4

Losing Ground at School

Roland G. Fryer, Jr., Harvard Society of Fellows and National Bureau of Economic Research and Steven D. Levitt, American Bar Foundation and University of Chicago

The black-white test score gap is a robust empirical regularity. A simple comparison of mean test scores typically finds black students scoring roughly one standard deviation below white students on standardized tests. Even after controlling for a wide range of covariates including family structure, socioeconomic status, measures of school quality, and neighborhood characteristics, a substantial racial gap in test scores persists.[1]

Gaining a better understanding of the underlying causes of the test-score gap is a question of great importance. Neal and Johnson and O'Neill find that most of the observed black-white wage differentials among adults disappears once lower eighth-grade test scores among Blacks are taken into account.[2] Thus, eliminating the test-score gap that arises by the end of junior high school may be a critical component of reducing racial wage inequality.[3]

A wide variety of possible explanations for the test-score gap have been put forth. These explanations include differences in genetic make-up (see Hernstein and Murray and Jensen), differences in family structure and poverty (see Armor, Brooks-Gunn and Duncan, Mayer, and Phillips, Crouse, and Ralph), differences in school quality (see Cook and Evans), racial bias in testing or teachers' perceptions (see Delpit, Ferguson, and Rodgers and Spriggs), and differences in culture, socialization, or behavior (see Cook and Ludwig, Fordham and Ogbu, Fryer, and Steele and Aronson).[4] The appropriate public policy choice (if any) to address the test-score gap depends critically on the underlying source of the gap.

In this paper, we use the Early Childhood Longitudinal Study Kindergar-

.................15683$ $CH4 10-05-05 08:50:08 PS PAGE 88

ten Cohort (ECLSK) to shed new light on the test-score gap. ECLS-K is a new data set administered by the Department of Education. The survey covers a sample of more than 20,000 children entering kindergarten in the fall of 1998. An enormous amount of information is gathered for each individual including family background, school and neighborhood characteristics, teacher and parent assessments, and test scores. The original sample of students has subsequently been reinterviewed in the spring of kindergarten and first grade.

The results we obtain using these new data are informative and in some cases quite surprising. As in previous data sets, we observe substantial racial differences in test scores in the raw data: black kindergartners score on average .64 standard deviations worse than Whites. In stark contrast to earlier studies (including those looking at kindergartners), however, after controlling for a small number of other observable characteristics (children's age, child's birth weight, a socioeconomic status measure, WIC participation, mother's age at first birth, and number of children's books in the home), we essentially eliminate the black-white test score gap in math and reading for students entering kindergarten.[5] Controlling for a much larger set of characteristics yields the same conclusion. This same set of covariates accounts for much but not all of the Hispanic-white difference in test scores, but cannot explain the high test-scores of Asians.

There are three leading explanations for why our results differ so sharply from earlier research such as Phillips, Crouse, and Ralph (1998): (1) nonrandom sampling in the data sets used in earlier studies, (2) real gains by recent cohorts of Blacks, and (3) better covariates in ECLS. Based on our analysis of the Children of the National Longitudinal Survey of Youth (CNLSY) data used by Phillips, Crouse, and Ralph, we conclude that real gains by recent cohorts of Blacks are an important part of the explanation. The raw black-white test-score gap for recent cohorts in CNLSY are comparable to those in ECLS, in sharp contrast to earlier cohorts in CNLSY. Real gains by Blacks born in recent years would appear to be the leading explanation. We cannot, however, fully eliminate the racial test score gap among recent CLNSY cohorts. This is due in part to better covariates in ECLS. Even when nearly identical covariates are included, differences persist between ECLS and CNLSY.

Despite the fact that we see no difference in initial test scores for observationally equivalent black and white children when they enter kindergarten, their paths diverge once they are in school. Between the beginning of kindergarten and the end of first grade, black students lose .20 standard deviations (approximately .10 standard deviation each year) relative to white students with similar characteristics.[6] If the gap in test scores for these children continues to grow at the same rate, by fifth grade the black students will be .50 standard deviations behind their white counterparts—a gap similar in magnitude to that found in previous analyses (see Jones et al., Phillips, and Phillips, Crouse, and Ralph).[7]

The leading explanation for the worse trajectory of black students in our sample is that they attend lower quality schools. When we compare the change in test scores over time for Blacks and Whites attending the same school, black students lose only a third as much ground as they do relative to Whites in the overall sample. This result suggests that differences in quality across schools attended by Whites and Blacks is likely to be an important part of the story. Interestingly, along ''traditional'' dimensions of school quality (class size, teacher education, computer:student ratio, etc.), Blacks and Whites attend schools that are similar. On a wide range of nonstandard school inputs (e.g., gang problems in school, percent of students on free lunch, amount of loitering in front of school by nonstudents, amount of litter around the school, whether or not students need hall passes, and PTA funding), Blacks do appear to be attending much worse schools even after controlling for individual characteristics.[8] Our story is incomplete, however, because the observable differences across schools do little to explain the widening black-white gap. This could be due to the coarseness of the school quality variables available in the ECLS.

We explore a range of other explanations as to why black children are losing ground, but find very little empirical support for these alternative theories. Black students do not appear to suffer bigger ''summer setbacks'' when school is not in session. The lower trajectories of black students are not simply an artifact of standardized testing. Subjective teacher assessments of student performance yield patterns similar to the test-score data. Having a black teacher provides no benefit to black students compared to their white classmates, calling into question the possible role of either overt discrimination or low expectations for black children on the part of white teachers. Finally, adding proxies for behavioral problems does not alter our findings.

The structure of the paper is as follows. The first section provides a brief review of the literature. The second section describes and summarizes the data set. Then we present the basic results for incoming kindergartners, demonstrating that the black-white test score gap disappears once other confounding factors are accounted for. In the next section we document the fact that a racial test-score gap emerges during the school-age years, and the following section analyzes the reasons for this divergence. We present our conclusions in the final section.

BACKGROUND AND PREVIOUS LITERATURE

The Coleman Report (Coleman et al.) was the first national study to describe ethnic differences in academic achievement among children at various stages of schooling. It documented that substantial differences in educational achievement between Blacks and Whites not only existed at every grade level, but increased with student age. Since then, substantial effort has been devoted to understanding what variables account for the gap, as well as how and why the magnitude of the gap has changed over time.[9] A number of stylized facts have emerged. Socioeconomic status and the effects of poverty are important factors in explaining racial differences in educational achievement (see Brooks-Gunn and Duncan, Mayer, Brooks-Gunn et al.).[10] Even after controlling for socioeconomic status in conventional regression analysis, a substantial gap still remains. That gap has generally been declining over time, although for high school students today, the gap is slightly larger than it was in the late 1980s (see Grissmer et al., Hedges and Nowell, and Humphreys).[11] Finally, the gap in test scores between Blacks and Whites historically emerges before children enter kindergarten and tends to widen over time (see Carneiro and Heckman and Phillips, Crouse, and Ralph).

