The Black-White Test Score Gap Through Third Grade

Roland G. Fryer, Jr., *Harvard University Society of Fellows and NBER,* and Steven D. Levitt, *University of Chicago and ABF*

This article describes basic facts regarding the Black-White test score gap over the first four years of school. Black children enter school substantially behind their White counterparts in reading and math, but including a small number of covariates erases the gap. Over the first four years of school, however, Blacks lose substantial ground relative to other races; averaging 0.10 standard deviations per school year. By the end of third grade, there is a large Black-White test score gap that cannot be explained by observable characteristics. Blacks are falling behind in virtually all categories of skills tested, except the most basic. None of the explanations we examine, including systematic differences in school quality across races, convincingly explain the divergent academic trajectory of Black students.

Decades after the landmark Supreme Court decision in *Brown v. Board of Education*, racial gaps in educational achievement remain substantial. Prior research shows Black children enter kindergarten lagging their White counterparts, and these differences grow throughout the school years (Campbell, Hombo, and Mazzeo, 2000; Caneiro and Heckman, 2003, Coleman, Campbell, and Hobson, 1966; Neal, 2005; Phillips, Crouse, and Ralph, 1998). On every subject at each grade level, there are

We are grateful to Cecelia Rouse and the participants of the Brown V. Board of Education Conference at Princeton University for helpful comments and suggestions.

Send correspondence to: Roland G. Fryer, Department of Economics, Harvard University, 1875 Cambridge Street, Cambridge MA, 02138; E-mail: rfryer@fas.harvard.edu

American Law and Economics Review
doi:10.1093/aler/ahl003
Advance Access publication July 25, 2006

substantial differences between Blacks and Whites (Campbell, Hombo, and Mazzeo, 2000; Neal, 2005). The typical Black 17-year old reads at the proficiency level of the typical White 13-year old (Campbell, Hombo, and Mazzeo, 2000). Even in affluent neighborhoods, achievement gaps are large (Ferguson, 2001, 2002; Ogbu and Davis, 2003). Including a myriad of controls, the test score gap remains essentially unchanged (Jencks and Phillips, 1998). Although the Brown decision provided unprecedented hope for a future of educational equality, that hope has yet to be realized.

Despite these disturbing differences, a recent analysis of a newly available data set, the Early Childhood Longitudinal Study (ECLS), provides two reasons for optimism (Fryer and Levitt, 2004). First, the raw test score differences for the recent cohort covered by ECLS are substantially below those found in earlier studies, suggesting the possibility of real gains by Blacks in recent cohorts. Second, in stark contrast to previous studies, Fryer and Levitt (2004) are able to eliminate the Black-White test score gap for incoming kindergarteners with the inclusion of just a parsimonious set of controls. Any optimism, however, is tempered by the fact that, by the end of first grade (the last data used in Fryer and Levitt, 2004), Black students have already lost substantial ground (the equivalent of almost three months of schooling) relative to Whites. If this trend were to continue, by the tenth grade, Blacks would be one standard deviation behind Whites—a number consistent with prior research (Jones, Burton, and Davenport, 1982, Phillips, Crouse, and Ralph, 1998; Phillips, 2000).

Fryer and Levitt (2004) were largely unsuccessful in pinpointing the mechanisms driving the divergent trajectories of Blacks and Whites. A number of leading hypotheses (the importance of parental and environmental contributions grow over time, Black students suffer worse summer setbacks, standardized tests are poor measures, interactions between Black students and schools interfere with learning) fail to explain why Blacks lost ground. The only hypothesis that received any empirical support was systematically lower quality schools for Blacks relative to Whites. The primary evidence in favor of this hypothesis emerged from comparisons of test score trajectories within versus across schools. Including school-fixed effects eliminates two-thirds of the difference in the learning trajectory of Blacks and Whites over the first two years of school. In other words, a White student attending the same school as a Black student loses two-thirds as much ground against the typical White student as does the Black

student. Nonetheless, the evidence on school quality as the driving force in the racial gaps in Fryer and Levitt (2004) was largely circumstantial and subject to numerous important caveats.[1]

In this article, we extend the analysis offered in Fryer and Levitt (2004) in three directions. First, data from ECLS through the third grade have recently become available, allowing us to extend the analysis from first grade to third grade. Second, we have obtained the restricted use version of the data that contain detailed information on additional geographic indicators down to the zip code level. Third, we investigate an additional explanation for the emerging Black-White test score gap, namely, that the set of skills tested in the third grade systematically differ relative to those in kindergarten and that Blacks perform worse on the skills emphasized in the later years.

A number of stylized facts emerge in this article. We find that Blacks continue to lose ground relative to Whites in second and third grade at a pace consistent with the losses observed between kindergarten and first grade. On average, Blacks are losing 10 standard deviations per year relative to Whites in the first four years of school. In contrast to Fryer and Levitt (2004), however, systematic differences in school quality appear much less important in explaining the differences in test-score trajectories by race, once the data are extended through third grade; Blacks lose substantial ground relative to Whites within the same school and even in the same classrooms. That is, including school- or teacher-fixed effects do little to explain the divergent trajectories of Black and White students between kindergarten and third grade. Hispanics continue to make up their inferior initial conditions relative to Whites, whereas Asians continue to make gains. Explanations other than school quality that we have explored also fail to convincingly account for the growing gap between Blacks and students of other races.

1. There are at least three limitations to the argument that school quality is the mechanism behind Black underachievement in Fryer and Levitt (2004). First, Hispanics also attend worse schools than Whites, yet their test scores converge to those of Whites. Second, because the assignment of children to schools depends in large part on residential location, school-fixed effects is in many ways equivalent to neighborhood-fixed effects. Third, controlling for a wide range of school inputs (which should capture important aspects of school quality) does little to lessen the gap.

By the end of third grade, even after controlling for observables, the Black-White test score gap is evident in every skill tested in reading and math except for the most basic tasks such as counting and letter recognition which virtually all students have mastered. The largest racial gaps in third grade are in the skills most crucial to future academic and labor market success: multiplication and division in math and inference, extrapolation, and evaluation in reading. Any initial optimism is drowned out by the growing gap.

The remainder of the article is structured as follows. Section II describes the data used in the analysis. Section III presents the basic facts and patterns in test scores in the first four years of school using these data. Section IV investigates the extent to which alternative hypotheses can account for the fact that Blacks are steadily losing ground. Section V concludes.

1. The Data

The Early Childhood Longitudinal Study Kindergarten Cohort (ECLS-K) is a nationally representative sample of over 20,000 children entering kindergarten in 1998. Thus far, information on these children has been gathered at five separate points in time. The full sample was interviewed in the fall and spring of kindergarten, spring of first grade, and spring of third grade. The sample will ultimately be followed through fifth grade.[2] Roughly 1,000 schools are included in the sample, with an average of more than 20 children per school in the study. As a consequence, it is possible to conduct within-school analyses.

A wide range of data is gathered on the children in the study, which is described in detail at the ECLS web site http://nces.ed.gov/ecls. We utilize just a small subset of the available information in our baseline specifications (although Fryer and Levitt [2004] show that similar results are obtained in a much more fully specified model). Students who are missing data on test-scores, race, or age are dropped from our sample.

Summary statistics for the variables we use in our core specifications are summarized by race in Table 1, with White referring solely to non-Hispanic Whites.[3] Our primary outcome variables are math and

2. In addition, there is an ECLS birth cohort that tracks a nationally representative sample of over 15,000 children born in 2001 through the first grade.

3. There are also a small number of children in the data whose racial status is classified as "other." These include Hawaiian, mixed race, and Native American students. Such students are included in our regressions but not shown in the summary statistics table.

