, Jeff and Eric Eide (1995). "Changes in College Skills and the Rise in the College Wage Premium," *Journal of Human Resources*, 30(2):280-310.

Hanushek, Eric A. (1996). "School Resources and Student Performance," in Gary Burtless, ed., *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success,* Washington, D.C.: Brookings Institution Press.

Hedges, Larry V. and Amy Nowell (1998). "Black-White Test Score Convergence Since 1965," in Christopher Jencks and Meredith Phillips, eds., *The Black-White Test Score Gap*, Washington, D.C.: Brookings Institution Press.

Ingels, Steven, Kathryn L. Dowd, John R. Taylor, Virginia H. Bartot, Martin R. Frankel, and Paul A. Pulliam (1995). *National Education Longitudinal Study of 1988 Second Follow-up: Transcript Component Data File User's Manual*, NCES 95-06, Washington, D.C.: National Center for Education Statistics.

Jencks, Christopher and Meredith Phillips (1998). "The Black-White Test Score Gap: An Introduction" in Christopher Jencks and Meredith Phillips, eds., *The Black-White Test Score Gap*, Washington, D.C.: Brookings Institution Press.

Lee, Valerie E., David T. Burkam, Todd Chow-Hoy, Becky A. Smerdon, and Douglas Geverdt (1998). *High School Curriculum Structure: Effects on Coursetaking* *and Achievement in Mathematics for High School Graduates. An Examination of Data from the National Education Longitudinal Study of 1998,* NCES Working Paper 98-09, Washington, D.C.: National Center for Education Statistics.

Murnane, Richard J., John B. Willett, and Frank Levy (1995). "The Growing Importance of Cognitive Skills in Wage Determination," *Review of Economics and Statistics*, 77(2):251-266.

The National Commission on Excellence in Education (1983). *A Nation at Risk*, Washington, D.C.

Quin, Peggy , Steven J. Ingels, Kathryn Dowd, James Stripe, John D. Baldridge, Virginia H. Bartot, and Martin R. Frankel (1994). *NELS:88: Second Follow-Up: Dropout Component Data File User's Manual*, NCES 94-375, Washington, D.C.: National Center for Education Statistics.

Rasinski, Kenneth A., Steven J. Ingels, Donald A. Rock, and Judith M. Pollack (1993). *America's High School Sophomores: A Ten Year Comparison,* NCES 93-087, Washington, D.C.: National Center for Education Statistics.

Rock, Donald A. and Judith M. Pollack (1995). *Psychometric Report for the NELS:88 Base Year Through Second Follow-Up*, NCES Statistical Analysis Report 95-382, Washington, D.C.: National Center for Education Statistics.

Rose, Heather and Julian R. Betts (2004). "The Effect of High School Courses on Earnings," *Review of Economics and Statistics*, 86(2).