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 four separate points in time. The full sample was interviewed in the fall and spring of kindergarten and spring of first grade. A random sample of one-fourth of the respondents were also interviewed in the fall of first grade. The sample will eventually be followed through fifth grade.[12] Roughly 1,000 schools are included in the sample, with an average of more than twenty children per school in the study. As a consequence, it is possible to conduct within-school analyses.

ESTIMATING RACIAL TEST SCORE GAPS FOR INCOMING KINDERGARTNERS

Table 4.1 presents a series of estimates of the racial test score gap for the tests taken in the fall of kindergarten. The specifications estimated are of the form

$$TESTSCORE_i = RACE_i\text{'}G + X_i\text{'}T + e_i (1)$$

where i indexes students. A full set of race dummies are included in the regression, with White as the omitted category. Consequently, the coefficients on race capture the gap between the named racial category and Whites. Our primary emphasis, is on the black-white test score gap. The vector of other covariates included in the specification, denoted X_i, varies across columns in table 4.1. As one moves to the right in the table, the set of covariates steadily grows. In all instances, the estimation is done using weighted least

Table 4.1. The Estimated Black–White Test Score Gap in Fall of Kindergarten

	Math					*Reading*				
Variables	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*	*(10)*
Black	−.638	−.368	−.238	−.094	−.102	−.401	−.134	−.006	.117	.093
	(.022)	(.022)	(.023)	(.023)	(.026)	(.024)	(.025)	(.026)	(.025)	(.030)
Hispanic	−.722	−.429	−.302	−.203	−.171	−.427	−.223	−.137	−.064	−.076
	(.022)	(.023)	(.024)	(.022)	(.028)	(.027)	(.026)	(.026)	(.025)	(.029)
Asian	.150	.070	.190	.265	.274	.335	.256	.371	.409	.375
	(.056)	(.051)	(.051)	(.048)	(.050)	(.064)	(.059)	(.059)	(.058)	(.060)
Other race	−.503	−.329	−.253	−.158	−.113	−.401	−.230	−.155	−.072	−.014
	(.041)	(.037)	(.036)	(.035)	(.035)	(.044)	(.040)	(.040)	(.038)	(.039)
Socioeconomic status composite measure		.456	.389	.302	.072		.451	.393	.299	.092
		(.014)	(.014)	(.014)	(.024)		(.014)	(.015)	(.015)	(.023)
Number of children's books			.007	.006	.005		.007	.006	.004	
			(.001)	(.001)	(.001)			(.001)	(.001)	(.001)
(Number of children's books)^2 (*1,000)			−.023	−.020	−.027			−.025	−.021	−.017
			(.003)	(.002)	(.016)			(.003)	(.003)	(.017)
Female				.010	.000				.159	.153
				(.015)	(.015)				(.017)	(.016)

Age at kindergarten fall (in months)				.056 (.002)	−2.680 (.542)				.041 (.002)	−2.409 (.483)
Birth weight (ounces) (*10)				.029 (.004)	.030 (.004)				.019 (.004)	.022 (.004)
Teenage mother at time of first birth				−.109 (.018)	−.029 (.021)				−.144 (.020)	−.069 (.022)
Mother at least thirty at time of first birth				.182 (.025)	.111 (.028)				.226 (.027)	.155 (.030)
WIC participant				−.211 (.019)	−.120 (.020)				−.184 (.021)	−.104 (.021)
R-squared	0.108	0.223	0.239	0.317	0.354	0.045	0.16	0.175	0.233	0.279
Number of observations			13290					12601		
Full set of covariates included in regression?	N	N	N	N	Y	N	N	N	N	Y

Notes: The dependent variable is the math or reading test score in the fall of kindergarten. Test scores are IRT scores, normalized to have a mean of zero and a standard deviation of one in the full, unweighted sample. Non-Hispanic Whites are the omitted race category, so all of the race coefficients are gaps relative to that group. The unit of observation is a student. Standard errors are in parentheses. Estimation is done using weighted least squares, using sample weights provided in the data set. In addition to the variables included in the table, indicator variables for students with missing values on each covariate are also included in the regressions. In addition, columns 5 and 10 report only a subset of the coefficients from regressions with ninety-eight covariates included in the specification. The full results for columns 5 and 10 are reported in Fryer and Levitt. Note that the specifications in columns 5 and 10 include age and age squared; that is why the coefficient on age changes so dramatically relative to other columns in the table.

squares, with weights corresponding to the sampling weights provided in the data set.

The first and sixth columns of table 4.1 presents the differences in means, not including any covariates. These results simply reflect the raw test score gaps. The next specification adds the composite indicator of socioeconomic status constructed by the ECLS survey administrators. Socioeconomic status is an important predictor of incoming test scores, carrying a t-statistic over forty. A one-standard deviation increase in the SES variable is associated with a .41 increase in both math and reading test scores. Controlling for socioeconomic status substantially reduces the estimated racial gaps in test scores (see also Coley). The black-white gap in math falls by more than 40 percent; the reading gap is reduced by more than two-thirds. The changes in the other race coefficients are not as large, but in every instance the estimated gaps shrink, and R-squared increases substantially.

The next set of specifications adds the number of children's books in the child's home, the square of that variable, and an indicator variable equal to one if the number of books takes on a missing value for that student. The number of books is strongly positively associated with high kindergarten test scores on both math and reading.[13] 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 (.115) in math and reading respectively. This variable seems to serve as a useful proxy for capturing the conduciveness of the home environment to academic success. Including number of books reduces the black-white gap on math to less than one-fourth of a standard deviation and completely eliminates the gap in reading. The gap for Hispanics also shrinks. The Asian-white gap, however, becomes even larger than the raw gap when number of books is added to the regression.

Columns 4 and 9 add controls for 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 twenties), and WIC participation. These covariates generally enter with the expected sign. Older children, those with higher birth weights, those with older mother's at the time of first birth all score better. 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. Adding these variables to the specification further improves the test scores of Blacks and Hispanics. In fact, the estimates suggest that, controlling for other factors, black children actually score slightly better than Whites in reading, and only slightly worse in math. We do not have a compelling explanation as to why there is a difference between reading and math achievement.

Only a small gap persists for Hispanics. The advantage enjoyed by Asians becomes even greater. R-squared increases substantially relative to the previous specification.

The final specifications in table 4.1 (columns 5 and 10) include an exhaustive set of roughly 100 covariates capturing city size, neighborhood characteristics, region of the country, parental education, parental income, parental occupational status, family size and structure, whether the mother worked, type of preschool program participation, whether English is spoken at home, and the extent of parental involvement in a child's life and school. We report only a subset of the covariates in table 4.1; full results can be seen in Fryer and Levitt.[14] Almost all of the controls enter in the predicted direction and with coefficients of plausible magnitude. Interestingly, none of the coefficients on race change appreciably. Only a few of the parameters on the controls included in the parsimonious specifications are greatly affected either, and these are easily explained. The socioeconomic status coefficient shrinks because the full set of covariates includes variables that go into the construction of the composite indicator such as parent's income and occupational status. The coefficient on age becomes highly negative because an age-squared term (which is positive and significant) is included in the full specification. The inclusion of these additional variables does little to improve the fit of the model.