Table 1. Summary Statistics by Race: Student Characteristics

Variable	Full Sample	White	Black	Hispanic	Asian
Test scores					
Fall kindergarten math	0 (1)	.307 (1.01)	−.356 (.759)	−.431 (.822)	.417 (1.13)
Spring kindergarten math	0 (1)	.304 (.975)	−.421 (.854)	−.377 (.914)	.392 (1.13)
Spring first grade math	0 (1)	.286 (.933)	−.472 (.974)	−.283 (.976)	.261 (1.03)
Spring third grade math	0 (1)	.275 (.908)	−.607 (.958)	−.264 (.885)	.340 (.956)
Fall kindergarten reading	0 (1)	.176 1	−.224 (.809)	−.273 (.928)	.487 (1.01)
Spring kindergarten reading	0 (1)	.182 (.982)	−.270 (.919)	−.162 (.965)	.537 (1.17)
Spring first grade reading	0 (1)	.216 (.952)	−.301 (.999)	−.103 (1.02)	.478 (1.16)
Spring third grade reading	0 (1)	.279 (.905)	−.491 (.968)	−.110 (.937)	.286 (.914)
Race					
White	.554 (.497)	1	0	0	0
Black	.152 (.359)	0	1	0	0
Hispanic	.178 (.382)	0	0	1	0
Asian	.065 (.246)	0	0	0	1
Other controls					
Female	.489 (.50)	.484 (.5)	.497 (.5)	.494 (.5)	.498 (.5)
Age (in months), fall kindergarten	67.013 (4.480)	67.45 (4.46)	66.73 (4.47)	66.38 (4.38)	65.98 (4.20)
SES composite measure, kindergarten	.005 (.782)	.212 (.731)	−.333 (.745)	−.381 (.694)	.183 (.811)
SES composite measure, first grade	−.002 (.700)	.175 (.682)	−.275 (.607)	−.309 (.630)	.071 (.728)
SES composite measure, third grade	−.001 (.642)	.149 (.627)	−.214 (.550)	−.277 (.599)	.055 (.668)
Number of children's books in the home, kindergarten	61.432 (60.706)	81.44 (63.9)	32.32 (39.0)	35.77 (45.1)	33.77 (47.61)

(Continued)

Table 1. Continued

Variable	Full Sample	White	Black	Hispanic	Asian
Number of children's books in the home, first grade	74.60 (133.6)	102.49 (156.8)	31.62 (50.1)	38.75 (68.37)	40.08 (70.61)
Number of children's books in the home, third grade	76.82 (156.9)	103.79 (186.9)	33.29 (97.86)	44.89 (92.64)	43.35 (84.17)
Mother's age at time of first birth	23.609 (5.472)	24.79 (5.35)	20.63 (4.77)	21.95 (4.97)	25.56 (5.49)
Child's birth weight (in ounces)	87.463 (54.852)	94.76 (52.61)	77.15 (54.45)	83.00 (55.99	61.06 (58.25)
WIC participant	.378 (.485)	.270 (.444)	.608 (.488)	.529 (.499)	.236 (.425)

Notes: The entries are means and standard deviations of student-level data for those students who do not have missing values for test scores, race, or age. Test scores are IRT scores, normalized to have a mean of 0 and a standard deviation of 1 in the full, unweighted sample. The category white includes only non-Hispanic Whites. Precise definitions of the variables are provided in the data appendix. The SES composite measure incorporates information on parental education, occupational status, and family income. The total number of students included in the sample is 11,201.

reading standardized test scores.[4] Standardized tests were administered to the full sample in the fall and spring of kindergarten and first grade and the spring of third grade.[5] The reading test includes questions designed to measure basic skills (print familiarity, letter recognition, beginning and ending sounds, rhyming sounds, and word recognition), vocabulary and comprehension, listening and reading comprehension, knowledge of the alphabet, phonetics, and so on. The math test evaluates number recognition, counting, comparing and ordering numbers, solving word problems, interpreting picture graphs, addition and subtraction, multiplying and dividing, place value and rate, and measurement. The values reported in the table are item response theory (IRT) scores provided in ECLS-K, which we have transformed to have mean 0 and a standard deviation of 1 for the overall sample on each of the tests and time periods.[6] In all instances, sample weights provided in ECLS-K are used.[7]

4. These tests were developed especially for the ECLS but are based on existing instruments including Children's Cognitive Battery (CCB); Peabody Individual Achievement Test-Revised (PIAT-R); Peabody Picture Vocabulary Test-3 (PPVT-3); Primary Test of Cognitive Skills (PTCS); and Woodcock-Johnson Psycho-Educational Battery-Revised (WJ-R). Students are administered the test questions orally, as it is not assumed that they know how to read. A "general knowledge" exam was also administered. The general knowledge test is designed to capture "children's knowledge and understanding of the social, physical, and natural world and their ability to draw inferences and comprehend implications." We limit the analysis to math and reading scores, primarily because of the comparability of these test scores to past research in the area. In addition, there appear to be some peculiarities in the results of the general knowledge exam. See Rock and Stenner (2005) for a more detailed comparison of ECLS to previous testing instruments.

5. The tests were also given in the spring of kindergarten, but we limit our focus to the endpoints of the available data. The kindergarten spring test results are in all cases consistent with the results presented in the article.

6. Because children were asked different questions depending on the answers they provided to the initial questions on the test, IRT-adjusted scores are preferable to simple test-score measures reflecting the number of correct answers a child provided. For more detail on the process used to generate the IRT scores, see chapter 3 of the ECLS-K User's Guide. Our results are not sensitive to normalizing the IRT scores to have a 0 mean and standard deviation equal to 1.

7. Because of the complex manner in which the ECLS-K sample is drawn, different weights are suggested by the providers of the data depending upon the set of variables used (BYPW0). We utilize the weights recommended for making longitudinal comparisons. None of our findings are sensitive to other choices of weights or not weighting at all.

White students on average score .307 standard deviations above the mean on the math exam in the fall of kindergarten, whereas Black students perform .356 standard deviations below the mean on that test, yielding a Black-White gap of .663 standard deviations. By the spring of third grade, that gap has increased to .882 standard deviations. The initial Black-White gap on reading is smaller (.400 standard deviations). Like math, however, the reading gap widens substantially to .771 standard deviations by the end of third grade.

The remainder of Table 1 summarizes summary statistics for the other variables used in the analysis. In contrast to the test score variables, for which we have observations at multiple points in time, many of the control variables are not time varying (e.g., birth weight), collected only once, or exhibit little variation over time for individual students. The most important of these covariates is a composite measure of socioeconomic status constructed by the researchers conducting the ECLS survey. The components used in the socioeconomic scale (SES) measure are parental education, parental occupational status, and household income. Other variables included as controls are gender, child's age at the time of enrollment in kindergarten, WIC participation (a nutrition program aimed at relatively low-income mothers and children), mother's age at first birth, birth weight, and the number of children's books in the home.[8] There are substantial differences across races on many of these variables. Black children in the sample are growing up under circumstances likely to be less conducive to academic achievement than White children: lower socioeconomic status, fewer children's books in the home, and so on. Hispanics are also worse off than Whites on average. For Asians, the patterns are more mixed. The set of covariates we include matches that used in Fryer and Levitt (2004). While this particular set of covariates might seem idiosyncratic, the results we obtain with this small set of variables mirrors the findings when we include an exhaustive set of over 100 controls. In light of past research that has had great difficulty making the Black-White test score gap disappear, we focus on the results from these very parsimonious regressions to highlight the fact that the sharp differences between our results and earlier studies are not primarily a consequence of the availability of different

8. A more detailed description of each of the variables used is provided in the appendix.

covariates in the ECLS. It is important to stress that a causal interpretation of the coefficients on the covariates is likely to be inappropriate; we view these particular variables as proxies for a broader set of environmental and behavioral factors.

2. Basic Facts about Racial Differences in Early Achievement

Table 2 summarizes a series of estimates of the racial test score gap in math for the tests taken over the first four years of school. The specifications estimated are of the form:

$$y_{it} = \rho_i \gamma + x_{it}\beta + \varepsilon_{it} \quad (1)$$

where y_{it} denotes an individual I's test score in grade t and x_{it} represents an array of student level social and economic variables describing each student's environment. The variable ρ_i is a full set of race dummies included in the regression, with White as the omitted category. Consequently, the coefficients on race capture the gap between the named racial category and the Whites. Our primary emphasis is on the Black-White test score gap. In all instances, the estimation is done using weighted least squares, with weights corresponding to the sampling weights provided in the data set. When there are multiple observations of social and economic variables (SES, number of books in the home, and so on), for all specifications, we only include the value recorded in the fall kindergarten survey.[9] Our analysis consists of a series of cross-sectional regressions; we do not use the panel structure of the ECLS for any of our analysis.