Table 4.2 explores the sensitivity of the estimated racial gaps in test scores across a wide variety of alternative specifications and subsamples 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 columns 4 and 9 of table 4.1).

Weighting all of the observations equally in the regressions leaves the black-white gap in math and reading remain virtually unchanged. Using an alternative test-score measure (T-scores, which are norm-referenced measurements of achievement) has very little impact on the results.

One might be concerned that restricting all the coefficient estimates to be identical across the entire sample may yield misleading results. Regressions on a common support (e.g., only on single mothers, region of the country, or only in rural areas) provide one means of addressing this concern. Almost every subset of the data examined yields results roughly similar to those for the overall sample. There is some slight evidence that black females do better relative to Whites than do black males. The results appear to be quite consistent across quintiles of the socioeconomic status distribution. Due in part to relatively imprecise estimates, the equality of black and white test scores on math and reading tests can rarely be rejected for any of the quintiles. Rural Blacks do somewhat worse relative to Whites than those in central cities.

Table 4.2. Sensitivity Analysis/Extensions of the Basic Model for Fall Kindergarten Test Scores

	Coefficient on Black for:		Coefficient on Hispanic for:		Coefficient on Asian for:	
Specification	*Math*	*Reading*	*Math*	*Reading*	*Math*	*Reading*
Baseline	−.094 (.023)	.117 (.025)	−.203 (.022)	−.064 (.025)	.265 (.048)	.409 (.058)
Unweighted	−.100 (.023)	.092 (.024)	−.206 (.021)	−.057 (.024)	.285 (.034)	.387 (.035)
Other test score measures						
T-scores	−.050 (.024)	.141 (.030)	−.057 (.022)	.065 (.028)	.176 (.040)	.298 (.048)
By Gender						
Males	−.126 (.034)	.093 (.037)	−.224 (.032)	−.095 (.035)	.338 (.078)	.385 (.087)
Females	−.058 (.030)	.147 (.035)	−.181 (.031)	−.035 (.036)	.203 (.059)	.433 (.077)
By SES Quintile:						
Bottom	−.092 (.044)	−.005 (.041)	−.202 (.044)	−.133 (.045)	.328 (.143)	.043 (.111)
Second	−.088 (.045)	.091 (.049)	−.179 (.046)	−.090 (.047)	.044 (.106)	−.001 (.090)
Third	−.097 (.049)	.068 (.045)	−.242 (.046)	−.106 (.051)	.249 (.121)	.351 (.167)
Fourth	−.082 (.058)	.292 (.077)	−.100 (.056)	.030 (.057)	.207 (.088)	.396 (.115)
Top	−.169 (.080)	.068 (.085)	−.323 (.078)	−.113 (.094)	.404 (.087)	.724 (.102)
By family structure:						
Single mother	−.087 (.043)	.070 (.043)	−.197 (.048)	−.119 (.047)	.086 (.149)	.114 (.144)
Two biological parents	−.127 (.034)	.141 (.042)	−.176 (.029)	−.033 (.033)	.291 (.054)	.456 (.064)
Teen mother at 1st birth	−. 101(.036)	.014 (.033)	−.199 (.036)	−.127 (.038)	.170 (.105)	.251 (.114)
Teen mother at child's birth	−.062 (.046)	−.021 (.043)	−.196 (.045)	−.105 (.052)	.279 (.141)	.281 (.135)

(continues)

Table 4.2. Continued

	Coefficient on Black for:		*Coefficient on Hispanic for:*		*Coefficient on Asian for:*	
Specification	*Math*	*Reading*	*Math*	*Reading*	*Math*	*Reading*
By region:						
Northeast	−.087 (.060)	.129 (.076)	−.159 (.054)	−.030 (.060)	.305 (.124)	.483 (.156)
Midwest	.004 (.053)	.093 (.057)	−.140 (.064)	−.031 (.061)	.337 (.119)	.562 (.133)
South	−.153 (.032)	.051 (.033)	−.217 (.040)	−.119 (.048)	.154 (.104)	.368 (.111)
West	.098 (.077)	.362 (.095)	−.200 (.044)	−.001 (.048)	.269 (.071)	.353 (.088)
By location type:						
Central city	−.110 (.035)	.147 (.040)	−.235 (.033)	−.073 (.037)	.271 (.061)	.439 (.075)
Suburban	−.135 (.039)	.030 (.041)	−.261 (.041)	−.145 (.042)	.146 (.102)	.310 (.119)
Rural	−.184 (.048)	−.032 (.050)	−.253 (.062)	−.124 (.072)	.255 (.130)	.126 (.102)
By school type:						
Public	−.106 (.024)	.098 (.027)	−.214 (.024)	−.081 (.027)	.260 (.051)	.392 (.064)
Private	.022 (.070)	.281 (.074)	−.152 (.058)	.015 (.066)	.296 (.135)	.479 (.137)
School >80% Black	.053 (.269)	−.016 (.215)	−.084 (.298)	.057 (.273)	.285 (.382)	.788 (.641)
School >80% White	−.105 (.047)	.059 (.053)	−.186 (.025)	−.061 (.028)	.288 (.054)	.436 (.065)

Notes: Specifications in this table are variations on those reported in columns 4 and 9 of table 4.1. Only the race coefficients are reported in this table. The top row of the table simply reproduces the baseline results in columns 4 and 9 of table 4.1. The remaining rows of the table correspond to different weights, test score measures, or particular subsets of the data. For further details of the baseline specification, see the notes to table 4.1

Blacks in private schools appear to do especially well, consistent with Neal and Grogger and Neal.[15]

The fact that the black-white test score gap essentially disappears with the inclusion of sufficient controls in ECLS is a very striking result given that in past research a substantial gap has persisted, regardless of the age of the individuals, the particular tests, or the covariates included (e.g., Hernstein and Murray, Neal and Johnson, Phillips, Crouse, and Ralph).[16] The most direct comparison to our research among previous studies is Phillips, Crouse, and Ralph, which looks at test outcomes for kindergartners in the early cohorts of CNLSY. Although Phillips, Crouse, and Ralph have the greatest success among earlier studies in explaining the racial differences in reading (they reduce the gap by two-thirds with their covariates), their raw gap is so large compared to ECLS that the *residual* gap in that paper is almost as large as the *raw* gap in ECLS.

Why our results differ so sharply from previous research, and Phillips, Crouse, and Ralph, in particular, is a question of critical importance. There are three leading explanations for the divergence: (1) the sample of births included in CNLSY, especially in the early years, may be nonrepresentative; (2) better covariates are available in ECLS; and (3) Blacks born into recent cohorts have made real gains relative to Blacks born a decade earlier. The first two explanations appear to play only a small role empirically. While it is true that the sample of births in early cohorts of CNLSY analyzed by Phillips, Crouse, and Ralph is heavily skewed toward teenage mothers, because of the way the sample is generated (i.e., by births to those included in NLSY), the nonrandom sampling, does not seem to provide the explanation for the differing results. When we restrict our ECLS sample to only include children born to teen mothers, our results are virtually unchanged.[17] When we try to estimate specifications in ECLS using only variables that are available in CNLSY, Blacks do somewhat worse than in our baseline sample (a gap of −.183 on math and .034 on reading), but this is nothing like the residual gap of −.67 on reading in Phillips, Crouse, and Ralph.