The first four columns of Table 2 summarize the differences in means, not including any covariates. These results simply reflect the raw test score gaps summarized in Table 1. Columns 5–8 mirror the main specification in Fryer and Levitt (2004). Controls include the composite indicator of socioeconomic status constructed by the ECLS survey administrators, number of children's books in the home, and that variable squared, gender, age, birth weight, indicator variables for having a mother whose first birth came when she was a teenager or over 30 (the omitted category is having a first birth in one's 20s), and WIC participation. These covariates

9. Including all the values of these variables from each survey or only those in the relevant years does not alter the results.

Table 2. Estimated Racial Achievement Gap over the First Four Years of School, Math

	(1) Fall-K	(2) Spring-K	(3) Spring-1st	(4) Spring-3rd	(5) Fall-K	(6) Spring-K	(7) Spring-1st	(8) Spring-3rd
Black	−0.663 (0.025)	−0.724 (0.027)	−0.758 (0.029)	−0.882 (0.031)	−0.099 (0.026)	−0.209 (0.028)	−0.279 (0.031)	−0.382 (0.033)
Hispanic	−0.738 (0.024)	−0.681 (0.025)	−0.568 (0.026)	−0.539 (0.026)	−0.197 (0.024)	−0.189 (0.026)	−0.122 (0.027)	−0.078 (0.028)
Asian	0.11 (0.058)	0.088 (0.056)	−0.025 (0.052)	0.066 (0.054)	0.258 (0.050)	0.226 (0.050)	0.092 (0.047)	0.163 (0.049)
Other race	−0.495 (0.047)	−0.481 (0.048)	−0.497 (0.050)	−0.541 (0.050)	−0.158 (0.040)	−0.175 (0.043)	−0.21 (0.046)	−0.244 (0.046)
Age (in months)					0.058 (0.002)	0.053 (0.002)	0.037 (0.002)	0.019 (0.002)
Birth weight (in ounces)					0.003 (0.000)	0.003 (0.000)	0.003 (0.000)	0.003 (0.000)
Female					0.005 (0.017)	−0.005 (0.017)	−0.044 (0.018)	−0.175 (0.018)
Number of children's books					0.006 (0.001)	0.006 (0.001)	0.005 (0.001)	0.006 (0.001)
Number of children's books (squared) (∗1000)					−0.021 (0.002)	−0.02 (0.003)	−0.019 (0.003)	−0.020 (0.003)
Mother over 30 at first birth					0.165 (0.026)	0.107 (0.025)	0.086 (0.022)	0.083 (0.024)
Socioeconomic status measure					0.306 (0.016)	0.282 (0.015)	0.256 (0.015)	0.288 (0.015)
Mother receives Wic benefits					−0.212 (0.021)	−0.191 (0.022)	−0.19 (0.023)	−0.208 (0.024)
Mother a teenager at first birth					−0.114 (0.021)	−0.118 (0.022)	−0.131 (0.025)	−0.132 (0.025)
Constant	0.307 (0.013)	0.304 (0.013)	0.286 (0.012)	0.275 (0.012)	−4.357 (0.154)	−3.952 (0.160)	−2.795 (0.168)	−1.576 (0.168)
Observations	11201	11201	11201	11201	11201	11201	11201	11201
R-squared	0.11	0.11	0.1	0.12	0.32	0.29	0.24	0.26

Notes: The dependent variable is the math test score at various points in a student's career. Test scores are IRT scores, normalized to have a mean of 0 and a standard deviation of 1 in the full, unweighted sample. Non-Hispanic Whites are the omitted race category, so all of the race coefficients are relative to that group. The unit of observation is a student. Standard errors in parentheses. Estimation is done using weighted least squares, using sample weights provided in the data set.

generally enter with the expected sign. Older children, those with higher birth weights, those with older mothers at the time of first birth all score better, although the benefit of entering school at a later age decreases steadily over time. Children on WIC do worse on the tests, suggesting that this variable is not capturing any real benefits the program might provide, but rather, the fact that eligibility for WIC is a proxy for growing up poor that the SES variable is not adequately capturing. Socioeconomic status and the number of children's books in the home are important predictors of test scores at each grade level. An one-standard deviation increase in the SES variable is associated with a .306 increase in fall kindergarten math scores and a .288 increase in spring first grade math scores. The number of books is also strongly positively associated with high kindergarten test scores in math.[10] Evaluated at the mean, a one-standard deviation increase in the number of books (from 72 to 137) is associated with an increase of .143 standard deviations in math and .115 standard deviations in reading. This variable seems to serve as a useful proxy for capturing the conduciveness of the home environment to academic success. The other variables tend to enter with the expected sign and have magnitudes that are similar to those reported in Fryer and Levitt (2004).[11]

The estimates in Table 2 suggest that, controlling for other factors, Black students score only slightly worse in math than Whites upon kindergarten entry, but their trajectories after entry into school are very different.[12] After controlling for our parsimonious specification, Blacks score .099 standard deviations below Whites in the fall of kindergarten. This deficit increases to .279 standard deviations by the spring of first grade and .382 by the spring of third grade. Thus, the Black-White test

10. The marginal benefit associated with one additional book decreases as more books are added. Beyond roughly 150 books, the marginal impact turns negative. Only 16% of the sample lies above this cutoff point.

11. As an additional robustness check, we have also entered the components of the SES measure, parental education, parental occupation, and income, individually. In all cases, the components enter with the expected sign. Income is the biggest contributor to test scores. The Black coefficient is unchanged.

12. The specifications in Table 2 are restricted to observations with valid test scores on all four tests. Because of this, our sample size is 2,089 observations smaller than Fryer and Levitt (2004). There is little change in the results when we restrict samples for a given test to students with valid scores on that test.

score gap grows by almost .30 percentiles between the fall of kindergarten and spring of third grade. The table also illustrates that the control variables included in the specification shrink the gap a roughly constant amount of approximately .50 standard deviations regardless of the year of testing. In other words, although Blacks systematically differ from Whites on these background characteristics, the impact of these variables on test scores is remarkably stable over time. Whatever factor is causing Blacks to lose ground is operating through a different channel.

In contrast to Blacks, Hispanics gain substantial ground relative to Whites, although they are plagued with many of the social problems that exist among Blacks—low socioeconomic status, inferior schools, and so on. One explanation for Hispanic convergence is increases in English proficiency, although we have little direct evidence on this question.[13] Calling into question that hypothesis is the fact, discussed below, that after controlling for other factors Hispanics do not test particularly poorly on reading, even upon school entry. Controlling for whether English is spoken in the home does little to affect the initial gap or the trajectory of Hispanics.[14] The large advantage enjoyed by Asians in the first two years of school is maintained. We also observe striking losses by girls relative to boys on math—over two-tenths of a standard deviation over the four year period. Although not the subject of this analysis, this is a finding deserving of further study.

Table 3 is identical to Table 2, but summarizes reading scores rather than math scores. Surprisingly, after adding our controls, Black children actually score slightly better than Whites in reading in the fall of kindergarten. Like math, however, Blacks lose substantial ground relative to other racial groups in the first four years of school. The coefficient on the indicator variable Black is .13 standard deviation above Whites in the fall of kindergarten and .282 standard deviations below Whites in the spring of third grade or a loss of over

13. Hispanics seem to increase their position relative to Whites in states where English proficiency is known to be a problem (Arizona, California, and Texas).

14. Hispanics are also less likely to participate in pre-school, which could explain their poor initial scores and positive trajectory. However, including controls for the type of program/care children have before entering kindergarten does nothing to explain why Hispanics gain ground.