Real gains by Blacks in recent cohorts, in contrast, does appear to be an important part of the divergence between our results and past research. Limiting the CNLSY to cohorts born in the same years as the ECLS sample, the raw test score gaps in the CNLSY are nearly half as large as in earlier cohorts of CNLSY used by Phillips, Crouse, and Ralph and are remarkably close to those found in the ECLS. On the math skills test, the raw gaps are .638 and .665 respectively in ECLS and CNLSY. For reading, the gap is .401 in ECLS and .540 in the CNLSY. Real gains by Blacks in recent years could explain this result. Interestingly, however, using the same set of controls that yield math and reading gaps in ECLS of −.183 and .034 respectively, in recent cohorts of the CNLSY the estimated black-white residual gaps are −.500 and −.41 on math and reading. Thus, although the raw gaps are similar in

ECLS and recent cohorts of CNLSY, larger residual gaps remain in CNLSY for reasons we cannot explain.

THE EVOLUTION OF THE RACIAL TEST SCORE GAPS AS CHILDREN AGE

The results of the previous section demonstrate that although black test scores lag Whites by a large margin, the inclusion of a small number of covariates eliminates any systematic differences in the math and reading performances of Whites and Blacks entering kindergarten. Hispanics somewhat lag Whites, and Asians exceed all of the other races. In this section, we explore how those racial gaps change over time.

In terms of raw test scores, black students lose some ground relative to Whites between the fall of kindergarten and the spring of first grade: .090 standard deviations on math and .128 standard deviations on reading. Table 4.3 presents regression results for those two time periods. We report results only from our "parsimonious" regression specification; similar racial gaps emerge when the exhaustive set of covariates is included. Controlling for other factors in the regressions, black students appear to lose much more ground than they do in the raw means: −.156 standard deviations on math and −.188 standard deviations on reading.[18] If black students in the sample continue to lose ground through ninth grade at the rate experienced in the first two years of school, they will lag white students on average by a full standard deviation in raw math and reading scores and over two-thirds of a standard deviation in math even after controlling for observable characteristics (substantially smaller for reading). Raw gaps of that magnitude would be similar to those found in previous studies of high school age children (see Grissmer, Flanagan, and Williamson, Hedges and Nowell, Humphreys, Phillips, and Phillips, Crouse, and Ralph).

In striking contrast to the black-white gap, Hispanics show gains relative to Whites between the beginning of kindergarten and the end of first grade. Asians lose roughly as much ground as Blacks on math (although they start ahead of Whites) and also fall slightly on reading. Thus, black students are not only losing ground relative to Whites, but even more so relative to Hispanics, and somewhat less compared to Asians.

WHY ARE BLACK STUDENTS LOSING GROUND IN THE FIRST TWO YEARS OF SCHOOL?

Understanding why black students fare worse in the first two years of school is a question of paramount importance for two reasons. First, knowing the

Table 4.3. The Evolution of Test Score Gaps by Race as Children Age

	Math			*Reading*		
Variables	*Fall kindergarten*	*Spring kindergarten*	*Spring first grade*	*Fall kindergarten*	*Spring kindergarten*	*Spring first grade*
Black	−.094 (.023)	−.201 (.025)	−.250 (.028)	.117 (.025)	−.009 (.027)	−.071 (.029)
Hispanic	−.203 (.022)	−.187 (.024)	−.120 (.026)	−.064 (.025)	−.005 (.027)	.001 (.029)
Asian	.265 (.048)	.221 (.049)	.115 (.044)	.409 (.058)	.434 (.054)	.345 (.045)
Other race	−.158 (.035)	−.166 (.039)	−.195 (.042)	−.072 (.038)	−.099 (.039)	−.163 (.042)
SES composite measure	.302 (.014)	.284 (.014)	.263 (.014)	.299 (.015)	.280 (.015)	.284 (.014)
Number of Books	.006 (.001)	.006 (.001)	.005 (.001)	.006 (.001)	.005 (.001)	.006 (.001)
(Number of Books) (squared) (*1000)	.020 (.002)	−.019 (.003)	−.019 (.003)	−.021 (.003)	−.020 (.003)	−.022 (.003)
Female	.010 (.015)	.003 (.016)	−.033 (.017)	.159 (.017)	.195 (.017)	.216 (.017)
Age at kindergarten fall (in months)	.056 (.002)	.051 (.002)	.036 (.002)	.041 (.002)	.034 (.002)	.021 (.002)
Birth weight (ounces) (*10)	.029 (.004)	.003 (.000)	.029 (.004)	.019 (.004)	.002 (.000)	.024 (.005)
Teenage mother at time of first birth	−.109 (.018)	−.112 (.021)	−.111 (.022)	−.144 (.020)	−.138 (.021)	−.131 (.024)
Mother in 30s at time of first birth	.182 (.025)	.127 (.024)	.093 (.022)	.226 (.027)	.158 (.025)	.085 (.024)
WIC Participant	−.211 (.019)	−.195 (.020)	−.201 (.021)	−.184 (.021)	−.152 (.02)	−.182 (.022)
R-squared	0.317	.282	.240	0.233	0.197	.194
Number of Obs.	13290	13,290	13,290	12601	12601	12,601

Notes: The dependent variable is fall kindergarten test scores in columns 1 and 3 and spring first grade test scores in columns 2 and 4. All specifications include the parsimonious set of controls corresponding to columns 4 and 9 of table 4.1. Test scores are IRT scores, normalized to have a mean of zero and a standard deviation of one in the full, unweighted sample. Non-Hispanic Whites are the omitted race category, so all of the race coefficients are gaps 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. In addition to the variables included in the table, indicator variables for students with missing values on each covariate are also included in the regressions.

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 schools on average; (2) the importance of parental/environmental contributions may grow over time. Since black children are on average disadvantaged in this regard, they fall behind; and (3) because of worse home and neighborhood environments, black students suffer worse ''summer setbacks'' when school is not in session.[19] We address each of these hypotheses in turn.

Are Black Students Losing Ground Because They Attend Worse Schools?

There is substantial racial segregation in school attendance in the United States. Our data samples roughly twenty children each from approximately 1,000 schools. In 35 percent of those schools, there is not a single black child in the sample.[20] The mean black student in our sample attends a school that is 59 percent black and 8 percent Hispanic. In contrast, the typical white student goes to a school that is only 6 percent black and 5 percent Hispanic. Given that Blacks and Whites have relatively little overlap in the schools they attend, differences in school quality are plausible explanations for why black students are losing ground.[21]

Because our data set has many individuals from each school included in the sampling frame, school-fixed effects can be included in the estimation. With school-fixed effects, the estimated black-white test score gap is identified off of the relative performance of Blacks and Whites attending the same school, as opposed to across schools. To the extent that differential average school quality across races is the complete explanation for the widening racial test score gap, one would predict that the gap should not widen over time when comparing Blacks and Whites attending the same school. There are, of course, thorny issues of sample selection that potentially complicate the interpretation of these results: white students who elect to attend schools with black students may have differential test score trajectories than other white students, even if they had gone to all white schools. Nonetheless, looking within schools provides a first attempt at testing this hypothesis.