Table 3. Estimated Racial Achievement Gap over the First Four Years of School, Reading

	(1) Fall-K	(2) Spring-K	(3) Spring-1st	(4) Spring-3rd	(5) Fall-K	(6) Spring-K	(7) Spring-1st	(8) Spring-3rd
Black	−0.4 (0.029)	−0.451 (0.029)	−0.517 (0.030)	−0.771 (0.032)	0.13 (0.030)	0.001 (0.03)	−0.078 (0.032)	−0.282 (0.034)
Hispanic	−0.45 (0.028)	−0.344 (0.029)	−0.319 (0.030)	−0.389 (0.030)	−0.071 (0.027)	−0.023 (0.029)	−0.014 (0.03)	−0.05 (0.03)
Asian	0.311 (0.069)	0.356 (0.062)	0.261 (0.052)	0.007 (0.046)	0.421 (0.063)	0.441 (0.058)	0.332 (0.048)	0.071 (0.042)
Other race	−0.37 (0.051)	−0.354 (0.048)	−0.405 (0.051)	−0.565 (0.055)	−0.06 (0.044)	−0.092 (0.044)	−0.153 (0.046)	−0.282 (0.049)
Age (in months)					0.044 (0.002)	0.036 (0.002)	0.023 (0.002)	0.013 (0.002)
Birth weight (in ounces)					0.002 (0.000)	0.002 (0.000)	0.002 (0.000)	0.001 (0.000)
Female					0.158 (0.018)	0.184 (0.019)	0.204 (0.019)	0.173 (0.018)
Number of children's books					0.007 (0.001)	0.006 (0.001)	0.006 (0.001)	0.006 (0.001)
Number of children's books (squared) (∗1000)					−0.024 (0.003)	−0.021 (0.003)	−0.022 (0.003)	−0.023 (0.003)
Mother over 30 at first birth					0.221 (0.029)	0.155 (0.027)	0.072 (0.025)	0.116 (0.023)
Socioeconomic status measure					0.3 (0.017)	0.275 (0.016)	0.277 (0.015)	0.294 (0.015)
Mother receives WIC benefits					−0.176 (0.023)	−0.141 (0.023)	−0.163 (0.024)	−0.194 (0.024)
Mother a teenager at first birth					−0.144 (0.022)	−0.136 (0.024)	−0.14 (0.026)	−0.151 (0.027)
Constant	0.176 (0.013)	0.182 (0.012)	0.216 (0.012)	0.279 (0.012)	−3.433 (0.161)	−2.927 (0.165)	−1.97 (0.173)	−1.139 (0.174)
Observations	10540	10540	10540	10540	10540	10540	10540	10540
R-squared	0.04	0.04	0.05	0.09	0.23	0.19	0.19	0.25

Notes: The dependent variable is the reading test score at various points in a student's career. Test scores are IRT scores, normalized to have a mean of 0 and a standard deviation of 1 in the full, unweighted sample. Non-Hispanic Whites are the omitted race category; so, all of the race coefficients are relative to that group. The unit of observation is a student. Standard errors in parentheses. Estimation is done using weighted least squares, using sample weights provided in the data set.

.40 standard deviations for the typical Black child relative to the typical White.[15] The impact of covariates—explaining about one-half of a standard deviation gap between Blacks and Whites at all ages—is similar to that in the math regressions. Hispanics experience a much smaller gap relative to Whites, and it does not grow. The early edge enjoyed by Asians diminishes by third grade. In stark contrast to the results on math, girls are not losing ground relative to boys in reading.

In an effort to uncover the factors that are associated with the divergent trajectory of Blacks, Table 4 explores the sensitivity of these losing ground estimates across a wide variety of sub-samples of the data. We report only the race coefficients and associated standard errors in the table. The top row of the table presents the baseline results using a full sample and our parsimonious set of controls (corresponding to Tables 2 and 3). In that specification, Blacks lose an average of .283 standard deviations in math and .41 in reading relative to Whites over the first four years of school. Blacks lose similar amounts of ground across most subsets of the data. In part because of imprecise estimates, only in a few cases can we reject the null hypothesis of no differences in the amount of ground lost by Blacks across sub-groups. Black females fare somewhat better relative to White females than Black males do relative to White males, but it is worth bearing in mind that White females do quite poorly relative to White males. The results appear to be similar across quintiles of the socioeconomic status distribution and by family structure. Blacks in schools with less than 50% Blacks lose somewhat less ground to the Whites in their schools than do Blacks in mostly Black schools. Blacks in private schools do not appear to do especially well or poorly.[16] The single greatest outlier we observe is among Blacks in the western region, who start school doing well but fall far behind. Hispanics in the West do not exhibit this same

15. The numbers here for third graders are similar in magnitude to those reported in NAEP scores (see Campbell, Hombo, and Mazzeo 2000).

16. We have also experimented with limiting the sample to the set of children for whom there is substantial overlap across races in background characteristics. More specifically, we ran probits with an indicator variable for Black as the dependent variable and the full set of covariates as predictors. When we drop from the sample the roughly 30% of students whose predicted probability of being Black is <10% or >90%, the Black-White gap on math rises slightly and the reading gap becomes closer to 0.

Table 4. Sensitivity Analysis for Losing Ground

Specification	Kindergarten Math	Third Grade Math	Lost Ground	Kindergarten Reading	Third Grade Reading	Lost Ground
Baseline	−0.099 (0.03)	−0.382 (0.03)	**0.28** (0.04)	0.13 (0.03)	−0.282 (0.03)	**0.412** (0.05)
By Gender						
Males	−0.123 (0.04)	−0.443 (0.05)	**0.32** (0.06)	0.11 (0.04)	−0.304 (0.05)	**0.414** (0.07)
Females	−0.07 (0.03)	−0.314 (0.05)	**0.244** (0.06)	0.154 (0.04)	−0.258 (0.05)	**0.412** (0.06)
By SES quintile						
Bottom	−0.099 (0.05)	−0.327 (0.07)	**0.228** (0.09)	−0.057 (0.05)	−0.259 (0.08)	**0.202** (0.09)
Second	−0.096 (0.05)	−0.376 (0.07)	**0.28** (0.09)	0.099 (0.06)	−0.286 (0.07)	**0.385** (0.09)
Third	−0.105 (0.05)	−0.382 (0.07)	**0.277** (0.09)	0.107 (0.05)	−0.241 (0.07)	**0.348** (0.08)
Fourth	−0.124 (0.07)	−0.447 (0.08)	**0.323** (0.10)	0.329 (0.10)	−0.31 (0.08)	**0.639** (0.12)
Top	−0.124 (0.10)	−0.318 (0.10)	**0.194** (0.14)	0.085 (0.10)	−0.339 (0.08)	**0.424** (0.13)
By family structure						
Single mother	−0.114 (0.05)	−0.408 (0.06)	**0.294** (0.08)	0.08 (0.05)	−0.316 (0.06)	**0.396** (0.08)
Two biological parents	−0.126 (0.04)	−0.363 (0.05)	**0.237** (0.06)	0.148 (0.05)	−0.248 (0.05)	**0.396** (0.07)
Teen mother at child's birth	−0.098 (0.04)	−0.375 (0.05)	**0.277** (0.07)	0.012 (0.04)	−0.28 (0.06)	**0.292** (0.07)
By Region						
Northeast	−0.064 (0.07)	−0.425 (0.09)	**0.361** (0.11)	0.194 (0.09)	−0.236 (0.08)	**0.43** (0.12)
Midwest	−0.067 (0.06)	−0.257 (0.08)	**0.19** (0.10)	0.076 (0.07)	−0.252 (0.07)	**0.328** (0.10)
South	−0.149 (0.04)	−0.4 (0.05)	**0.251** (0.06)	0.043 (0.04)	−0.292 (0.05)	**0.335** (0.06)
West	0.144 (0.09)	−0.458 (0.10)	**0.602** (0.14)	0.47 (0.12)	−0.319 (0.10)	**0.789** (0.16)

(Continued)

Table 4. Continued

Specification	Kindergarten Math	Third Grade Math	Lost Ground	Kindergarten Reading	Third Grade Reading	Lost Ground
By Location type						
Central city	−0.116 (0.04)	−0.429 (0.05)	**0.313** (0.06)	0.17 (0.05)	−0.287 (0.05)	**0.457** (0.07)
Suburban	−0.15 (0.04)	−0.359 (0.06)	**0.209** (0.07)	0.032 (0.04)	−0.324 (0.06)	**0.356** (0.07)
Rural	−0.178 (0.05)	−0.53 (0.08)	**0.352** (0.09)	−0.04 (0.06)	−0.392 (0.09)	**0.352** (0.11)
By school type						
Public	−0.117 (0.03)	−0.389 (0.04)	**0.272** (0.04)	0.102 (0.03)	−0.287 (0.04)	**0.389** (0.05)
Private	0.057 (0.08)	−0.29 (0.09)	**0.347** (0.12)	0.331 (0.09)	−0.262 (0.07)	**0.593** (0.11)
School > 50% Black	−0.144 (0.10)	−0.473 (0.11)	**0.329** (0.15)	0.09 (0.09)	−0.352 (0.11)	**0.442** (0.14)
School < 50% Black	−0.13 (0.04)	−0.362 (0.05)	**0.232** (0.06)	0.09 (0.05)	−0.223 (0.05)	**0.313** (0.06)

Notes: The table summarizes estimates on the gap between Black and White test scores, controlling for other factors, upon entry to kindergarten and at the end of third grade. The column labeled "Lost Ground" is the change in that gap over the first four years of school. The first three columns correspond to math, and the last three columns are for reading. The top row provides baseline estimates for the entire sample and the remaining rows focus on subsets of the overal sample. The estimates are from regressions using our parsimonious set of controls. Standard errors are in parentheses.

pattern; they remain steady versus Whites. We have not found a compelling explanation for the poor performance of Blacks in the West. One point worth noting is that Blacks in the West are losing the great majority of this ground relative to Whites in the same classrooms; so, differential school quality across races does not appear to be the answer.

The results summarized in Tables 2–4 maintain the assumption that children of different races are equally responsive to changes in covariates. Cross-race differences in coefficients are potentially important because they affect the interpretation of the racial test score gap estimates in the preceding tables. Black children experience worse environments on average. If Black children do not derive as much benefit from improvements in socioeconomic status, number of children's books, higher birth weight, and so on, then our earlier results suggesting that including covariates lessens the racial gaps may be exaggerated. Furthermore, within-race analysis allows one to see how the relationship between particular covariates (e.g., number of children's books) and achievement varies over time within race.

Tables 5 and 6 summarize within-race estimates of our basic specifications in math and reading, respectively. Columns 1 and 6 replicate the coefficient estimates from the full sample. The remaining columns present results within a specific race category. For the most part, responsiveness to covariates appears similar across races. One difference is that the Black children in our sample may be somewhat less responsive to changes in socioeconomic status than the Whites: a one-standard deviation improvement in socioeconomic status for a Black child is associated with a .192 standard deviation increase in math scores compared to .343 for a White child, but the results are more similar on reading scores.

3. Why are Black students losing ground?

Understanding why Black students fare worse in the first four years of school is a question of paramount importance for two reasons. First, knowing the source of the divergence may aid in developing public policies to alleviate the problem. Second, determining the explanation for the widening gap will help to determine whether the simple linear extrapolation over the academic career is a plausible conjecture.

There are a number of plausible explanations as to why the racial gap in test scores grows as children age: (1) Black children attend lower quality

Table 5. Estimates of the Responsiveness of Math Scores to Covariates by Race

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Fall Kindergarten					Spring Third Grade				
	Full Sample	**Whites**	**Blacks**	**Hispanics**	**Asians**	**Full Sample**	**Whites**	**Blacks**	**Hispanics**	**Asians**
Black	−0.099	–	–	–	–	−0.382	–	–	–	–
	(0.026)					(0.033)				
Hispanic	−0.197	–	–	–	–	−0.078	–	–	–	–
	(0.024)					(0.028)				
Asian	0.258	–	–	–	–	0.163	–	–	–	–
	(0.050)					(0.049)				
Other race	−0.158	–	–	–	–	−0.244	–	–	–	–
	(0.040)					(0.046)				
female	0.005	−0.001	0.055	0.03	−0.135	−0.175	−0.197	−0.068	−0.151	−0.233
	(0.017)	(0.023)	(0.039)	(0.034)	(0.099)	(0.018)	(0.022)	(0.054)	(0.044)	(0.094)
Age at K–Fall (in months)	0.058	0.064	0.046	0.052	0.076	0.019	0.017	0.027	0.019	0.017
	(0.002)	(0.003)	(0.005)	(0.004)	(0.017)	(0.002)	(0.003)	(0.006)	(0.006)	(0.014)
Socioeconomic status	0.306	0.343	0.192	0.227	0.419	0.288	0.297	0.243	0.266	0.356
	(0.016)	(0.021)	(0.035)	(0.033)	(0.071)	(0.015)	(0.018)	(0.041)	(0.041)	(0.070)
Number of children's books	0.006	0.005	0.005	0.01	0.008	0.006	0.005	0.007	0.007	0.005
	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)
Number of children's books (squared)	−0.021	−0.017	−0.019	−0.035	−0.027	−0.020	−0.017	−0.025	−0.025	−0.018
	(0.002)	(0.004)	(0.010)	(0.006)	(0.015)	(0.003)	(0.004)	(0.010)	(0.008)	(0.014)
Birth weight in ounces	0.003	0.004	0.003	0.001	0.009	0.003	0.004	0.004	0.001	0.005
	(0.000)	(0.001)	(0.001)	(0.001)	(0.003)	(0.000)	(0.001)	(0.001)	(0.001)	(0.003)

(Continued)

Teenage mother at time of first birth	−0.114 (0.021)	−0.136 (0.034)	−0.125 (0.042)	−0.077 (0.035)	−0.054 (0.132)	−0.132 (0.025)	−0.155 (0.036)	−0.127 (0.058)	−0.106 (0.049)	−0.24 (0.222)
Mother at least 30 at age of first birth	0.165 (0.026)	0.157 (0.031)	0.098 (0.081)	0.174 (0.070)	0.243 (0.121)	0.083 (0.024)	0.073 (0.026)	0.06 (0.115)	0.137 (0.075)	0.018 (0.108)
Wic participant	−0.212 (0.021)	−0.212 (0.030)	−0.174 (0.056)	−0.188 (0.040)	−0.163 (0.113)	−0.208 (0.024)	−0.211 (0.031)	−0.168 (0.072)	−0.201 (0.052)	−0.021 (0.121)
Constant	−4.357 (0.154)	−4.758 (0.214)	−3.648 (0.325)	−3.988 (0.322)	−5.974 (1.081)	−1.576 (0.168)	−1.498 (0.209)	−2.678 (0.457)	−1.471 (0.418)	−1.497 (0.961)
Observations	11201	6808	1370	1945	478	11201	6808	1370	1945	478
R-squared	0.32	0.23	0.2	0.31	0.3	0.26	0.18	0.13	0.15	0.19

Notes: The table reports estimates from our parsimonious specification, done separately by race, for math test scores in the fall of kindergarten and spring of third grade. The first and sixth columns provide baseline estimates for the entire sample. Standard errors are in parentheses.

Table 6. Estimates of the Responsiveness of Reading Scores to Covariates by Race

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Fall Kindergarten					Spring Third Grade				
	Full Sample	Whites	Blacks	Hispanics	Asians	Full Sample	Whites	Blacks	Hispanics	Asians
Black	0.13	–	–	–	–	−0.282	–	–	–	–
	(0.030)					(0.034)				
Hispanic	−0.071	–	–	–	–	−0.05	–	–	–	–
	(0.027)					(0.030)				
Asian	0.421	–	–	–	–	0.071	–	–	–	–
	(0.063)					(0.042)				
Other race	−0.06	–	–	–	–	−0.282	–	–	–	–
	(0.044)					(0.049)				
Female	0.158	0.168	0.156	0.146	0.172	0.173	0.174	0.196	0.199	0.104
	(0.018)	(0.024)	(0.046)	(0.046)	(0.125)	(0.018)	(0.022)	(0.055)	(0.052)	(0.078)
Age at K-Fall (in months)	0.044	0.047	0.04	0.035	0.044	0.013	0.017	0.001	0.016	0.006
	(0.002)	(0.003)	(0.005)	(0.005)	(0.020)	(0.002)	(0.003)	(0.007)	(0.006)	(0.012)
Socioeconomic Status	0.3	0.315	0.228	0.222	0.582	0.294	0.303	0.255	0.268	0.348
	(0.017)	(0.022)	(0.038)	(0.043)	(0.084)	(0.015)	(0.018)	(0.041)	(0.047)	(0.056)
Number of Children's Books	0.007	0.006	0.005	0.008	0.01	0.006	0.006	0.007	0.005	0.003
	(0.001)	(0.001)	(0.002)	(0.002)	(0.004)	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)
Number of Children's Books (squared)	−0.024	−0.023	−0.013	−0.028	−0.039	−0.023	−0.021	−0.026	−0.018	−0.01
	(0.003)	(0.004)	(0.012)	(0.008)	(0.020)	(0.003)	(0.004)	(0.011)	(0.009)	(0.013)
Birth Weight in ounces	0.002	0.002	0.002	−0.001	0.009	0.001	0.001	0 (0.001)	0 (0.001)	0.005
	(0.000)	(0.001)	(0.001)	(0.001)	(0.005)	(0.000)	(0.001)			(0.002)