The comparison of changes in the black-white test score gap over time including and excluding school-fixed effects is presented in table 4.4. 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. The first three columns reflect the full sample of students. The remaining columns restrict the sample to schools that have both black and white chil-

Table 4.4. Does Differential School Quality Explain Black Students Losing Ground: A Comparison of Cross-school and Within-school Estimates of the Test Score Trajectory by Race (Values reported in table are the coefficient on the variable Black)

	Full Sample of Students			*Excluding Students Attending All-White Schools*					
Subject	*(1) Fall kindergarten*	*(2) Spring first grade*	*(3) Difference (2)-(1)*	*(4) Fall kindergarten*	*(5) Spring first grade*	*(6) Difference (5)-(4)*	*(7) Fall kindergarten*	*(8) Spring first grade*	*(9) Difference (8) − (7)*
Math	−.094	−.250	−.156	−.136	−.261	−.125	−.175	−.222	−.047
	(.023)	(.028)	(.036)	(.028)	(.034)	(.044)	(.034)	(.040)	(.052)
Reading	.117	−.071	−.188	.072	−.084	−.156	−.007	−.057	−.05
	(.025)	(.029)	(.038)	(.030)	(.035)	(.046)	(.038)	(.042)	(.057)
Include school-fixed effects in regression?	N	N	N	N	N	N	Y	Y	Y
Number of Obs.	13,290			6,532					

Notes: Entries in the table are estimates of the Black-White test score gap, controlling for the parsimonious set of regressors. Columns 3, 6, and 9 represent the estimated change in the gap between kindergarten fall and first grade spring. The first three columns include all students. The remaining columns restrict the data set to schools that had students of different races included in the ECLS-K sample. The final three columns include school-fixed effects. Estimation is done using weighted least squares, using sample weights provided in the data set.

dren in our sample. This set of students is relevant because only mixed-race schools provide useful variation to identify the racial test score gap when school-fixed effects are included.

Column 3 of the table shows the baseline results reflecting the fact that Blacks are losing ground in the full sample (−.156 standard deviations relative to Whites in math, −.188 standard deviations in reading). When we eliminate students attending all-white schools from the sample, but otherwise estimate identical specifications, the results are not greatly affected (nor are they affected by eliminating students attending all black schools). Blacks continue to lose substantial ground by the end of first grade. When school-fixed effects are included in the regression (columns 7–9), the black-white test-score gap is identified off of differences between Blacks and Whites attending the same school. The estimates of ground lost by Blacks shrinks to less than one-third of the magnitude in the full sample, and is not statistically different from zero in these specifications.[22]

These findings are consistent with—but not definitive proof of—the argument that systematic differences in school quality for Blacks and Whites may explain the divergence in test scores. An alternative explanation is that Whites who choose to attend schools with Blacks are systematically worse than other Whites. Note, however, that a comparison of columns 1 and 4 show that in the fall of kindergarten black students actually fare somewhat worse relative to Whites who attend schools with Blacks then they do with the full sample of Whites. This finding suggests that the Whites who go to school with Blacks (controlling for observables) actually achieve at a slightly higher level than do those who attend all-white schools, which is consistent with previous research. Moreover, comparing columns 4 and 7, in kindergarten fall, Blacks do even worse relative to Whites attending the same school than they do compared to other Whites. Thus, a simple selection story in which low-achieving Whites are more likely to go to school with Blacks is not consistent with the data. On the other hand, we cannot rule out a priori the possibility that Whites who attend school with Blacks are on lower academic trajectories, despite the fact that they initially score better on tests than other Whites.

If Blacks attend worse schools than Whites on average, one might expect that this would be reflected in observable characteristics of the schools. Table 4.5 analyzes this issue. Each row of the table corresponds to a different measure of school quality. Column 1 presents means and standard deviations of each variable in the data, some of which are standard measures of school inputs (e.g., average class size, teacher education) and others that are nontraditional (e.g., measures of gang problems and loitering). Unfortunately, the nontraditional measures are subjective responses by the school principal, administrator, or other person in charge to questions of how seri-

Table 4.5. Differences across Races in Measurable School Inputs

School Input	*Mean of School Input*	*Coefficient on Race in Predicting Level of School Input:*			
		Black	*Hispanic*	*Asian*	*Other*
Average Class Size	20.673 (3.875)	.591 (.340)	.699 (.271)	.799 (.349)	−.259 (.343)
Teacher Has Master's Degree	.280 (.449)	.037 (.028)	.012 (.025)	−.001 (.032)	−.080 (.032)
Computer:Student Ratio	1.257 (2.050)	.003 (.156)	−.131 (.140)	.040 (.119)	.683 (.443)
Internet Hookup:Student Ratio	.344 (.627)	−.048 (.037)	−.032 (.038)	.020 (.035)	.377 (.186)
Percent of Students in School with Free Lunch	29.83 (27.98)	19.32 (2.64)	8.17 (2.00)	3.27 (2.08)	6.81 (2.78)
Gang Problems in School (1–3)	1.409 (.585)	.261 (.058)	.338 (.044)	.128 (.044)	.336 (.069)
Problems with Teacher Turnover (1–5)	1.811 (.943)	.263 (.083)	.227 (.064)	.062 (.078)	.132 (.092)
Litter Around School (0–3)	.741 (.759)	.492 (.065)	.369 (.053)	.240 (.063)	.412 (.087)
People Loitering Around School (0–3)	.524 (.747)	.497 (.079)	.331 (.064)	.171 (.063)	.368 (.088)
Receives PTA Funding	.733 (.442)	−.048 (.033)	−.050 (.026)	.000 (.029)	−.133 (.050)
Hall Pass Required	.425 (.494)	.194 (.037)	.100 (.034)	.010 (.041)	.059 (.046)

Notes: The values in the first column of the table are the means and standard deviations of the named school input. The entries in the remaining columns are estimated coefficients on race (with non-Hispanic Whites as the omitted categories) from regressions of the named school inputs on the race dummies and other covariates included in the parsimonious set of controls. The method of estimation is weighted least squares using sample weights provided by ECLS. The reported standard errors have been corrected to take into account within-school correlation in the school-level measures.

ous problems such as gangs are at the school. Consequently, these measures are likely to be of poor quality. Columns 2–5 report the race coefficients from regressions that are parallel to those elsewhere in the paper, except that school inputs are the dependent variable rather than test scores. Thus, the entries in columns 2–5 reflect the extent to which children of other races attend higher or lower quality schools on each of the measures, controlling for our parsimonious set of covariates. On traditional measures of school quality such as class size, teacher's education, computers in class, and Internet connections, differences between Blacks and Whites are small. On the other hand, the percentage of students eligible for free lunch, the degree of gang problems in school, the amount of loitering in front of the school by nonstudents, and the amount of litter around the schools are much higher for Blacks.