(Continued)

Teenage mother at time of first birth	−0.144	−0.114	−0.215	−0.141	0.023	−0.151	−0.155	−0.144	−0.129	−0.154
	(0.022)	(0.033)	(0.048)	(0.045)	(0.161)	(0.027)	(0.037)	(0.061)	(0.060)	(0.164)
Mother at least 30 at age of first birth	0.221	0.209	0.182	0.342	0.104	0.116	0.097	0.213	0.192	−0.025
	(0.029)	(0.033)	(0.108)	(0.089)	(0.158)	(0.023)	(0.026)	(0.101)	(0.076)	(0.080)
Wic Participant	−0.176	−0.171	−0.146	−0.172	−0.272	−0.194	−0.2	−0.08	−0.205	−0.212
	(0.023)	(0.030)	(0.070)	(0.050)	(0.134)	(0.024)	(0.031)	(0.070)	(0.058)	(0.110)
Constant	−3.433	−3.702	−3.071	−2.669	−3.93	−1.139	−1.382	−0.715	−1.201	−0.854
	(0.161)	(0.211)	(0.358)	(0.385)	(1.212)	(0.174)	(0.213)	(0.501)	(0.441)	(0.858)
Observations	10540	6788	1344	1342	477	10540	6788	1344	1342	477
R-squared	0.23	0.18	0.19	0.24	0.28	0.25	0.18	0.13	0.16	0.25

Notes: The table reports estimates from our parsimonious specification, done separately by race, for reading test scores in the fall of kindergarten and spring of third grade. The first and sxith columns provide baseline estimates for the entire sample. Standard errors are in parentheses.

schools on average, (2) the importance of parental/environmental contributions may grow over time. Because Black children are on average disadvantaged in this regard, they fall behind, (3) Black-White differences in earlier test scores were masked because of the type of material asked or the difficulties in testing skills among the youngest school-age children. Differences may only manifest themselves in more involved, higher order problems.[17] We address each of these hypotheses in turn.[18]

3.1. Are Black students losing ground because they attend worse schools?

Fryer and Levitt (2004) tested six theories to explain why Black children lost ground relative to Whites between fall kindergarten and spring first grade. The only hypothesis that received any empirical backing was differences in school quality. It was shown that, relative to Whites attending the same schools, Blacks lost only a small amount of ground. Both the Blacks and the Whites who attended schools with Blacks lost substantial amount on tests relative to Whites at other schools. But, evidence for the theory was far from conclusive. For instance, Fryer and Levitt (2004) were not able to explain the divergent trajectories of Blacks with any of the extensive observable school inputs provided by ECLS-K.[19] And, because of the link

17. There is some suggestive evidence in this regard, although we are unable to empirically validate the claims—thus treating it as an open question. Rock and Stenner (2005) argue that kindergarten test scores measure a child's "product" (e.g., readiness), whereas later tests measure a child's "process" (e.g., ways of learning). And it is thought that the latter is more correlated with "intelligence."

18. We have also tested whether racial differences in family disruption (i.e., compositional changes in the household or the number of times that a child changes schools) or frequency of repeating grades can explain the divergent trajectory of Blacks. Neither receives any empirical backing.

19. Indeed, Fryer and Levitt (2004) write, "There are important weaknesses in the argument that differential school quality explains the divergent trajectories of Whites and Blacks. First, the observable measures of school inputs included in Table 7 explain only a small fraction of the variation in student outcomes. For instance, adding the school input measures to our basic student-level test-score regressions only increases the R-squared of the regression by .05. Second, even after the school input measures are added to the test-score regressions, the gap between Blacks and Whites continues to widen. Third, both Hispanics and Asians also experience worse schools than Whites, but neither of those groups is losing ground. Because of these important weaknesses in the story—perhaps as a consequence of poor school quality measures in the data—the evidence linking school quality

between residential location and school attendance, school-fixed effects also capture neighborhood effects. Thus, it is not obvious how to separately identify the effect of school quality from one in which the influence of neighborhood quality on student outcomes grows with age.

When we revisit this hypothesis armed with more years of data, the empirical support for school quality being the primary source of divergent Black-White test scores weakens, as summarized in Table 7. This table compares estimates of the Black-White test score gap over time, with and without school-fixed effects. All of the specifications in the table include the parsimonious set of covariates, although only the coefficient on the Black-White gap is shown in the table. We eliminate students attending racially homogeneous schools from the sample. Blacks continue to lose substantial ground by the end of third grade. When school-fixed effects are included in the regression (columns 6–10), the Black-White test-score gap is identified off of differences between Blacks and Whites attending the same school. As reported in Fryer and Levitt (2004), the estimates of ground lost by Blacks shrinks to less than one-third of the magnitude in the full sample when comparing fall kindergarten and spring first grade test scores, and is not statistically different from zero in these specifications.[20] The additional data on third graders lead us to believe that school quality is less important than we had initially conjectured. A comparison of columns 5 and 10 make this clear. After including school-fixed effects, two-thirds of the difference between Blacks and Whites remain. Indeed, all of the ground lost between first grade and third grade by Blacks is within rather than across schools.

One explanation is that, because of tracking within schools, the educational experiences of Blacks and Whites might nonetheless be different even at the same school. The results presented in the table, however, are essentially unchanged when we include teacher-fixed effects, so that the

differences to the divergent trajectories of Blacks can be characterized as no more than suggestive."

20. This finding in some ways parallels Currie and Thomas' (1995) finding that early gains for students who attend Head Start tend to disappear because of low-quality schools that these students later attend. Consistent with Currie and Thomas (2005), we do not find a positive effect of Head Start on student test scores even in kindergarten, once other factors are controlled for. This finding is also related to Krueger and Whitmore (2002) and Phillips, Crouse, and Ralph (1998), who find that the Black-White gap widens as a result of poorer quality schools.

Table 7. Does Differential School Quality Explain Black Student's Losing Ground?

Subject	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Fall Kindergarten	**Spring First Grade**	**Spring Third Grade**	**Difference (2) – (1)**	**Difference (3) – (1)**	**Fall Kindergarten**	**Spring First Grade**	**Spring Third Grade**	**Difference (7) – (6)**	**Difference (8) – (6)**
Math	−.132	−.27	−.375	−0.138	**−0.243**	−.186	−.249	−.366	−0.063	**−0.180**
	(.033)	(.037)	(.040)	(.050)	**(.052)**	(.039)	(.043)	(.047)	(.058)	**(.061)**
Reading	.094	−.062	−.249	−0.156	**−0.343**	−.011	−.067	−.225	−0.056	**−0.214**
	(.037)	(.039)	(.041)	(.054)	**(.057)**	(.045)	(.046)	(.047)	(.064)	**(.065)**
Include school –fixed effects?	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes

Notes: Entries in columns (1)–(3) and columns (6)–(8) are estimates of the black-white test score gap at various points in the academic career. Columns (6)–(10) include school-fixed effects, so that the gaps are identified using within-school differences between black and white performance. Estimation is done using weighted least squares, with sample weights provided in the data set.

differences are identified off of children in the very same classroom. We conclude that neither school quality nor tracking within schools is the primary explanation for Black digression.

3.2. Does the importance of parental/environmental inputs grow as children age?

Black children may tend to grow up in environments less conducive to high educational attainment. If the importance of parental/environmental inputs grows as children age, one would expect to observe the *raw* gaps widening between Blacks and Whites, but to the extent our control variables adequately capture a child's environment, the residual gap after including all the covariates would remain constant. In fact, however, the residual gap increases more than the raw gap contradicting this explanation. Indeed, from a theoretical perspective, one might expect that the importance of parental inputs declines with age. Before reaching school age, the relative share of educational inputs provided by parents is very large. Once school starts, much of the burden for educating is shifted to the schools. Our empirical evidence does not, however, provide much support for this conjecture either.[21]

3.3. Did the type of material tested change to the detriment of Blacks?

One possible explanation for the divergent trajectories of Blacks and Whites relates to the nature of the material tested. Rock and Stenner (2005), for instance, hypothesize that the skills tested at kindergarten entry are less correlated with general intelligence then the later tests, and Blacks typically score better on achievement tests than on tests of aptitude. When starting school, kids know very little—irrespective of their

21. In a recent article, Todd and Wolpin (2004) have argued a slightly different point, which is that current test scores may reflect both current home and environmental inputs, as well as lagged values of these inputs. To the extent that current and lagged environments are highly correlated, controlling for current inputs is likely to yield similar estimates on average of the racial test score gap, and indeed, that is the case. Controlling for our parsimonious specification yields a coefficient on Black of –.382 in the third grade math regression and –.249 in reading. Adding lagged socioeconomic and home environment variables changes these coefficients to –.373 and –.273, respectively.

environment. As children age, their ability and social environment might matter more as they start to learn increasingly abstract concepts. It is in this regard that racial differences in home environment, parenting, peer group norms, and economic status could play a major role.[22] Arguing against such a hypothesis, as Rock and Stenner (2005) note, is the fact that test scores in the fall of kindergarten are strongly predictive of test scores in spring of third grade; if the two tests are capturing very different sets of skills, this would not be expected.[23]

To further explore whether "higher order" thinking could potentially explain the puzzle of Black underachievement, we investigate Black-White learning trajectories by specific types of skills. Table 8 summarizes the unadjusted means, by race, of children in fall kindergarten, spring first grade, and spring third grade on questions assessing specific sets of skills. In the raw data, Blacks lag Whites somewhat on virtually all types of questions, except those which are mastered by virtually students of a given grade level. That is true both at entry to school and even more so by the end of third grade. Note that the only skills with much variance by the end of third grade are those that are associated with concepts virtually no kindergartner had mastered: multiplication and division, rates and measurement, extrapolation, and so on.

Table 9 summarizes the results of probits, controlling for other covariates, of Black-White differences in mastery rates for particular skills. The dependent variable in the analysis is set equal to 1 if a student is assessed as having a 90% plus probability of having mastered a subject and is equal to 0 otherwise. The same set of covariates used earlier in the article is also included here, although the coefficients on these variables are not summarized in the table. The coefficients summarized in the table are the marginal effect of being Black, evaluated at the sample mean. Standard errors are reported in parentheses, and the mean level of mastery among Whites is reported in square brackets. Controlling for observables, upon entry to school, the gaps between Whites and Blacks tend to be small. This is true

22. This theory, if true, also re-introduces the possibility that genetics could play a role. Because we have little evidence on this either way, we choose to exclude it while noting that it is a possibility.

23. Nor does the answer appear to be that the kindergarten test is especially noisy. As Rock and Stenner (2005) note, the reliability of the kindergarten test is high.

Table 8. Unadjusted Means on Questions Assessing Specific Skills

Skill Tested	Fall Kindergarten				Spring First Grade				Spring Third Grade			
	White	Black	Hispanic	Asian	White	Black	Hispanic	Asian	White	Black	Hispanic	Asian
Math												
Count, number, shapes	.965	.897	.856	.966	0.999	0.998	0.999	1 (0)	1 (0)	1 (0)	1 (0)	1 (0)
	(.121)	(.211)	(.252)	(.116)	(.005)	(.025)	(.009)					
Relative size	.675	.423	.385	.667	.993	.970	.981	.996	.999	.999	.999	1 (0)
	(.340)	(.349)	(.363)	(.340)	(.044)	(.109)	(.074)	(.023)	(.001)	(.002)	(.001)	
Ordinality, Sequence	.291	.096	.106	.307	.970	.890	.916	.975	.999	.998	.999	.999
	(.349)	(.208)	(.232)	(.375)	(.123)	(.246)	(.207)	(.102)	(.004)	(.014)	(.013)	(.01)
Add/Substract	.058	.011	.014	.079	.808	.577	.638	.803	.984	.935	.959	.986
	(.157)	(.06)	(.067)	(.195)	(.157)	(.354)	(.348)	(.274)	(.064)	(.131)	(.107)	(.051)
Multiply/Divide	.005	.000	.001	.007	.339	.098	.151	.317	.857	.585	.705	.857
	(.054)	(.009)	(.019)	(.056)	(.363)	(.205)	(.263)	(.356)	(.247)	(.362)	(.334)	(.254)
Place Value	.000	.000	.000	.000	.045	.006	.011	.045	.516	.19	.303	.556
	(.003)	(.000)	(.000)	(.002)	(.137)	(.039)	(.061)	(.138)	(.390)	(.302)	(.357)	(.400)
Rate and Measurement	0 (0)	0 (0)	0 (0)	0 (0)	.004	0 (.003)	.001	.003	.206	.047	.085	.252
					(.031)		(.006)	(.022)	(.316)	(.160)	(.208)	(.342)
Reading												
Letter Recognition	.749	.600	.546	.790	.999	.994	.998	.999	1 (0)	1 (0)	1 (0)	1 (0)
	(.378)	(.424)	(.446)	(.349)	(.029)	(.060)	(.034)	.004)				
Beginning Sounds	.370	.206	.227	.415	.984	.949	.970	.988	.999	.999	.999	.999
	(.377)	(.307)	(.329)	(.397)	(.076)	(.146)	(.100)	(.062)	(.001)	(.003)	(.002)	(.001)
Ending Sounds	.216	.105	.121	.265	.956	.885	.922	.965	.999	.996	.998	.999
	(.302)	(.216)	(.233)	(.341)	(.120)	(.214)	(.165)	(.109)	(.006)	(.016)	(.009)	(.004)
Sight Words	.032	.013	.013	.077	.861	.696	.763	.891	.995	.975	.988	.997
	(.148)	(.095)	(.094)	(.245)	(.288)	(.397)	(.3630	(.261)	(.042)	(.099)	(.065)	(.028)

(Continued)

Table 8. Continued

Skill Tested	Fall Kindergarten				Spring First Grade				Spring Third Grade			
	White	Black	Hispanic	Asian	White	Black	Hispanic	Asian	White	Black	Hispanic	Asian
Words in Context	.012	.004	.004	.523	.322	.394	.624	.972	.890	.944	.987	
	(.093)	(.052)	(.176)	(.410)	(.380)	(.399)	(.406)	(.127)	(.248)	(.180)	(.079)	
Literal Inference	.004	.001	.002	.009	.19	.077	.112	.285	.861	.636	.762	.870
	(.047)	(.025)	(.041)	(.068)	(.315)	(.201)	(.246)	(.369)	(.266)	(.375)	(.330)	(.234)
Extropolation	0	0 (.004)	0 (.006)	0.001	0.032	0.011	0.017	0.049	0.353	0.144	0.234	0.329
	(.006)			(.007)	(.084)	(.040)	(.053)	(.106)	(.282)	(.178)	(.239)	(.272)
Evaluation	0	0 (.012)	.001	.002	.056	.017	.029	.106	.560	.267	.406	.501
	(.013)		(.020)	(.016)	(.154)	(.083)	(.108)	(.220)	(.365)	(.309)	(.367)	(.364)

Notes: The values reported in the table are unadjusted mean scores on specific areas of questions in fall of kindergarten, spring of first grade, and spring of third grade. These entries are proficiency probability scores, which are constructed using IRT scores and provide the probability of mastery of a specific set of skills.