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 4.5 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 differences to the divergent trajectories of Blacks can be characterized as no more than suggestive. *Does the importance of parental/environmental inputs grow as children age?*

Black children tend to grow up in environments less conducive to high educational attainment. If the importance of parental/environmental inputs grows as children age, Black students would be expected to lose ground relative to Whites. The evidence in table 4.3, however, argues just the opposite. If that were true, than 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.[23] Also, the magnitude of the coefficients on socioeconomic status, age at kindergarten entry, and mother's age at first birth are smaller in the first-grade test-score regressions. That suggests that the relative importance of nonschool factors decreases over time, presumably because schools become a critical input into educational gains once children enter school.[24] Interestingly, the importance of school safety measures (e.g., gang problems, metal detectors, etc.) seem to become more important as children age.

Do Black Children Suffer Worse Summer Setbacks when School Is Not in Session?

Entwisle and Alexander and Heyns[25] have argued that black students lose more ground over the summer than white students as a consequence of worse home and neighborhood environments, and they gain ground over the school year while in school. If this were the explanation for the falling performance of Blacks, then public policies should be aimed not at schools, but rather, summer interventions. Our data provide a unique opportunity to test this hypothesis because a subset of the sample is tested both in the spring of kindergarten and in the fall of first grade, shortly after students return to class, allowing us to isolate the relative summer setbacks for Blacks and Whites. The results are reported in table 4.6. For the randomly chosen subset of the sample that is tested in the fall of first grade (about one-fourth of the students), we report at each point in time both the raw test score gap and the residual gap controlling for our parsimonious set of covariates. For the regression results, only the coefficient reflecting the black-white test score gap is shown in the table, and each entry in the table is from a separate regression. The test score gaps in the fall of kindergarten (column 1) and spring of first grade (column 4) for this subset of the sample are similar to those for the sample as a whole, suggesting that the subsample is indeed representative. Of greater interest is a comparison of the test scores in the spring of kindergarten versus the fall of first grade, since most of the intervening

Table 4.6. Do Black Students Suffer a Greater Summer Setback when School Is Not in Session? Estimates of the Black-White Test Score Gap for the Subset of the Sample Tested in Fall of First Grade (Values in the table are coefficients on the variable Black)

	Date test administered:			
Subject	*Fall kindergarten*	*Spring kindergarten*	*Fall first grade*	*Spring first grade*
Raw Gaps				
Math	−.601 (.040)	−.640 (.044)	−.631 (.045)	−.696 (.048)
Reading	−.376 (.042)	−.421 (.044)	−.390 (.043)	−.548 (.048)
With Controls				
Math	−.052 (.040)	−.097 (.044)	−.134 (.045)	−.236 (.052)
Reading	.142 (.043)	.054 (.045)	.071 (.044)	−.081 (.051)

Notes: Table entries are estimated Black-White test score gaps at different points in time for the subset of the sample that has all four test scores. Only a small fraction of the sample was tested in fall of first grade. The total number of observations in the subsample is 5,223. The top panel of the table reflects raw test score gaps; the bottom panel is the residual test score gap, controlling for the parsimonious set of control variables. The observations demarcated by the heavy border represent the tests given shortly before and shortly after summer break. Standard errors are in parentheses.

time was spent outside of school. On the raw scores, there is little difference before and after the summer break; to the extent there is any gap, it favors black students. With controls, black students lose slightly relative to Whites over the summer on math (the gap rises from -.097 to -.134), but the null hypothesis of no change cannot be rejected. The point estimates for reading show slight gains by black students relative to Whites over the summer. Thus, the empirical results lend little support to the hypothesis that differential summer setbacks explain the lost ground of black students in our sample. We do observe Blacks losing ground during the school year in both subjects in both years, in direct conflict with Entwisle and Alexander.

CONCLUSION

Previous efforts to explain the black-white test score gap have generally fallen short—a substantial residual remained for black students, even after controlling for a full set of available covariates. Using a new data set, we demonstrate that among entering kindergartners, the black-white gap in test scores can be essentially eliminated by controlling for just a small number of observable characteristics of the children and their environment. Once students enter school, the gap between white and black children grows, even conditional on observable factors. We test a number of possible explanations for why Blacks lose ground. We speculate that Blacks are losing ground relative to Whites because they attend lower quality schools, though we recognize that we have not provided definitive proof. This is the only hypothesis that receives any empirical support. To convincingly test this hypothesis, we need more detailed data on schools, neighborhoods, and the general environment kids grow up in.

Compared to previous studies, our results provide reason for optimism. Research on earlier cohorts of children found much greater black-white test score gaps, both in the raw scores and controlling for observables. When we attempt to mimic the nonrandom sample frames in earlier research (for example only looking at low birth-weight babies as in IHDP), we continue to find much smaller gaps in our sample. One plausible explanation for the differences between the current sample and cohorts attending kindergarten ten to thirty years ago is that the current cohort of Blacks has made real gains relative to Whites. Recent cohorts show smaller black-white gaps in the raw data, across multiple data sets, which gives us reason for optimism.

DATA APPENDIX

The Early Childhood Longitudinal Study Kindergarten Cohort (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 recoded some of the ECLS variables used in our analysis.

Socioeconomic Composite Measure

The socioeconomic scale variable (SES) 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

Parents 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.

Average Class Size

We computed each child's average class size over their kindergarten year by adding their class size in the fall and spring and dividing by two.

Teacher Has Master's Degree

We coded a dummy variable equal to one if the child's teacher has a master's degree or above.

Computer-Student Ratio

The number of computers in each school and the total enrollment of each kindergarten program is provided by the ECLS based on a survey given to each school. We divided the number of computers in each school by the total enrollment in kindergarten to produce this ratio.

Internet Hook-Up Student Ratio

This was constructed similar to the Computer:Student ratio, except the numerator consists of Internet/LAN connections in the school.

Percent of Students in Child's School Available for Free Lunch

Schools provided the percent of students in their school who were eligible for free lunch.

Gang Problems

Schools were asked: ''How much of a problem are gangs in the neighborhood where the school is

located?'' We coded this variable so that 1 implies ''no problem,'' 2 implies ''somewhat of a problem,'' and 3 implies ''big problem.''

Teacher Turnover

Schools were asked how much they agreed with the statement ''teacher turnover is a problem in this school.'' Answers range from 0 to 5, 0 indicating they strongly disagree and 5 indicating they strongly agree.

Litter around School

The ECLS interviewer was asked to report the amount of litter around each school. The variable ranges from 0 to 3. 0 indicates no litter and 3 indicates ''a lot.''

People Loitering around School

The ECLS interviewer was asked to report the amount of loitering by nonstudents around the school. The variable ranges from 0 to 3, 0 indicating no loitering and 3 indicating "a lot."

PTA Funding

Schools reported whether or not they receive supplemental funding from their PTA. We recoded this variable so that 1 implies yes and 0 implies no.

Hall Pass Required

Schools were asked: "Are hall passes required to ensure the safety of the children in your school?" This variable is coded 1 if yes and 0 if no.