Table 9. Performance Gaps on Questions Assessing Specific Skills

Skill Tested	Coefficient on Black		
	Fall Kindergarten	Spring First Grade	Spring Third Grade
Math			
Count, number, shapes	.020	–	–
	(.012)	–	–
	[.965]	[.999]	[1.000]
Relative size	−.051	−.002	–
	(.015)	(.003)	–
	[.675]	[.993]	[.999]
Ordinality, Sequence	−.023	−.017	–
	(.006)	(.009)	–
	[.349]	[.970]	[.999]
Add/Subtract	−.000	−.152	−.039
	(.000)	(.018)	(.009)
	[.058]	[.808]	[.984]
Multiply/Divide	–	−.027	−.179
	–	(.005)	(.019)
	[.005]	[.339]	[.857]
Rate and Measurement	–	–	−.016
	–	–	(.005)
	[0.000]	[.004]	[.206]
Place Value	–	–	−.090
	–	–	(.013)
	[0.000]	[.045]	[.516]
Reading			
Letter Recognition	.025	–	–
	(.019)	–	–
	[.749]	[.999]	[1.000]
Beginning Sounds	.035	−.008	–
	(.013)	(.005)	–
	[.370]	[.984]	[.999]
Ending Sounds	.014	−.019	−.002
	(.007)	(.011)	(.001)
	[.216]	[.956]	[.999]
Sight Words	.008	−.031	−.004
	(.005)	(.019)	(.003)
	[.032]	[.861]	[.995]
Words in Context	–	−.028	−.032
	–	(.017)	(.009)
	[.012]	[.523]	[.972]
Literal Inference	–	−.004	−.149
	–	(.009)	(.019)
	[.004]	[.190]	[.861]
Extrapolation	–	–	−.103

(Continued)

Table 9. Continued

Skill Tested	Coefficient on Black		
	Fall Kindergarten	Spring First Grade	Spring Third Grade
	–	–	(.020)
	[.000]	[.032]	[.353]
Evaluation	–	–	−.015
	–	–	(.003)
	[.000]	[.056]	[.560]

Notes: Coefficients are from probit regressions, with values reported in the table being marginal effects evaluated at the sample mean. The dependent variable is a dichotomous measure of skill mastery, defined to be equal to one if a student is assessed as having a 90% or greater likelihood of mastery in a given skill and equal to zero otherwise. The particular skill tested is reported in the left-hand column of the table. Although not reported in the table, the specifications include the full set of other controls used in regressions reported in prior tables. Standard errors are in parentheses. The mean of the dependent variable for White students is reported in square brackets. In cases where virtually no students or virtually all students have mastered a subject, we do not report results.

on both math and reading skills, and regardless of whether the skill is mastered by many of the students or relatively few of the students. Over time, Black students lose ground in virtually every skill area, except the most basic skills that are mastered by virtually all students in the grade. In addition and subtraction, which is challenging for many first graders regardless of race, the Black students lag significantly in first grade, but both Blacks and Whites achieve almost complete mastery by third grade. In that subject, as well as some of the basic reading skills like "words in context," a few percent of Blacks fail to master the material, even though almost all Whites do. Multiplication and division, as well as "literal inference," display a pattern that is far more disturbing. By the spring of third grade, over 85% of White students have mastered these subjects, but mastery rates are 15–20% lower for Blacks, even after controlling for other factors.

It is difficult to know precisely what conclusion to draw from these results. To the extent that the pattern of Black skill acquisition as student age follows the path of the basic skills, that is, Black students master the material, but at a somewhat later age than White students, the patterns may be construed as encouraging. The implication would be that Black students, although lagging Whites at any particular point in time, are on parallel trajectories. Much more troubling, it would seem, is the possibility

that as the skills become more difficult, for example, division, a non-trivial fraction of the Black students may never master the skills. If these skills are inputs into future subject matter, then the racial gap may be further magnified. While the data available thus far cannot speak definitively in determining which of these scenarios is more likely, the patterns in Table 9 do raise the specter of the latter scenario being possible.[24]

4. Conclusion

The racial achievement gap remains a stubborn reality. Using newly collected data on a recent cohort from the ECLS, we document substantial Black-White test score gaps in both math and reading that grow at approximately 10 standard deviations per year that children are in school. The divergence in test scores relative to Whites is not apparent for either Hispanics or Asians.

The explanation as to why Blacks are losing ground proves elusive. Fryer and Levitt (2004) test a wide range of hypotheses, finding some empirical support for only one explanation: differential school quality across races. When the data are extended to cover an additional two years of schooling, however, the support for even this hypothesis weakens. We also explore whether the growing racial test score gap could be attributed to the inherent difficulties in testing achievement at especially young ages or the possibility of increasing importance of home inputs for the development of higher order thinking but can provide no compelling evidence confirming these hypotheses either.

Data Appendix

The ECLS-K is a nationally representative sample of 21,260 children entering kindergarten in 1998. Thus far, information on these children has been gathered at four separate points in time. The full sample was interviewed in the fall and spring of kindergarten and spring of first grade. All of our regressions and summary statistics are weighted, unless otherwise noted, and we include dummies for missing data. We describe below how we combined and re-coded some of the ECLS variables used in our analysis.

24. Arguing against that hypothesis is the fact that the estimated labor market returns to a marginal year of education are typically found to be at least as high for Blacks as for Whites, suggesting that Blacks continue to learn skills valued by the labor market at a pace equal to Whites throughout the course of their education.

Socioeconomic Composite Measure

The SES variable was computed by ECLS at the household level for the set of parents who completed the parent interview in Fall Kindergarten or Spring Kindergarten. The SES variable reflects the socioeconomic status of the household at the time of data collection for spring kindergarten. The components used for the creation of SES were Father/male guardian's education; Mother/female guardian's education; Father/male guardian's occupation; Mother/female guardian's occupation; and Household income.

Number of Children's Books

Parents/guardians were asked "How many books does your child have in your home now, including library books?" Answers ranged from 0 to 200.

Child's Age

We used the Child's Age at Assessment Composite variable provided by ECLS. The Child's age was calculated by determining the number of days between the child assessment date and the child's date of birth. The value was then divided by 30 to calculate the age in months.

Birth Weight

Parent's were asked how much their child weighed when they were born. We multiplied the pounds by 16 (and added it to the ounces) to calculate birth weight in ounces.

Mother's Age at First Birth

Mothers were asked how old they were at the birth of their first child.

References

Caneiro, Pedro, and James Heckman. 2003. "Human Capital Policy," NBER Working Paper No. 9495.

Coleman, James, Ernest Q. Campbell, and Carol J. Hobson. 1966. *Equality of Educational Opportunity*. Washington DC: U.S. Government Printing Office.

Campbell, J. M., C. M. Hombo, and J. Mazzeo. 2000. *NAEP 1999 Trends in Academic Progress: Three Decades of Student Performance.* NCES 2000469.

Currie, Janet, and Duncan Thomas. 1995. "Does Head Start Make a Difference?" 85(3) *American Economic Review* 341–64.

Ferguson, Ronald. 2001. "A Diagnostic Analysis of Black-White GPA Disparities in Shaker Heights, Ohio." Brooking Papers on Education Policy.

———. 2002. "What Doesn't Meet the Eye: Understanding Racial Disparities in Fifteen Suburban School Districts," mimeo. Harvard University.

Fryer, Roland, and Steven Levitt. 2004. "Understanding the Black-White Test Score Gap in the First Two Years of School," 86(2) *The Review of Economics and Statistics* 447–64.

Jencks, Christopher, and Meredith Phillips. 1998. "America's Next Achievement Test: Closing the Black-White Test Score Gap," 9(40) *American Prospect* 44–53.

Jones, Lyle V., Nancy Burton, and Ernest Davenport. 1982. *Mathematics Achievement Levels of Black and White Youth*. Chapel Hill: University of North Carolina, L.L. Thurstone Psychometric Laboratory.

Krueger, Alan, and Diane Whitmore. 2002. "Would Smaller Classes Help Close the Black White Achievement Gap?" In John Chub and Tom Loveless, eds., *Bridging the Achievement Gap*. Washington, DC: The Brookings Institute.

Neal, Derek. 2005. "Why Has Black-White Skill Convergence Stopped?" NBER Working Paper No. 11090.

Ogbu, John and Astrid Davis. 2003. *Black American Students in an Affluent Suburb: A Study of Academic Disengagement.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Phillips, Meredith. 2000. "Understanding Ethnic Differences in Academic Achievement: Empirical Lessons from National Data." In David Grissmer and Michael Ross, eds., *Analytic Issues in the Assessment of Student Achievement*.Washington DC: U.S. Department of Education, National Center for Education Statistics, 103–132.

Phillips, Meredith, James Crouse, and John Ralph. 1998. "Does the Black-White Test Score Gap Widen After Children Enter School?" In Christopher Jencks and Meredith Phillips, eds., *The Black-White Test Score Gap*. Washington, DC: The Brookings Institute.

Rock, Donald, and A. Jackson Stenner. 2005. "Assessment Issues in the Testing of Children at School Entry," 15(1) *The Future of Children* 15–34.

Todd, Petra and Kenneth Wolpin. 2004. "The Production of Cognitive Achievement in Children: Home, School and Racial Test Score Gaps," PIER Working Paper No. 04–019.