NOTES

This chapter originally appeared as "Understanding the Black-White Test Score Gap in the First Two Years of School," *The Review of Economics and Statistics* 86, no. 2 (May 2004):465–480. We are grateful to Josh Angrist, Janet Currie, Michael Greenstone, Christopher Jenks, Alan Krueger, James Heckman, Susan Mayer, Derek Neal, Meredith Phillips, Barbara Schneider, and two anonymous referees for helpful comments and suggestions. Financial support was provided by the National Science Foundation (Fryer and Levitt). Correspondence can be addressed either to Roland Fryer, Jr., American Bar Foundation, 750 N. Lake Shore Drive, Chicago, IL 60611, or to Steven Levitt, Department of Economics 1126 E. 59th Street, University of Chicago, Chicago, IL 60637. E-mail: roland@uchicago.edu, slevitt@midway.uchicago.edu.

1. See W. Baughman and W. Dahlstrom, *Negro and White Children: A Psychological Study in the Rural South* (New York: Academic Press, 1968); B. A. Braken, E. Sabers, and W. Insko, "Performance of Black and White Children on the Bracken Basic Concept Scale," *Psychology in Schools* 24, no. 1 (1987):22–27; Jeanne Brooks-Gunn et al., "Do Neighborhoods Influence Child and Adolescent Development?" *American Journal of Sociology* 99, no. 2 (1993):353–395; Jeanne Brooks-Gunn, Greg J. Duncan, and Pamela Klebanov, "Economic Deprivation and Early-Childhood Development," *Child Development* 65, no. 2 (1994):296–318 and "Ethnic Differences in Children's Intelligence Test Scores: Role of Economic Deprivation, Home Environment and Maternal Characteristics," *Child Development* 67 (1996):396–408; James Coleman et al., *Equality of Educational Opportunity* (Washington DC: U.S. Government Printing Office, 1966); Richard J. Coley, "An Uneven Start: Indicators of Inequality in School Readiness," *Educational Testing Service Report*, Princeton, NJ, March 2002; Richard J. Hernstein and Charles Murray, *The Bell Curve: Intelligence and Class Structure in American Life* (New York: The Free Press, 1994); Lloyd Humphreys, "Trends in Levels of Academic Achievement of Blacks and Other

Minorities," *Intelligence* 12 (1988):231–260; Arthur Jensen, "How Much Can We Boost IQ and Scholastic Achievement?"*Harvard Educational Review* 39 (1969):1–123 and *Educability and Group Differences* (New York: The Free Press, 1973); A. Kaufman and N. Kaufman, *K-ABC: Kaufman Assessment Battery for Children* (Circle Pines, MN: American Guidance Services, 1983); E. Krohn and R. Lamp, "Current Validity of the Stanford-Binet Fourth Edition and K-ABC for Head Start Children," *Journal of Psychology* 27 (1989):59–67; J. Naglieri, "WISC-R and K-ABC Comparison for Matched Samples of Black and White Children," *Journal of Social Psychology* 24 (1986):81–88; Meredith Phillips et al., "Family Background, Parenting Practices, and the Black-White Test Score Gap," in *The Black-White Test Score Gap*, Christopher Jencks and Meredith Phillips, eds. (Washington, DC: Brookings Institution Press, 1998), pp. 103–145; Meredith Phillips, "Understanding Ethnic Differences in Academic Achievement: Empirical Lessons from National Data," in *Analytic Issues in the Assessment of Student Achievement*, David Grissmer and Michael Ross, eds. (Washington DC: U.S. Department of Education, National Center for Education Statistics, 2000), pp. 103–132; and Sandra Scarr, *Race, Social Class and Individual Differences in I. Q.* (Hillsdale, NJ: Lawrence Erlbaum Associates, 1981).

2. See Derek Neal and William R. Johnson, "The Role of Pre-Market Factors in Black-White Wage Differences," *Journal of Political Economy* 104 (1996):869–895 and June O'Neill, "The Role of Human Capital in Earnings Differences between Black and White Men," *Journal of Economic Perspectives* 4, no. 4 (1990):25–46.

3. To this effect, Jenks and Phillips write: "Reducing the black-white test score gap would do more to promote racial equality than any other strategy that commands broad political support."

4. See Hernstein and Murray, *The Bell Curve*; Jensen, *Educability and Group Differences*; and Arthur Jensen, *The G Factor: The Science of Mental Ability* (Westport, CT: Greenwood Publishing Group, 1998); Greg Armor, "Why Is Black Educational Achievement Rising?" *Public Interest* (September 1992):65–80; Jeanne Brooks-Gunn and Greg J. Duncan, eds., *The Consequences of Growing Up Poor* (New York: Russell Sage, 1997); Susan E. Mayer, *What Money Can't Buy: Family Income and Children's Life Chances* (Cambridge, MA: Harvard University Press, 1997); Phillips et al. 1998; Michael Cook and William Evans, "Families or Schools? Explaining the Convergence in White and Black Academic Performance," *Journal of Labor Economics* 18, no. 4 (2000):729–754; Lisa Delpit, *Other Peoples Children: Cultural Conflict in the Classroom* (New York: The New Press, 1995); Ronald F. Ferguson, "Teachers' Perceptions and Expectations and the Black-White Test Score Gap," in *The Black-White Test Score Gap*, Jencks and Phillips, eds., 273–317; William Rodgers and William Spriggs, "What Does AFQT Really Measure: Race, Wages, Schooling and the AFQT Score," *The Review of Black Political Economy* 24, no. 4 (1996):13–46; Phillip Cook and Jens Ludwig, "The Burden of 'Acting White': Do Black Adolescents Disparage Academic Achievement?" in *The Black-White Test Score Gap*, Jencks and Phillips, eds., 375–400; Signithia Fordham and John Ogbu, "Black Students' School Successes: Coping with the Burden of Acting White," *The Urban Review* 18, no. 3 (1986):176–206; Roland Fryer, "An Economic Approach to Cultural Capital," 2002, Working Paper (Chicago: University of Chicago Press); C. Steele and J. Aronson, "Stereotype Threat and the Test Performance of Academically Successful African Americans," in *The Black-White Test Score Gap*, Jencks and Phillips, eds., 401–430.

5. On a test of general knowledge, a racial test-score gap persists. On a subjective teacher assessment of general knowledge, however, there is no difference between Blacks and Whites in fall of kindergarten.

6. Neither Hispanics nor Asians experience this widening test score gap over time. Indeed, Hispanic children systematically close the gap relative to Whites, presumably because their initial scores are artificially low as a consequence of limited English proficiency among some Hispanic parents.

7. Lyle V. Jones, Nancy Burton, and Ernest Davenport, *Mathematics Achievement Levels of Black and White Youth* (Chapel Hill: University of North Carolina, L.L. Thurstone Psychometric Laboratory, 1982); Phillips, ''Understanding Ethnic Differences''; Meredith Phillips, James Crouse, and John Ralph, ''Does the Black-White Test Score Gap Widen after Children Enter School?'' in *The Black-White Test Score Gap*, Jencks and Phillips, eds., 229–272.

8. This pattern is also consistent with self-selection of low-achieving Whites into schools attended by Blacks. Casting doubt on this alternative explanation is the fact that Whites who go to school with Blacks have baseline test scores upon entering kindergarten that are similar to those who are in all-white classes (Humphreys, ''Trends in Levels of Academic Achievement,'' documents a similar finding among high school students). When we eliminate from the sample Whites who have no black children in their class (more than 60 percent of all white children fall into this category), we obtain similar results.

9. In particular, Hernstein and Murray's controversial book, *The Bell Curve*, published in 1994, ignited interest in the subject by arguing that genetic differences are the primary explanation for the differences between Blacks and Whites in achievement test scores. For excellent summaries of the book, see James J. Heckman, ''Lessons from The Bell Curve,'' *Journal of Political Economy* 103, no. 5 (1995):1091–1120 and Arthur Goldberg and Charles Manski, ''Review Article: The Bell Curve,'' *Journal of Economic Literature* 33, no. 2 (1995):762–776. Examples of the discussion that emerged include Bernie Devlin, Daniel Resnick, and Kathryn Roeder, *Intelligence, Genes, and Success: Scientists Respond to the Bell Curve* (New York: Copernicus Books, 1998); Steven Fraser, *The Bell Curve Wars: Race Intelligence, and the Future of America* (New York: Basic Books, 1995); and Joe Kincheloe, Shirley Steinberg, and Aaron Gresson, *Measured Lies: The Bell Curve Reexamined* (New York: St. Martins Press, 1997).

10. See Brooks-Gunn and Duncan, *The Consequences of Growing Up Poor*; Mayer, *What Money Can't Buy*; Jeanne Brooks-Gunn, P. K. Klebanov, and Greg J. Duncan, ''Ethnic Differences in Children's Intelligence Test Scores: Role of Economic Deprivation, Home Environment, and Maternal Characteristics,'' *Child Development* 67, no. 2 (1995):396–408; Jeanne Brooks-Gunn and Greg J. Duncan, ''Family Poverty, Welfare Reform and Child Development,'' *Child Development* 71, no. 1 (2000):188–196.

11. David Grissmer, Ann Flanagan, and Stephanie Williamson, ''Why Did the Black-White Score Gap Narrow in the 1970's and 1980's?'' in *The Black-White Test Score Gap*, Jencks and Phillips, eds., 182–228; Larry Hedges and Amy Nowell, ''Black-White Test Score Convergence since 1965,'' in *The Black-White Test Score Gap*, Jencks and Phillips, eds., 149–181; Humphreys, ''Trends in Levels of Academic Achievement.''

12. 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.

13. 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 percent of the sample lies above this cutoff point.

14. Roland Fryer and Steven Levitt, "The Black-White Test Score Gap in the First Two Years of School," *The Review of Economics and Statistics* 86 (2004):447–464, test a more exhaustive set of possibilities.

15. 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 percent of students whose predicted probability of being black is less than 10 percent or greater than 90 percent, the black-white gap on math rises slightly and the reading gap becomes closer to zero.

16. The exceptions we are aware of in which the black-white test score gap has been made to disappear are **Crane (1998)[[PLEASE SUPPLY ALL INFO FOR THIS]]**; Kai Li and Dale J. Poirier, "The Roles of Birth Inputs and Outputs in Predicting Health, Behavior, and Test Scores at Age Five or Six," **2002[[OR 2001. WHICH IS CORRECT??**),Working Paper, **[[LOCATION??]]**; and Pedro Carneiro and James Heckman, "Human Capital Policy," 2002, Working Paper, The University of Chicago. Li and Poirier, "The Roles of Birth Inputs and Outputs," using a Bayesian structural model, find no systematic differences between Blacks and Whites using the NLSY. Hernstein and Murray, *The Bell Curve*, and Meredith Phillips, James Crouse, and John Ralph, "Does the Black-White Test Score Gap Widen after Children Enter School?" in *The Black-White Test Score Gap*, Jencks and Phillips, eds., 229–272, using different methods on the same data, find large gaps still persists. Using CNLSY, **Crane (1998)[[NOT IN REFS]]** and Carneiro and Heckman, "Human Capital Policy," find that on some tests, racial gaps disappear with controls, although large gaps remain on other tests designed to capture similar sets of skills.

It is important to note that on the test of general knowledge in ECLS, the black-white gap does not fully disappear. Black students test almost one full standard deviation behind Whites in a raw comparison of means. That gap falls to .3 when controls are included. On the subjective teacher assessments, the raw gap in general knowledge between Blacks and Whites is much smaller (.25 standard deviations) and does shrink almost to zero with the inclusion of controls.

17. Our results are also unchanged when we limit our ECLS sample to low birthweight babies, who are oversampled in IHDP, another data set analyzed by Phillips, Crouse, and Ralph, "Does the Black-White Test Score Gap Widen?"

18. Similar results (not shown in the table) are obtained when we include the full set of nearly 100 covariates. In those specifications, black students lose .136 standard deviations on math and .109 standard deviations on reading. Including the fall kindergarten test score as a covariate predicting the spring first grade test score also has little impact on the results: black students lose .192 (.140) standard deviations in math (reading).

19. Fryer and Levitt, "The Black-White Test Score Gap," test a more exhaustive set of possibilities.

20. Black students may attend these schools, but just not be in the classrooms sampled.

21. Because elementary school students attend schools close to home, there is no way for us to distinguish between the impact of neighborhood and school quality in our data set. Note, however, that we are able to explain racial gaps upon entry to school without using controls for the neighborhood environment. For neighborhoods rather than schools to explain the racial divergence in test scores, the quality of the neighborhood would need to have a large impact on test scores after entry into school, but not before.

22. This finding in some ways parallels the findings in Janet Currie and Duncan Thomas, "School Quality and the Longer Term Effects of Head Start." *Journal of Human Resources* 35, no. 4 (2000):755–774, that early gains for students who attend Head Start tend to disappear due to low-quality schools that these students later attend. Consistent with Currie and Thomas, 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 Alan Krueger and Diane Whitmore, "Would Smaller Classes Help Close the Black White Achievement Gap?" 2001, Working Paper #451 Industrial Relations Section, Princeton University and Phillips, Crouse, and Ralph, "Does the Black-White Test Score Gap Widen?," who find that the black-white gap widens as a result of poorer quality schools.

23. Indeed, from a theoretical perspective, one might expect that the opposite hypothesis would hold true: the importance of parental inputs declines with age. Prior to 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.

24. An alternative explanation for the shrinking coefficient on the SES variable is that socioeconomic status varies over time. Therefore, using the kindergarten value of the SES variable in the first grade regression induces measurement error. That explanation cannot explain the declining coefficients on age at school entry and mother's age at birth. Moreover, for other variables that are time varying, like number of books and WIC participation, the coefficients do not shrink in the first-grade regression.

25. Doris Entwisle and Karl Alexander, "Summer Setback: Race, Poverty, School Composition, and Mathematics Achievement in the First Two Years of School," *American Sociological Review* 57 (1992):72–84 and "Winter Setback: The Racial Composition of Schools and Learning to Read," *American Sociological Review* 59 (1994):446–460; Barbara Heyns, *Summer Learning and the Effects of Schooling* (New York: Academic Press, 1978).